GEOMATICS WORKBOOKS 12

FOSS4G EUROPE COMO 2015

Free and Open Source Software for Geospatial

Open Innovation for Europe

EDITORS:
Maria Antonia **Brovelli**
Marco **Minghini**
Marco **Negretti**

ISSN 1591-092X

Geomatics Workbooks, volume 12

FOSS4G Europe, Como 2015

Geomatics Workbooks 12 - July 2015
Editor: Maria Antonia Brovelli, Marco Minghini, Marco Negretti
Cover: Silvia Guerini
2015-07-21

Geomatics Workbooks
ISSN 1591-092X
http://geomatica.como.polimi.it/workbooks/

Laboratorio di Geomatica
Politecnico di Milano - Polo Territoriale di Como
via Valleggio 11, 22100 Como

Index

Open data track

Water track

General track

Free and Open Source Software for Geospatial

Arnulf Christl[1,2]

[1] metaspatial

[2] OSGeo

1 The FOSS4G Conference Series

FOSS4G is the premium Open Source Geospatial conference event of the Open Source Geospatial Foundation (OSGeo). The first edition under the auspices of OSGeo dates back to 2006 where it took place in Lausanne, Switzerland.
Before this the GRASS users organized a forerunner named the FOSS/GRASS User Conference in Bangkok, Thailand in 2004. The very first international Conference was the Open Source GIS - GRASS users conference in 2002 in Trento http://www.ing.unitn.it/~grass/. And in the year 2000, 15 years ago, the "Primo Meeting degli Utenti Italiani di GRASS" took place in Como, Italy (http://geomatica.como.polimi.it/grass_2000/). And now we are back – a long long tradition of Open Source in Como!
The first issues of FOSS4G were also dubbed "Meeting of the Tribes" alluding to the different programming languages (C++, Java, JavaScript, PHP, etc.) and the fact that OSGeo managed to provide a neutral ground and have them all talk to each other. As Free and Open Source software became more and more commonly used, the need for broader conferences for common users also grew. Initially the annual FOSS4G conference mostly attracted core developers of the main OSGeo projects. Over time it grew to quite a considerable size and at the next edition in Seoul, South Korea we expect 1000+ attendees. These numbers also mark a size that requires increasingly larger venues and professional logistics which result in a considerably higher price tag. The venue for FOSS4G rotates around the world starting with Lausanne, Switzerland in 2006, followed by Victoria in Canada, Cape Town in South Africa, Sydney in Australia, Barcelona in Spain, Denver in the USA and Nottingham in the UK (2013). The last edition in 2014 took place in North America (Portland, Oregon, USA) and this year FOSS4G will take place in Seoul, South Korea. The 2016 edition of FOSS4G was awarded to the German OSGeo Local Chapter (FOSSGIS) which will organize it in Bonn, Germany. This means that attending a FOSS4G event is still a rare chance considering that our earth is still quite a sizable place, regardless of how well we are virtually connected.
Over time more and more OSGeo Local Chapters have started their own conferences, often in their local languages. Just one example between many is the German Local Chapter association which organizes the annual FOSSGIS conferences in German language since 2003. The need for smaller conferences could be answered by these local OSGeo chapters but there was still a gap between the global FOSS4G (commonly held in English language) and the local

conferences held in their native languages.

2 FOSS4G Europe

The first regional conference titled "FOSS4G Eastern Europe" took place in 2012 at the Faculty of Civil Engineering in the Czech Technical University in Prague. It catered for a broader audience than the original "meeting of the tribes" and also included more attendees from research and education. Especially for local and neighboring countries travel costs were much lower than a global FOSS4G taking place in Sydney, Victoria or Cape Town. This first conference was such a good success that in 2013 the Romanian Local Chapter hosted the FOSS4G Central and Eastern Europe at the National Library of Romania in Bucharest, Romania. In 2014 the conference was renamed FOSSG Europe and moved to Bremen in Germany at the Jacobs University, again with a broad audience from research, education and public authorities.

3 FOSS4G Europe 2015 in Como, Italy

This year FOSS4G Europe brings us to beautiful Como in Italy. The main tag line reads: "Open Innovation for Europe". The web site continues to state that "The Conference aims to bring together FOSS4G users and developers worldwide and foster closer interactions with and amongst European communities in order to share ideas for improving geodata, software and applications openness."

This makes this edition of FOSS4G a first to explicitly mention "geodata" before "software". Interesting. Maybe it shows that Open Source software has been commonly accepted in the geospatial domain and we can turn to more pressing issues?

4 Open Data

Back in 2010 Schuyler Erle wrote (http://de.slideshare.net/sderle/how-crowdsourcing-changed-disaster-relief-forever):

> Your software is awesome,
>
> (but) your software is useless
>
> ...without data

Since then many things have improved but we are still far away from a generally accepted Open Data policy. Especially in Europe the INSPIRE regulations have helped pave the way towards a more open policy but it is a top down effort which still lacks support from those who actually have the data. There are great exceptions, for example swisstopo (http://map.geo.admin.ch from Switzerland (who are not even bound by the INSPIRE regulations). Austria maintains http://basemap.at and opens up topographic maps down to a scale of 1:1000 including all street addresses and even plot numbers. Finland has opened up most of the geospatial data and even in Germany at least the federal level of data is now available as map services. Somehow not yet really Open Data, but we appreciate the effort anyway.

All of this can and should be improved. To achieve this we need to complement these top down efforts with a bottom up approach. Many smaller initiatives already go a lot further than national programs and some small local administration make their geospatial data wide and openly available for anybody to use. Hand in hand with this growing sense of distributing geospatial data freely is a growing tendency to recognize OpenStreetMap as a valuable addition to national geodata infrastructure instead of perceiving it as the natural enemy of the surveying trade. One nice example is the state of Bavaria (Germany) where the authoritative map offering includes a layer named "Biergarten" (http://geoportal.bayern.de/bayernatlas/). The locations have been taken from OpenStreetMap and are also updated on a regular basis. This makes sure that the locations of the "beer gardens", a cultural asset of Bavaria, are up to date and get maintained by the people who actually go to these places to drink beer. Not something you would expect from the officially endorsed surveyor. Well, who knows, Bavaria is known for its open approach to beer.

Whatever else, Open Data is surely the way to go. This is a fact and this edition of FOSS4G-Europe will show many examples of how this can be done, what needs to be taken into account when starting off and how both the public and the economy can profit from it.

Open Source Software

This shift in priorities towards Open Data in the FOSS4G Europe Program also reflects on the maturity of Open Source geospatial software in general. OSGeo software, libraries and tools are nowadays commonly used in many public authorities. Be it QGIS or gvSIG on the desktop, MapServer, GeoServer and deegree on the dynamic serving end, the obligatory Postgres database with PostGIS in the basement and a plethora of Web clients and client frameworks including OpenLayers, MapBox, Leaflet and many more.

This basic architecture now also scales by adding caches (TileCache) and proxies (MapProxy), NoSQL databases (MongoDB, CouchDB) and so many more good things that this introduction cannot make justice to even a small number of them. With the uptake of Big Data we see Hadoop introduced and spatial indexes are created using Solr. This shows that the artificial gap between geospatial and "other" software is slowly closing. Geospatial is not special anymore. Well, it is still spatial and probably also special but people recognize that both "standard IT" and geospatial software have to work together to address current needs.

Application Openness

Honestly, this one should get you thinking, even if you are a long time Open Source developer. Why should FOSS4G Europe explicitly mention "Application Openness "? What is that anyway? My take on asking for more "application openness" is to reduce the barriers implemented in "Apps" (archaic: "software") and to make access to the maps and underlying data more open. Many applications these days are built with Open Source software but access is still restricted to named users and the data is purposefully obfuscated. The crux of this problem is that Apps as we know them from our beloved mobile devices

follow a newish business model which creates a new vendor-lock-in. Why that? Because vendors of the Apps cannot live off "selling" software usage licenses (an outdated business model from the late nineties). There is also no sizable servicing or maintenance business and nobody needs trainings or consulting for using an App. If an App does not work we stop using it. A regular 24/7/364 Service Level Agreement is bluntly expected – for no extra fees.
So where is the money in App development? It is our private data, the new currency of the Web. We (apparently) get to use services and Apps "for free" (as in free beer). But we do pay. We pay with our location tracks, information about friends or what we search on the Web. We let others read all our private emails, address books, calendars, event listings and even our private photo album. Wow! Unheard of?
Most people righteously rebound in outrage whenever any national secret service has been found guilty of peeking into our private lives. But if Facebook wants to know all details about our friends and Google wants to read every single email we send – then it is sort of OK. At least acceptable. This is really weird but a fact. In Great Britain people will rather want to see my Credit Card as proof of identity than my official state issued German identity card or passport. Come on?! This is a German identity card and it cost a billion to develop... But we are digressing.
Or maybe we are not? With the erosion of privacy in the personal continuum why are we making such a fuss about not being able to publish geospatial data openly?

GeoForAll – Education

One of the most important aspects in the Free and Open World was and still is is education. Many curricula are still focused on old tools or are limited to one specific brand. OSGeo and ICA/ACI (the International Cartographic Association) have partnered to improve this by starting the GeoForAll Initiative – to promote and enhance education, research and service activities in the area of Open Geospatial Science & Applications – all over the world. http://www.geoforall.org.
FOSS4G Europe will mark a milestone for the GeoForAll Initiative because it will bring together a wide range of interested parties who want to contribute and profit from GeoForAll. There will be meetings and sessions focused entirely on developing this initiative and to explore ways in which to spread word and competence in teaching geospatial tools. GeoForAll has been started as an extensive outreach program to consistently improve education all around geospatial topics.

The Tracks

FOSS4G Europe has one general track and additional five themes including the Academic session, Library track, Open Data track (see above), Positioning and Water tracks.

The Academic Session

The Academic Session focuses on original research contributions on all aspects of open-source geospatial software and its application. Submissions focusing on

INSPIRE, Big Data and Societal Challenges were particularly encouraged. Papers were selected in a blind review process based on their merits in scientific novelty, relation to the state of the art, presentation of the paper and overall quality. Authors have received feedback and suggestions for improvement to their paper before submission of the final camera ready version of the paper. Selected papers will be published online in the Geomatics Workbooks (ISSN 1591-092X), the conference journal of GEO Laboratory at Politecnico di Milano, Como Campus. Outstanding contributions will be further invited to submit an extended version to high-impact journals.

The Libraries Track

Libraries are evolving from repositories for printed information to engaged community centers which provide places and services for discovery in science and technical information. Open source geospatial software has a great potential for advanced location-enabled library services. The expertise of librarians on the other hand can provide significant benefits to access and preserve the wealth of knowledge, software and data generated by the OSGeo project communities. Metadata and data blur and intermingle depending on the perspective and questions asked. This track focuses on sharing experience between librarians and GIS experts to encourage collaboration and discussion of new and emerging concepts in the library domain.

The Positioning Track

The objective of this track is to bring together researchers, developers and practitioners to present and discuss new advances in research, development and application in the area of Indoor and outdoor positioning, mapping and location based services.

Precise positioning by low cost sensors is a big research topic. Over the past years, increasingly accurate indoor localization technologies, based on new technologies are fundamental for many applications. This includes social efforts (support for impaired people), commercial applications (fleet control and guidance), personal Apps (sport and outdoor activities, social networks), indoor safety & security and many more.

Topics of interest include indoor maps, localization techniques, routing and navigation but also GNSS positioning for navigation, integration of sensors for positioning and georeferencing, mobile client-side technologies and applications, system architecture, experiences and best practice.

The Open Data Track

The scope of the Open Data Track is to provide insights on success stories, best practices and guidelines in releasing public geodata, fostering community driven geodata and developing and maintaining products, services and apps based on open geodata. The Open Data Track will reflect on the growth of Open Data initiatives in the geospatial realm and highlight advantages, controversies and questions it brings forward.

The track will showcase the state-of-the-art of Open Data and how Open Source Geospatial software is a perfect technological fit. Presentations will explore significant implications of open geodata and present best practices and guidelines for releasing and using these data within the FOSS4G community and beyond.

Summary and Thanks

The submissions for the 2015 edition of FOSS4G Europe in Como, Italy show a growing sense on the importance of Open Data for everything geospatial. Proven and working software and architectures are presented side by side with completely new approaches coming out of research. Diversity wins.

The FOSS4G Europe Organizing Committee has done a great job in providing a broad approach to everything open in the geospatial domain. Many thanks to the Organizing Committee and Conference Committees for the many hours of tedious work that went into making this conference and proceedings a great success. A special thanks for her tireless work goes to our conference chair Maria Brovelli, without her we would not have this wonderful conference. And another special thanks goes to our conference secretary Marco Minghini for putting up with so many open ideas and still managing to keep everything together.

We hope that you enjoyed the conference and profit form the proceedings and welcome you to all the FOSS4G successors in coming years. In case you did not have a chance to come in person why not consider to Roll your own FOSS4G? If you are interested you should first subscribe to the OSGeo Mailing Lists and find out whether there are others interested in joining you – either on a specific theme, a software, a region or a common language. The OSGeo Wiki also provides a cookbook (http://wiki.osgeo.org/wiki/FOSS4G_Cookbook) which highlights what you have to take into account if you want to start your own series of events. Or maybe you just plan to have a one-off meeting on a specific topic...

Whatever your take on openness, keep in mind that diversity wins.

Have fun,

Arnulf Christl

(OSGeo President Emeritus)

LIMEWISE: A deterministic line-based interpolation methodology for a realistic multimodal accessibility representation

Palmier Patrick

CEREMA – Nord-Picardie territorial division

Abstract

The representation of territorial accessibility need to combine two technical fields, transport modeling in order to calculate realistic and accurate multimodal travel times, and advanced GIS features to visualize the results through isochrones or other geographic indicators. Unfortunately results are too often produced by either transport or GIS specialists, ignoring each other. This paper proposes an original methodology, using free and geo open-source software to produce accurate indicators and realistic maps combining advanced techniques of both fields, in order to analyze territories accessibility, especially at a large scale. The article focuses on an implementation of a deterministic interpolation based on linear objects taking into account access times varying along polylines objects, and other criteria as impassibility, possibility of spread from linear objects, taking advantage of the open source geographic software Qgis to implement and validate the methodology through a plugin called "networks" available on Qgis repository.

Keywords

accessibility, multimodal, isochrone, interpolation, linear, transport

1 Territorial accessibility indicators

Generally, territorial accessibility maps based on isochrones are often produced to show accessibility inequalities. This mode of representation is attractive, especially for a non specialist public. Isochrones enable users to produce accessibility indicators, such as the number of inhabitants or jobs within a duration period a time, or based on gravity models (Hansen , 1959).

Transport modeling practitioners, generally used specific modeling tools from proprietary software which are designed to do efficient calculations, based on shortest path algorithms taking into account transport demand, and where the generation of isochrones is, when it exists, an optional layer, often based on simplistic rules for interpolation. In parallel, in most GIS software, isochrones are often produced with their own shortest path algorithms. Often they don't take into account the complexity of transport problem, especially inter-modality combining individual modes defined by arc-based travel times and timetable based-modes. Interpolation for isochrones generation, is in general based on kriging from travel times associated to point objects, located at junctions. But kriging is designed to interpolate results from point objects where the spatial distribution of the variable we want to interpolate is not known "a priori". Kriging is then not well adapted to travel times, which are varying along the itineraries, and where time values outside lines don't fill with kriging variograms, but can be estimated with a deterministic approach

2 The multimodal issues

2.1 Road issues

Road accessibility calculation can be divided in two groups:

- Transport modeling group where accessibility is based on traffic assignment results. This approach needs a trip demand matrix in order to take into account road congestion defined by volume-delays functions on links. As traffic assignment is a converging iterative process, generally to reach Wardrop equilibrium (Wardrop, 1952), the road network taken into account isn't exhaustive to be adapted to the demand zones and for calculation speed. Transport modeling software are proprietary software where pricing is depending on the number of zones or links, which also limit the size of the network. Thus, if travel times at equilibrium reflect correctly real travel time conditions, this method can't estimate times for local network, which could be problematic for isochrones visualization at large scale. Static models are also based on peak hour to take into account road congestion, which is not adapted for long trips exceeding 1 hour, where only part of the trip is occurring at peak hour.
- Road accessibility based on GIS calculations are generally based on exhaustive with constant travel times on links defined by average speed on links depending on road and environment characteristics.

If, in order to calculate a particular itinerary between an origin and a destination, as in road navigator, A* like algorithms are adapted and powerful, Dijkstra or Graph Growth Algorithm like algorithms are more suitable for 1 to all destinations calculations as need for isochrones.

2.2 Timetable based networks

Public transport accessibility could be computed either on frequency based networks or timetable-based networks. Transport models are currently dealing with both approaches. Frequency based algorithms , as implemented in most modeling software, give an estimation of average travel times during a period, often based on strategies (Spiess & Florian , 1989). Timetable based algorithms give exact travel times or costs, but for starting or arriving at a precise time. These shortest path algorithms are different from road ones, and are subject currently of much research, in order to optimize public transport routing search engines (Delling, Pajor & Werneck , 2014) and (Dibbelt, Pajor, Strasser, & Wagner, 2013). Nevertheless, certain algorithms lose a significant part of their performance if they had to deal both timetable and road based networks.

2.3 Multimodal networks

Today, concerns related to inter-modality and modal shift, require accessibility tools that combine different networks, either road networks (walking, cycling, driving) , timetable based ones (public transport, airplane), and hybrid ones as bike or car share services for example. These networks are built combining several networks connected in inter-modality points. Shortest path algorithms need to deal with characteristics of road and time table based algorithms. For planning purpose, and desktop tools don't need to have very quick algorithms as multimodal itineraries API need, but to be able to carry out complex studies with specific and multimodal options in order to evaluate complex public policies. Such algorithms as developed by IFSTTAR (Bousquet, 2010) in Tempus project or Musliw (Palmier, 2010) respond to these requirements. Multi-objective algorithms are designed to propose users several optimizations depending of their wishes, as implemented by Gräbener, Berro & Duthen (2010)

3 Multimodal calculations

3.1 The need of link based graphs calculations

For realistic accessibility maps visualizations, we need first powerful shortest paths algorithms which can manage several networks layers (road based and timetable based). In addition, the graph topology must be designed at a link level and not at a node level. It implied that the graph is composed of links with incoming and outgoing links, and not a succession of nodes connected with links. This prerequisite is essential to visualize access time, in particular at large scale, because access times are varying along links. It implies that the access time for a particular node is generally not unique. It depends on the node incomings link. This property allow to take into account in standard, specific turning movements, or transfer penalties. More complex algorithms are taking into account complete sub-path prohibitions or cost-addition as in Tempus (Bousquet), 2010. The accessibility calculator tool must provide detailed access times of general costs for every link in order to build accessibility maps.

3.2 The Musliw Tool

Musliw is a tool designed by CEREMA (Palmier, 2010) to compute multimodal itineraries calculation for planning studies. The network internal structure is based on nodes and links. Links are defined by a starting and ending node and a line identifier. Links could be either road links defined by a travel time which could be set for different time periods and calendar days, or either timetable links defined by a line identifier, and a set of services (days of services, starting time and ending time). The internal graph topology is link-based and produces complete statistics as detailed output as a result. These outputs concern generalized costs including in-vehicle times, individual modes travel times, waiting time and boarding times which can be set depending of a user-defined link attribute. The user can ask either for a date and a time of latest arrival or either a date and time of earliest departure, and can specify specific check-in delays for flight or Eurostar. The tool produces as outputs, generalized costs for every node and arc, detailed multimodal itineraries, link flows, transfer flows and trip summaries.

4 Isolines representation issues

4.1 Point-based interpolation

As access times are often only available at node level, the creation of isolines or iso-polygons has to be carried out with point objects. Many techniques are existing to predict a value at a given point by processing an interpolation. Several techniques are commonly used, like inverse distance weighting, triangulated irregular network, nearest neighbors or kriging. If these techniques are sufficient at small-scale with a high density of points, they are unadapted for analysis at large scale, especially for measuring the walking accessibility of a site. Even though, these techniques of interpolation are often used, because research in this area is very active and many interpolation algorithms are implemented in commercial and open-source GIS software and are easy to use. In addition, many of GIS users don't know the theoretical concepts of interpolation and kriging they are using and the set of algorithm parameters to be determined, so they use default ones. This implies that the produced maps could be wrong and it could be impossible to explain the represented access times.

4.2 Fixed grid or buffer around stations

Several other procedures are used to represent territorial accessibility, as generating a fixed grid and to extend the graph by creating virtual arcs from the center of each cell to the nearest arc orthogonally. This method gives a good representation of accessibility but has some disadvantages such as increasing the size of the network graph, being only adapted in a specific range of scales. A grid is necessarily associated to a networks graph that requires to build as many grid as different network graphs you have. Grids could be replaced by territorial zones or voronoï regions based on the node layer.

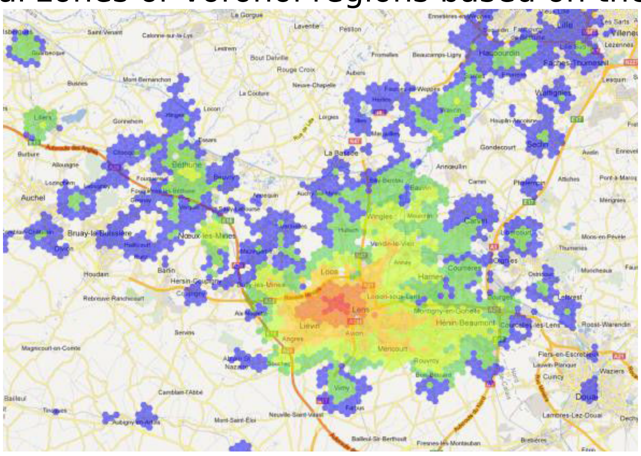


Figure 1: Example of PT and walking accessibility rendering with grid - Source CEREMA/DterNP – (Palmier 2014)

Instead of generating polygons, it could be sufficient in some cases to make a thematic map based on access times on nodes, represented by colored circles. Finally, we can mention the possibility of generating isolines concave hulls, which are subject of much research, but have the disadvantage of not being unique.

4.3 Line-based interpolation

Two main methods are used to create iso-lines from polylines objects. The first, which is the default method used by ArcGis called “polygon based networks buffer” (Frank, Schmid, Sallis, Chapman & Saelens, 2005) consists in building polygons from vertices which are exactly at a specific access distance of time from an origin. Generally, this implies that the shortest path calculations are made with the software algorithm, which doesn’t enable most of the time complex multimodal networks mixing timetable based and individual mode networks. This method is disputed because of the importance of the buffer type which influences significantly the precision of walkability measurement. Oliver, Schuurman & Hall (2007) prefer introduced the “buffered line-based network buffer” which corresponds to create a buffer of 50m wide around every road that can be reached in a predefined distance or time. But this method don’t perform very well to take into account data which is not point data, as density or public space areas. The limits of these different methods to build isolines for multimodal accessibility highlight the need of a method that leverage the best of, on one hand, the quality of multimodal calculation, and on the other hand, a specific interpolation method that answer completely the objectives of multimodal accessibility specifics.

5 The LIMEWISE method

5.1 Presentation

Limewise (Linear Interpolation Method With Impassability and Spread for Evaluation) method origin, was designed to define, in the mobility plan of Lille Metropole, high walking access areas around public transport stations (at 5 minutes and 10 minutes).

Basic principles of the Limewise method were implemented in ZAP maps in Lille (Palmier, 2001). The methodology has been constantly improved since 2001, to reach its current state of development, including the consideration of lateralized spread and impassability. Here is a comparative of three different methods (Chatalic 2012).

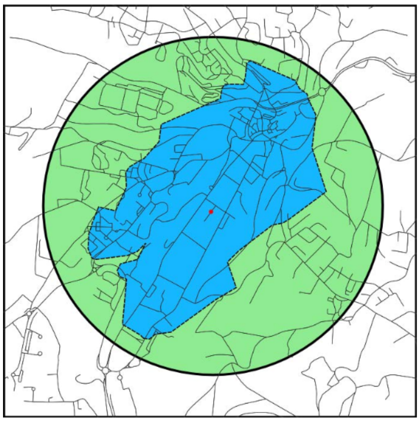


Figure 2: Buffer line-based network (Chatalic 20112

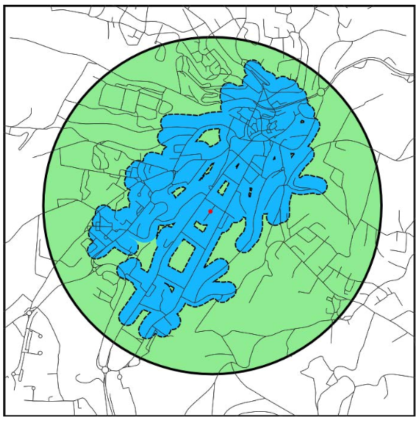


Figure 3: Buffer line-based network buffer (Chatalic 2012)

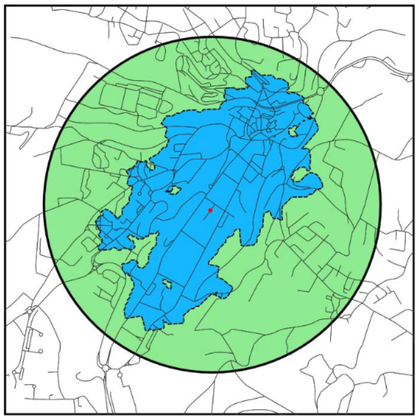


Figure 4: Zap method (Chatalic 2012)

5.2 The questioning of some assumptions

First, multimodal accessibility specificities encouraged to question some assumptions. Common methods supposed that accessibility is universal (defined everywhere), continuous, isotropic and permanent.

5.2.1 Universal

Multimodal accessibility is not defined everywhere, because some places are inaccessible, as rivers, lake, prohibited places, etc. This property should be taken into account in a realistic method. Tools have to take into account uncrossable areas of infrastructures.

5.2.2 Continuous

Mulimodal accessibility isn't continuous everywhere, because some roads are impassable. It implies that access time on one side of the road isn't the same as the one on the other side. For example, from either side of a motorway access time could be very different, if we suppose that we can't cross the lanes. In addition, even if you consider that where you're in a car you can park everywhere on the side of the road, it's no more true in highway, and you can't jump out of a bus, a metro or a train until it stops at a station. Public transport accessibility is by nature discontinued, as you can access or egress only at stops or stations. Tools had to take into account the possibility of spread from links (prohibited for motorway links or public transport link, allowed on both sides in general for walking, cycling and car with on-street parking links)

5.2.3 Isotropic

In parallel, multimodal accessibility isn't always isotropic. First, concerning public transport the frequencies and travel time may vary from a bus stop to another depending of the service direction. Even, for individual modes, when you're walking on the sidewalk of an avenue with many cars, you can spread on the same side of you sidewalk, but not on the other side, because you need to cross the road, and it's impossible except at junctions due to heavy traffic. Tools had to deal with anisotropy as travel times may be different from one side to another side of the road, especially when people can cross the road only at junctions.

5.2.4 Permanent

The accessibility of individual modes is permanent, that is it is defined all day-long and any day, even if travel time may vary during the day, as by car in peak hours. Public transport are discontinuous along the day, because people have to wait for the bus in order to board in. In addition, frequencies varies along the day, and often there is no service available at night between 1am and 4Am. The level of service also depends of the type of the day (school, period, holidays, sunday,...)

5.2.5 The need of an adapted methodology

The questions raised require a methodology that can deal with an oriented graph to take into account travel times which could be different depending of the flow direction. In addition, the methodology has to take into account impassability of links and the possibility to define if people can spread from the road either to the left, to right, in both directions or not.

5.3 Description of the method

The Limewise method is based on the key idea that in order to represent realistic multimodal accessibility, you need an efficient an adapted multimodal shortest path algorithm combined with a line-based interpolation algorithm which interact each other to take advantages of the best of both techniques.

5.3.1 Access times for any link at start and end nodes

The multimodal shortest path algorithm should provide access times for any link of the multimodal oriented graph. It means that each reachable link has an access time (or cost) from origin. This access time corresponds to the endpoint, if the query was to start from origin at a certain time, or it corresponds to the start point, if the query was to arrive at the destination a a certain. In both cases, you need to calculate the time at the other end, by subtracting the link travel time or cost.
This implies that access time at a node is not unique. They are as many access times at a node as incoming links.

5.3.2 Limewise necessary link attributes

In order to create realistic isochrones maps, you need to fill to each link three attributes

- Flow direction: This field is necessary to know if each link is one-way , two-ways, or prohibited

'0'	Prohibited link
'1'	One way link in the same direction of digitizing
'2'	One way link in the opposite direction of digitizing
'3"	Two way links

- Side of spread: This attribute indicates if when you are traveling in a certain link, you can spread in adjacent blocks, left or right, or not. For example, if you are circulating on a motorway you can't spread because you need to reach an interchange to exit. It's the same thing when you are traveling in a bus, a train or a metro, your need that the vehicle arrives at a stop to alight. If you're walking on a street with heavy traffic, you can spread only in one side, because you can't cross the street, except at a junction.

'0'	Spread is prohibited in both sides
'1'	Spread only to the right side
'2'	Spread only to the left side
'3”	Spread allowed in both sides

- Impassability: This field indicates whether you can cross the link or not. For example, if you are on a side of a river, your can't access the other side directly, you need to go to a bridge that crosses. Conversely, I you are on a side of an underground metro line at a ground level, you can cross easily as you are not on the same level. In theory, you can imagine barriers that you can cross only in one direction, for example at a border or to enter in some areas where entrance are not in the same place as exit.

'0'	You can cross the link from any side
'1'	You can cross only from left to right
'2'	You can cross only from right to left
'3”	You can't cross the link

5.3.3 Examples of network links codification

The next table gives examples of links codification.

Type of link	Flow direction	Spread	Impassibility
Bus (on road)	'1'	'0'	'0'
Metro (undeground)	'1'	'0'	'0'
Train (road level)	'1'	'0'	'3'
Highway	'0' or '1' or '2' or '3'	'0'	'3'
Streets (low trafic)	'1' or '2' or '3'	'3'	'0'
Streets (high trafic)	'1' or '2' or '3'	'1' or '2'	'0'
River or railways	'0'	'0'	'3'
Bridge or tunnel	'1' or '2' or '3'	'0'	'0'

5.3.4 Method of interpolation

- Firstly, you have to grid your study area.
- Secondly, you had to fetch each link in your study area where you can spread from.
- Then for each grid cell in a search radius distance from the link, you find the closest point on link from the center of the cell. In Qgis, this is done with the closestSegmentWithContext API function. It corresponds to the orthogonal projection of the cell center on the link. If this projection isn't on the link, the closest point is either the link starting point or ending point. You have to check the right/left side position of the cell center from

the link, if spread from link is allowed only to the left or to the right side.

- If the distance between the grid cell and the link is lower than the smallest distance between the grid cell and all previous links, and the if the line between the link and the grid cell doesn't cross a link which is impassable, then the grid cell minimum distance is updated and the cell access time is calculated as follow.

$$T_{i,j}=(1-\lambda)t_{k,endpoint}+\lambda t_{k,startpoint}+v_{spread}\,.\,d_{(i,j),k}$$

i,j :refers to cell i,j
k :refers to link k
λ :refers to fraction of line length defined from the link startpoint to the orthogonal projection of the grid cell
$t_{k,startpoint}$ and $t_{k,endpoint}$:refers to acces times for link k at startpoint and endpoint
v_{spread} :refers to the speed of spread can be a constant or a link attribute
$d_{(i,j),k}$:refers to the shortest distance between the grid cell and the link

- After having fetched all links, you have a geographic grid representing access times, which can be either displayed as a raster or either as isolines with contour generation algorithms as proposed in Qgis or Grass.

6 Application

6.1 Implementation in Qgis

Limewise method in Qgis is implemented in networks plugin, available in the Qgis repository, and takes into account flow directions, side of spread described with a global or a link based spread speed, impassability and travel times varying along each link length. This plugin also proposes a set a tools to help users to prepare networks or graphs. As Limewise method works with oriented graphs, users have to choose between dealing only with one-way links or two-ways links. Working with two ways links requires that user use four access time fields, two for each direction (end node and start node times). Dealing with one-way links only, needs to duplicate two-ways links and reverse the direction (for example using the “reverse” plugin command) in order to facilitate its use.

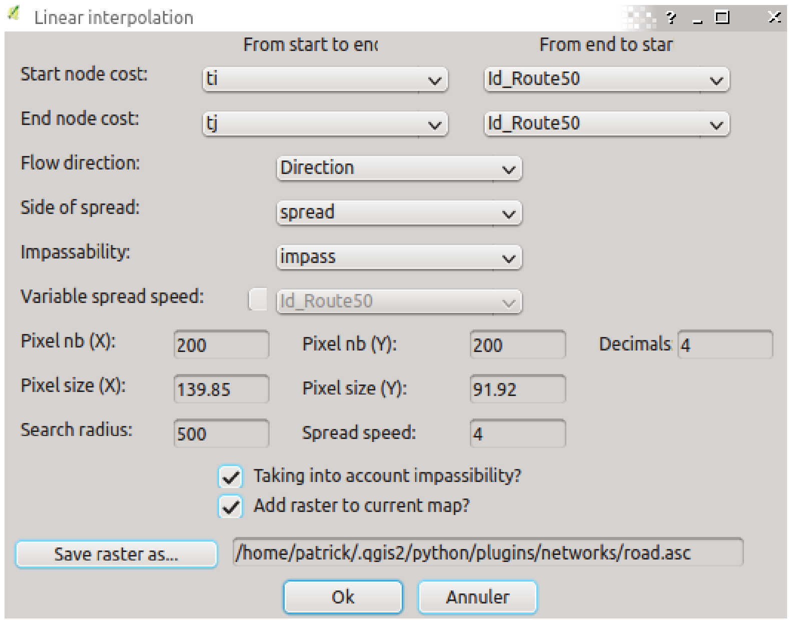


Figure 5: Limewise dialog in Qgis

6.2 Examples

The following examples shows differences using different settings in Limewise algorithms. They are related to a road accessibility from an origin in the center of Lille

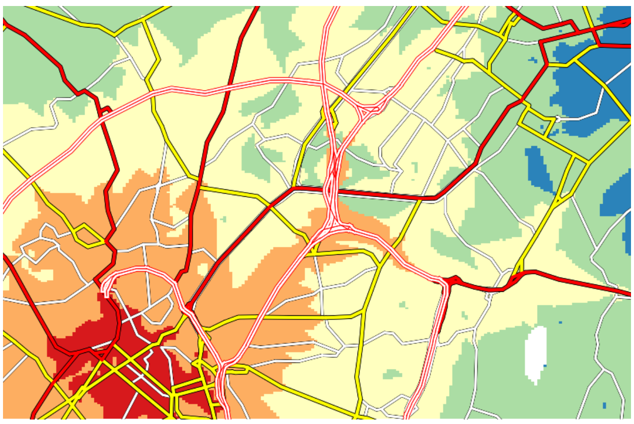


Figure 6: Limwise with basic settings (Palmier 2015)

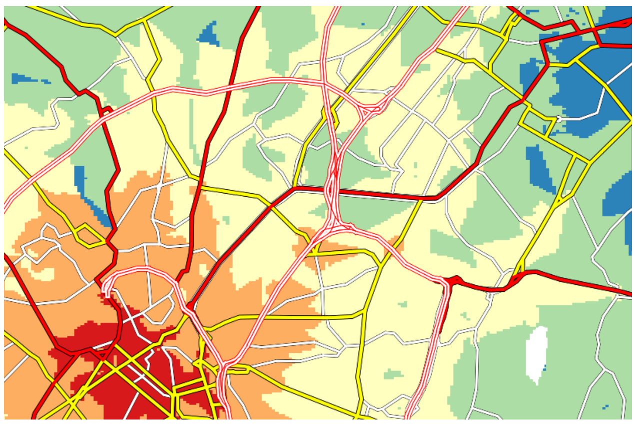


Figure 7: Limewise with user defined settings (adapted to local environment) (Palmier 2015)

On the left, the map shows Limewise basic settings, where spread is possible on each side for every link. On the right, spread is prohibited for motorway links, and only allowed on right side for primary links (in red). In the same time, motorway links are impassable. Color gradient from red to blue, is divided by steps of 1O minutes with a spread speed of 20km/h.

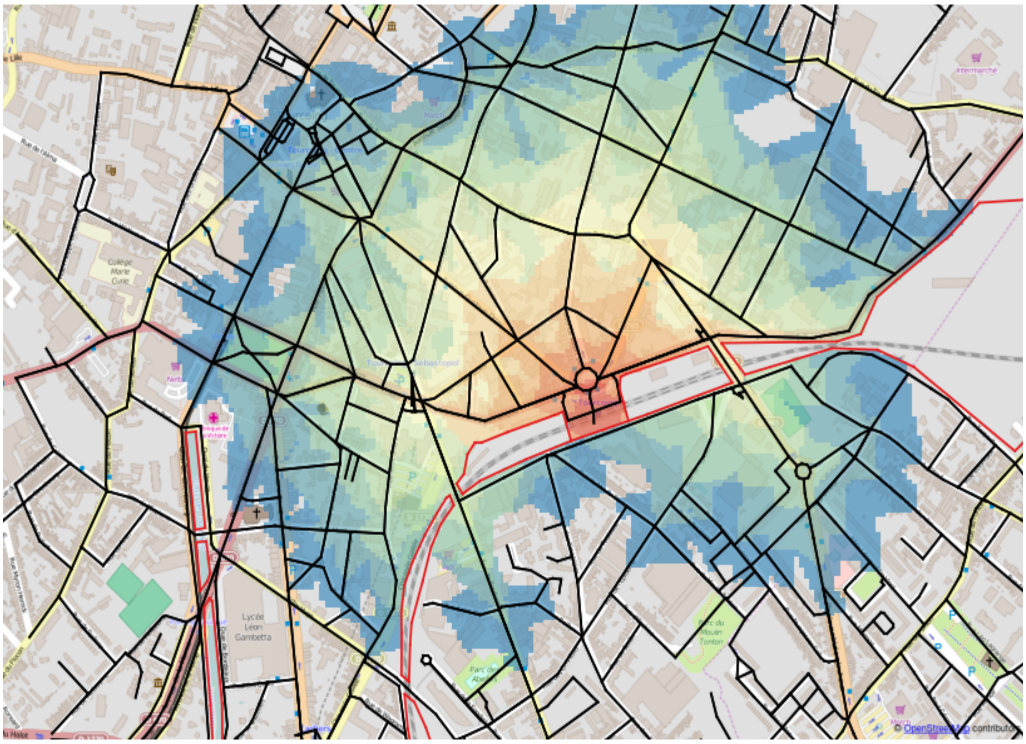


Figure 8: Limewise with user defined settings (adapted to local environment) Palmier (2015)

The map above shows the Limewise method applied to the Tourcoing station which is only north-oriented. Streets are colored in black, and railways corridors limits in red, are considered as impassable. Color gradient from red to blue describes the walking accessibility to the station by steps of 100m. This method points out very clearly dead ends (in the bottom middle a the map) where accessibility to the station is very poor. A map analysis, shows the advantages provided by the Limewise method. First, access times on links are realistic, and those inside blocks are easy to explain or interpret as rule of spread are well known and set by the use.

6.3 Variability over a time period

The remaining question, about accessibility concerns the variability over a time period, as public transport accessibility strongly depends on timetables, especially with low frequencies lines. (Richer & Palmier, 2011) propose a method called TIP to represent multimodal accessibility over a time period defined by a combination of three indicators (travel time, intensity, hardness). These indicators are build by simulating 60 departures in an hour, by steps of a minute. Travel time is defined by the average of the every travel times, intensity by the number of different alternatives and the standard deviation on travel times, and hardness by the average number of boardings. These indicators are synthesized in a statistical classification.

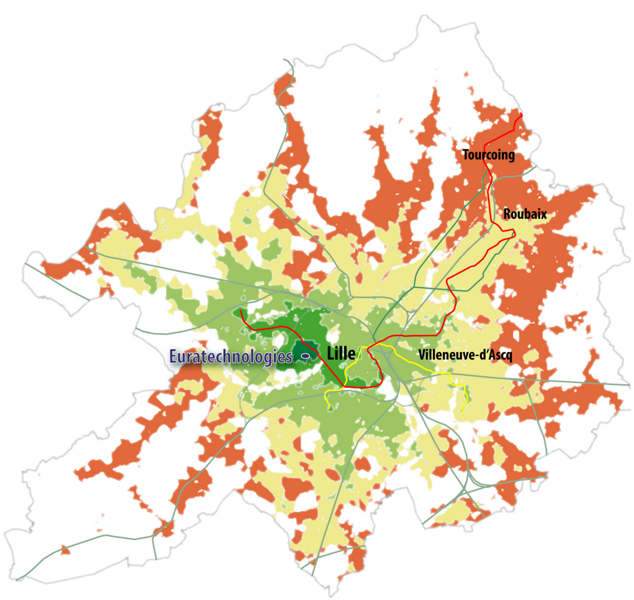


Figure 9: minimum travel time method (Richer – Palmier 2011)

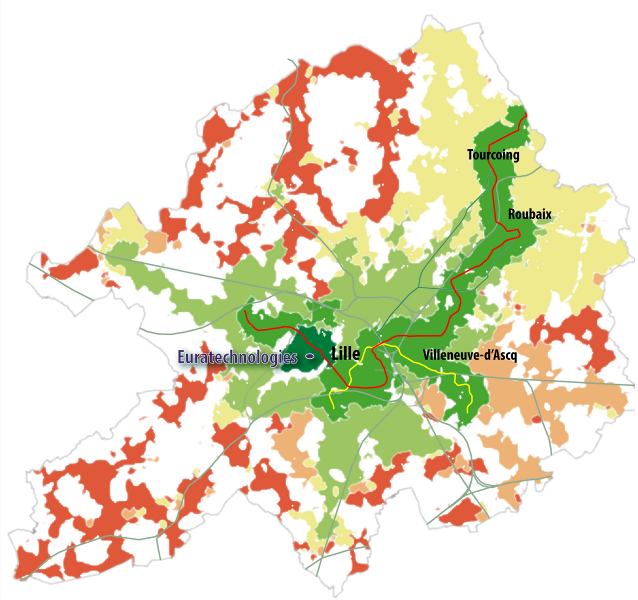


Figure 10: TIP method (Richer – Palmier 2011)

The TIP method on the right shows that the minimum travel time indicator is not very adapted for public transport accessibility. The TIP method shows the huge impact of VAL metro lines, which with a frequency between 1 and 2 minutes, give a high accessibility level compared to some areas, which could have a smallest access times, corresponding to a quick service with low frequency, therefore with poor access time during the rest of the period.

7 Conclusion

This article outlines the need for strong interactions between multimodal routing calculation processes and isochrone generation algorithms, while today each one works in a rather separated way. Isochrone creations algorithms must take advantage of all the information that can provide multimodal routing algorithms , and routing algorithms must adapt their output to produce indicators that are designed for the generation of accessibility maps, which will be precise, accurate and easy to interpret for decision makers. The implementation in Qgis of the Limewise methodology associated with the free Musliw tool provides a set of tools that fully responds to these requirements. These tools are part of a regional project in Nord - Pas de Calais region, where CEREMA has developed a database of theoretical multimodal transport supply. This database is provided with free tools as Qgis, and Musliw, in order to allow users to exploit the database without the need to purchase expensive software they can't afford. Several training sessions were conducted with local authorities and agencies, so that they can take full benefit of their own data, which was unfortunately, most of the time, dedicated to the operator.

References

✔ Bousquet, A. (2010). Optimisation d'itinéraires multimodaux fondée sur les temps de parcours à l'échelle d'une agglomération urbaine dense (Doctoral dissertation, Ecole Nationale des Travaux Publics de l'Etat).

- ✔ Chatalic, V. (2012) Université Paris1-Panthéon Sorbonne Master 2 aménagement du territoire et urbanisme Mémoire de recherche.
- ✔ Delling, D., Pajor, T., & Werneck, R. F. (2014). Round-based public transit routing. Transportation Science.
- ✔ Dibbelt, J., Pajor, T., Strasser, B., & Wagner, D. (2013). Intriguingly Simple and Fast Transit Routing. SEA, 7933, 43-54.
- ✔ Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ. American journal of preventive medicine, 28(2), 117-125.
- ✔ Gräbener, T., Berro, A., & Duthen, Y. (2010). Time dependent multiobjective best path for multimodal urban routing. Electronic Notes in Discrete Mathematics, 36, 487-494.
- ✔ Hansen, W. G. (1959). How accessibility shapes land use. Journal of the American Institute of Planners, 25(2), 73-76.
- ✔ Oliver, L. N., Schuurman, N., & Hall, A. W. (2007). Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. International journal of health geographics, 6(1), 41.
- ✔ Palmier, P. (2001). Cartes ZAP: mode d'emploi. Lille Métropole Communauté Urbaine.
- ✔ Palmier, P. (2011). Musliw. Logiciel de calcul d'accessibilité multimodale. SETRA
- ✔ Richer, C., & Palmier, P. (2011). Mesurer l'accessibilité en transport collectif aux pôles d'excellence de Lille Métropole. In Mobilités spatiales et ressources métropolitaines: l'accessibilité en questions/11ème colloque du groupe de travail" Mobilités Spatiales et Fluidité Sociale" de l'AISLF.c
- ✔ Spiess, H., & Florian, M. (1989). Optimal strategies: a new assignment model for transit networks. Transportation Research Part B: Methodological, 23(2), 83-102.
- ✔ Wardrop, J. G. (1952, June). ROAD PAPER. SOME THEORETICAL ASPECTS OF ROAD TRAFFIC RESEARCH. In ICE Proceedings: Engineering Divisions (Vol. 1, No. 3, pp. 325-362). Thomas Telford.

The integration of land change modeling framework FUTURES into GRASS GIS 7

Anna Petrasova, Vaclav Petras, Douglas A. Shoemaker, Monica A. Dorning, Ross K. Meentemeyer

Center for Geospatial Analytics, North Carolina State University, Raleigh, NC, USA

Abstract

Many valuable models and tools developed by scientists are often inaccessible to their potential users because of non-existent sharing infrastructure or lack of documentation. Case in point is the FUTure Urban-Regional Environment Simulation (FUTURES), a patch-based land change model for generating scenario-based regional forecasts of urban growth pattern. Despite a high-impact publication, few scientists, planners, or policy makers have adopted FUTURES due to complexity in use and lack of direct access. We seek to address these issues by integrating FUTURES into GRASS GIS, a free and open source GIS and research platform for geospatial domain. This integration will enable us to take advantage of GRASS GIS tools for landscape structure analysis, and thus eliminate the need to use proprietary software for data preprocessing. Moreover, integration into GRASS GIS simplifies the distribution of FUTURES across all main operating systems and ensures maintainability of our project in the future. We will present our use case of integrating this advanced land change model into GRASS GIS platform and discuss the current state of the integration as well as the planned steps to achieve our vision of simple-to-use and fully free and open source FUTURES.

Keywords

geospatial, urbanization, open science, reproducibility, free and open source

1 Introduction

Despite all currently available technologies and tools for collaborative research and software development, many academic projects, although published in high-impact peer-reviewed journals are not adopted by disciplinary communities and thus fail to have broader impact for which they were intended. In the case of sophisticated geospatial models and analyses, acceptance of particular tools requires more than releasing the code with appropriate license online. Detailed and updated documentation, along with sample data of sufficient complexity, is required to demonstrate the model features and evaluate its suitability for use in research. A well defined user and programming interface implemented in a programming language widely used by disciplinary practitioners is crucial for further usage and extensions. Availability of the model across different operating systems is not only convenient for various users but it also simplifies model coupling.

The introduction of the land change modeling framework FUTURES (Meentemeyer et al., 2012) created to map regional projections of urban growth is an example of disconnect between modelers-developers and potential users. The authors of the FUTURES model first published a study of land development dynamics in the rapidly expanding metropolitan region of Charlotte, North Carolina, and later an analysis of the impacts of urbanization on natural resources under different conservation strategies (Dorning et al., 2015). Despite the recognition of high accuracy and novelty of the model, its use has been limited to FUTURES' authors and close collaborators. Limited access to the model together with missing documentation have slowed the progress of updating and adding new features, effectively barring the land change community from adopting the model.
Publishing the model as open source software is a solution that can address many of these disconnects, and create new opportunities for the scientific community to explore, apply and modify FUTURES for their own research. In this work-in-progress we rework the FUTURES urban growth model as an open source geospatial tool, highly relevant for the land change community, by integrating it into GRASS GIS, a free and open source GIS (Neteler et al., 2012). GRASS GIS natively provides many features, including landscape structure analysis, efficient large raster data processing and spatio-temporal visualizations, making it a suitable geospatial research platform for FUTURES.

2 GRASS GIS as a research platform for FUTURES

The FUTURES framework consists of several interconnected submodels (Figure 1). The core component is a stochastic, patch-growing algorithm (PGA) that bridges field-based and object-based representations of landscape change by constructing discrete land conversion events understood as urban growth. The Potential submodel feeds PGA with urban growth suitability based on statistical relationships of historic land change and significant environmental, infrastructural, and socioeconomic factors. The Demand submodel estimates population demand for urban development based on increases in population concurrent with the rate of landscape change. PGA, as the FUTURES engine, was written in C for performance reasons, whereas Potential and Demand were originally informally implemented as R scripts and ArcGIS workflows.
From the perspective of FUTURES architecture, GRASS GIS is a highly qualified geospatial platform for deployment. GRASS GIS functionality is divided into independent modules written in C and Python, which makes GRASS GIS compatible with FUTURES modular structure. This architecture allows for GRASS GIS to run on different operating systems and platforms including high performance computing (HPC) clusters (Metz, Rocchini & Neteler, 2014). Distribution of contributed scientific models is ensured through the GRASS GIS Addons repository, a browsable revision control repository maintained by OSGeo Foundation.
The source code in all GRASS GIS repositories is maintained by the GRASS GIS Development team. GRASS GIS has a long history of integrating and preserving scientific analytical tools (Chemin et al., 2015). Models integrated in the past are still maintained, improved and used with or without contribution from original authors. References to the original work are part of the documentation

and the original authors are cited in the derived works, see for example Di Leo et al. (2013). This long-term maintenance solves the issues occurring with different operating system, compilers, libraries and application programming interface (API) changes. GRASS GIS libraries and API offer standardized ways to effectively read, write, and organize geospatial data including optimizations for reading and processing of large raster and vector files which relieves the burden of implementing these low-level operations from the modeler-programmer. In addition, GRASS GIS allows to create a standardized interface for modules; using this mechanism GRASS GIS automatically generates a GUI, command line interface and HTML documentation for every module.
GRASS GIS users can easily work with R statistical software. Smooth transitioning between GRASS GIS and the R environment is possible thanks to spgrass6 and rgrass7 packages for R. The same applies for Python where GRASS GIS itself provides the API for Python so that Python libraries such as Numpy, SciPy, and Matplotlib can be easily used (Zambelli, Gebbert, S., & Ciolli, 2013).

3 Integrating FUTURES into GRASS GIS

We focused first on the patch-growing algorithm as the core component of FUTURES. Currently, the PGA part of FUTURES is fully integrated into GRASS GIS and available as part of a growing addon meta-module r.futures that will ultimately incorporate all the components. Since it is the most computationally demanding part of FUTURES we are currently exploring optimization options. However, just using efficient GRASS GIS I/O libraries resulted in significant speedup. The implementation of the two statistical submodels will be formalized as GRASS GIS addon modules using GRASS GIS Python API and NumPy to perform geospatial tasks and simple statistics. Potential as the more complex statistical model in R will be wrapped to provide a GRASS GIS interface to simplify chaining the modules. Large amounts of time and effort have been spent in manual preprocessing of input data, therefore we are currently automating the process by adding a new Calibration submodel which evaluates the patch size, shape and distribution needed for the PGA submodel (Figure 1).

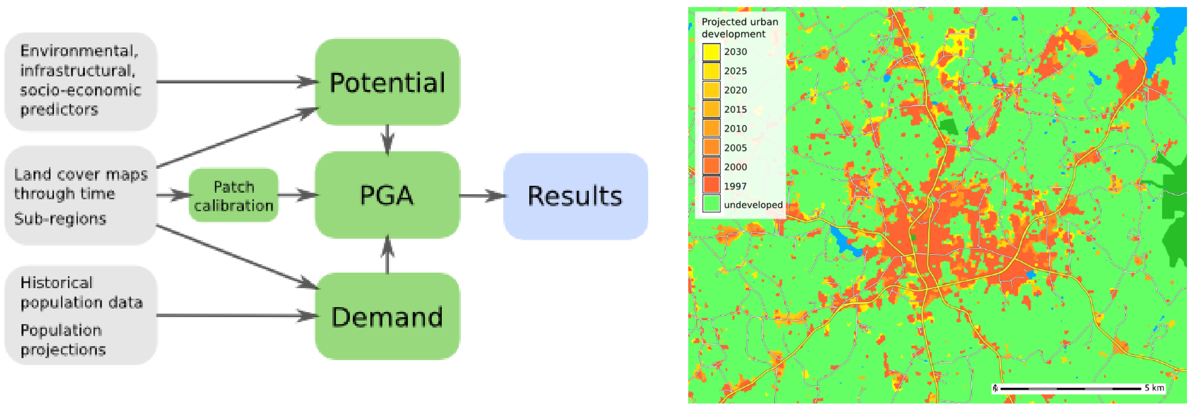


Figure 1: *On the left FUTURES simplified schema with inputs in gray and submodels identified for integration into GRASS GIS as modules in green. On the right the result of one run of FUTURES showing the incremental urban growth in part of Charlotte region, NC in thirty years.*

To support different geographic scales and wider range of applications we plan to design the GRASS GIS interface of the FUTURES addon modules to be flexible enough to allow users to input data at different levels of complexity or even substitute a selected FUTURES submodel with their own implementation. Moreover, the PGA could be reused for different applications than urban growth (e.g., epidemiology), and therefore we aim at generalizing the interface to accommodate such cases. Once all the FUTURES components are integrated into GRASS GIS, we will automate the entire procedure by chaining the FUTURES modules into a meta-module, useful not just for new FUTURES users but also for running FUTURES on HPC systems.

During the GRASS GIS FUTURES implementation we will prepare sample data for testing purposes based on publicly available data. All documentation, including parameter description and examples based on the dataset, will be distributed as manual pages to FUTURES addon modules. Upon completion and testing of each component, new addons will be published in the GRASS GIS repository (grass.osgeo.org/grass70/manuals/addons/r.futures.html).

References

- ✔ Bivand, R. (2007). Using the R–Grass interface. OSGeo Journal, 1, 36-38.
- ✔ Chemin, Y., Petras, V., Petrasova, A., Landa, M., Gebbert, S., Zambelli, P., Neteler, M., Löwe, P., & Di Leo, M. (2015). GRASS GIS: a peer-reviewed scientific platform and future research repository. Poster at General Assembly 2015 of European Geosciences Union, Vienna, Austria.
- ✔ Di Leo, M., de Rigo, D., Rodriguez-Aseretto, D., Bosco, C., Petroliagkis, T., Camia, A., & San-Miguel-Ayanz, J. (2013). Dynamic data driven ensemble for wildfire behaviour assessment: a case study. In Environmental Software Systems. Fostering Information Sharing (pp. 11-22). Springer Berlin Heidelberg.
- ✔ Dorning, M. A., Koch, J., Shoemaker, D. A., & Meentemeyer, R. K. (2015). Simulating urbanization scenarios reveals tradeoffs between conservation planning strategies. Landscape and Urban Planning, 136, 28–39. doi:10.1016/j.landurbplan.2014.11.011
- ✔ Meentemeyer, R. K., Tang, W., Dorning, M. A., Vogler, J. B., Cunniffe, N. J., & Shoemaker, D. A. (2012). FUTURES: Multilevel Simulations of Emerging Urban–Rural Landscape Structure Using a Stochastic Patch-Growing Algorithm. Annals of the Association of American Geographers, (October 2012), 23. doi:10.1080/00045608.2012.707591
- ✔ Metz, M., Rocchini, D., & Neteler, M. (2014). Surface temperatures at the continental scale: tracking changes with remote sensing at unprecedented detail. Remote Sensing, 6(5), 3822-3840.
- ✔ Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. Environmental Modelling & Software, 31, 124-130. doi:10.1016/j.envsoft.2011.11.014
- ✔ Zambelli, P., Gebbert, S., & Ciolli, M. (2013). Pygrass: An object oriented python application programming interface (API) for geographic resources analysis support system (GRASS) geographic information system (GIS). ISPRS International Journal of Geo-Information, 2(1), 201-219.

Evaluation of wind, solar and hydro energy potential using GRASS

Giulia Garegnani, Pietro Zambelli, Gianluca Grilli, Daniele Vettorato

European Academy of Bozen/Bolzano, Italy

Abstract

One of the aim of the European directive 2009/28/EC is the promotion of the use of energy from renewable sources in order to decrease the fossil fuel dependency. In the Alpine region the application of this directive is particularly relevant because of the complex morphology of the territory with a high renewable energy potential due to the presence of several renewable sources (Vettorato & al., 2011). On the other hand, the production of energy from renewable sources creates new environmental pressures on the Alps that are sensible and fragile from the biodiversity, the landscape value and in general for their ecosystems services preservation point of view. The understanding of the trade-off between renewable energy potential and ecosystem services is needed in order to support decision makers in the sustainable energy planning phases . Usually the spatial information about energy potential is missed or data are inhomogeneous and depend on specific study of each alpine country. The aim of this article is to present 3 GRASS modules (r.green.wind; r.green.solar; r.green.hydro) that help to obtain a clear understanding of the spatial distribution of hydro, solar and wind-power in complex territories like the Alps. We evaluated the theoretical potential with a transparent methodology starting from physical parameters describing the availability of the renewable source (i.e. wind velocity, irradiation, discharge etc.). The methodology, thanks to the customizable GRASS modules, can fit also other regions . The main outputs of the r.green package is given by energy potential maps that consider theoretical, economic, technical and legislative aspects as well as the ecosystem services preservation. In this paper only the first step of the process (theoretical potential calculation) is presented.

Keywords

Energy potential maps, Alpine region, Hydro-power, Solar-power, Wind-power

1 Introduction

Several studies have been done in order to compute the energy potential from renewable energy sources (REs) by means of GIS systems. Studies are often not comparable, they differently define the potential and they regard only one of the REs available in the Alpine region. Several authors suggest a definitions of energy theoretical and technical potential, especially (Chatzimouratidis & Pilavachi, 2009; de Vries, van Vuuren, & Hoogwijk, 2007; Resch et al., 2008;

Sacchelli, Zambelli, Zatelli, & Ciolli, 2013; Sacchelli et al., 2013). From the theoretical potential, i.e. all the available energy in the nature, the technical, sustainable and economical potentials can be derived by applying different constrains, i.e. by excluding too steepest areas, natural parks, etc. Assumptions on average values and trends can strongly influence the potential for this reason a transparent definition of potential and a clear description of the methodology should be mandatory. This work focuses on the methodology to compute the theoretical potential and some technical constrains in a quite large region as the Alps. The first step is to define the physical variables necessary to compute theoretical potential maps. Regarding the hydro-power, it is necessary to know the availability of water and the gross head, i.e the elevation model and discharge data in order to compute the potential energy. In the case of wind power, the energy is related to the wind velocity, the annual irradiation is instead necessary to compute the solar power.

It is not the aim of the article a deeper analysis of sector studies to collect, compute and recalibrate all these physical parameters. In the literature, (Huld, Müller, & Gambardella, 2012; Šúri, Huld, Dunlop, & Ossenbrink, 2007) develop an European solar radiation database. Several studies are related to hydrological models and computation of the run-off (Formetta, Antonello, Franceschi, David, & Rigon, 2014; Formetta, David, & Rigon, 2012; Norbiato, Borga, Merz, Blöschl, & Carton, 2009; Sauquet & Leblois, 2001; Schmidt & Morche, 2006). In the case of wind velocity, data for the alpine region are available and analyzed (Schaffner, Remund, Cattin, & Kunz, 2006). We try to use a unique information source covering the entire Alpine region for each physical parameters. The resolution, precision and accuracy of the data are coarser than in the case of a limited area but the methodology is still suitable for more specific studies and results provide a useful description of a more general Alps potential without entering in the details of a local case study.

The second step is the choice of the algorithms to compute the energy potential of the REs. The hydro-power potential depends on the river discharge and the gross head (Bódis, Monforti, & Szabó, 2014; Kusre, Baruah, Bordoloi, & Patra, 2010; Palomino Cuya, Brandimarte, Popescu, Alterach, & Peviani, 2013; Punys & Pelikan, 2007). The potential usually refers to a single water basin and should be considered has a mean value. Palomino Cuya et al. (2013) investigate also the possibility to transfer water, i.e. energy power, from one river-basin to another ones. In the case of wind power, (Baban & Parry, 2001; Draxl & Mayr, 2011; Mari et al., 2011; van Haaren & Fthenakis, 2011) build a decision support system starting from wind velocity maps. Draxl & Mayr (2011) and Mari et al. (2011) compute the energy potential starting from wind distribution functions and the power curve of available turbines.

The aim of this contribution is then to identify a methodology that can be applied in other areas in order to provide an essential input, i.e. the theoretical potential maps, for decision support systems (DSS) in the field of REs planning. In fact, DSS are very important multi-disciplinary tool to plan a very complex territory as the alpine area and to find alternative energy solutions. In this work, firstly we clarify the terminology based on the literature (Chatzimouratidis & Pilavachi, 2009; de Vries, van Vuuren, & Hoogwijk, 2007; Resch et al., 2008; Sacchelli et al., 2013) in order to have comparable results

for the three sources. After reviewing available data in the alpine area for the REs, we report the algorithms developed to compute the theoretical potential in the first section. The methodology is applied for the Alpine region and results are analyzed in order to provide first recommendation to decision makers and stakeholders in the last section.
The paper presents an innovative methodology to identify where and what kind of energy potential it is available at regional scale, considering not only the geographical and physical factors, but also the constraints due to the current technology available on the market and their exploitability in the Alpine region.

2 Methodology and GIS algorithms

In the first section, we define the theoretical potential based on the work of Resch et al. (2008), we clearly define the terminology and the criteria to evaluate the potential.
We consider the theoretical limit as the maximum energy that can be theoretically converted in the ideal case. For the theoretical potential general physical parameters are taken into account.
Then, theoretical limits cannot be reach in practice. Secondly, another limit can be defined by the actual technology and by the efficiency limit that can be reached with the actual technology. Besides, overall technical limitations (i.e. the available land area, steepest areas, etc.), planning and legislation constrains have to be added. Ecological constrains and sustainable criteria can limit the technical potential to the sustainable potential. Finally, we can derive the economical potential, i.e the market and the realization cost.
In this work we focus on wind velocity, solar irradiation and water availability and their theoretical based on physical law and current technology without any other constrains.
The methodology presented in this paper is based on GRASS GIS 7. The algorithms are implemented in Python and split in several sub-modules to preserve the granularity and independence of task for each renewable energy source. All the developed modules use and combine whenever possible the existing GRASS modules, but for some specific tasks, especially when we need to use other python libraires such as Numpy or Scipy, the PyGrass library were used (Zambelli, Gebbert, & Ciolli, 2013).
The physical parameters considered to assess the energy potential for the four REs are reported in table 1.

RES	Physical parameters	Variable	Unit
Hydro-power	Gross head Discharge	ΔH Q	m m^3s^{-1}
Wind-power	Mean velocity	v	$m\,s^{-1}$
PV	Solar irradiation	E_{sol}	$KW\,m^{-2}$

Table 1: Physical parameters for theoretical potential estimation.

2.1 Hydro-power

In the case of hydro-power, the maximum potential is the energy that can be produced under the assumption that all the water resource is used. The potential does not consider the environmental flow and other use of water (aqueducts, irrigation, etc.).

The base unit for hydro-power estimation is the river basin. All the physical variables necessary for the potential estimation refers to the base unite and they are the afferent discharge Q_{aff} , the total discharge at the closure point of the basin $Q_{closure}$, the mean elevation h_{mean} and the closure elevation $h_{closure}$ as shown in figure 1. In the case of referent basin 3, the discharge $Q_{3,aff}$ is equal to the difference between the discharge at the closure point of the basin 3 and the sum of the closure discharges of basins 1 and 2, Figure 1:

$$Q_{3,aff}=Q_{3,closure}-(Q_{1,closure}+Q_{2,closure})$$

The hydro-power potential of a referent basin *i* is evaluated in a qualitative way as

$$P_{i,aff}=g\,Q_{i,aff}\,\Delta h$$

where P is the power, g the gravity acceleration, $\Delta h=h_{i,mean}-h_{i,closure}$ the gross head equal to the difference between the average elevation of the basin and the elevation at the river basin closure.

The power $P_{i,aff}$ is the gross theoretical potential, i.e. the annual energy potentially available in the catchment if all natural flows were turbined down to the closure point of the catchment.

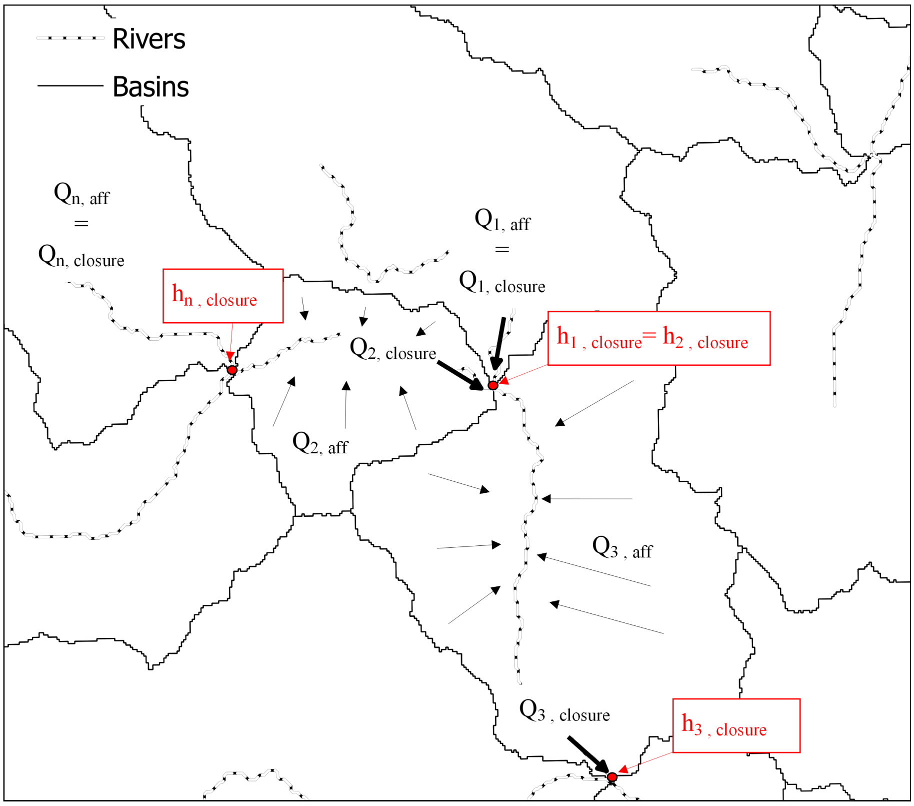


Figure 1: Sketch of variables for hydro-power evaluation.

In order to calculate the gross head we assume that it depends on the difference between the mean elevation of the referent basin and the elevation at the closure point. This gross head can be an average of several run-of the river power station.
Following Mari et al. (2011) we don't consider only the own power of the catchment but also the interaction we upper basins as shown in figure 1. The hydro-power potential depends on two variables, the discharge and the gross head. The discharge coming from upper basins can be used to produce energy in the lower basin. This can be done by means of a system of weirs and reservoirs. In this case the potential related to the base unit and due to the upper discharge is equal to:

$$P_{i,up}=\sum_{j\in UP}\left(g\,Q_{j,closure}\left(h_{j,closure}-h_{i,closure}\right)\right)$$

where UP is the set of the upper basin, we set a maximum number equal 3. In order to identify the network of basins in the Alps region and their relationships, a search algorithm is developed in python language starting from the output of *r.watershed* by means of PyGrass library (Zambelli, Gebbert, & Ciolli, 2013).
Finally, the total power of the basin is given by:

$$P_i=P_{i,up}+P_{i,aff}$$

2.2 Solar-power

The photovoltaic (PV) effect is the basis of conversion of light to electricity in photovoltaic, or solar, cells. Hersch & Zweibel (1982) simply describe the PV effect as follows: light, which is pure energy, enters a PV cell and imparts enough energy to some electrons to free them. A built-in-potential barrier in the cell acts on these electrons to produce voltage which can be used to drive a current through a circuit.
According to european norm EN 15316-4-6 for the solar photovoltaic, the electricity produced by the photovoltaic system is

$$E_{el,pv,out}=E_{sol}f_{perf}\frac{P_{pk}}{I_{ref}}$$

where
E_{sol} is the annual solar irradiation on the photovoltaic system, P_{pk} the peak power, represents the electrical power of a photovoltaic system with a given surface and for a solar irradiance of 1 kWm^2 on this surface (at 25°C); f_{perf} is the system performance factor; I_{ref} is the reference solar irradiance equal to 1 kWm^2. With actual technologies the performance factor f_{perf} can reach value of 0.80.
Most solar cells on the market are based on silicon wafers with typical efficiency of 10-15\% (P_{pk} 100-150 W_pm^{-2}). Conversion efficiency needs to be increased in the future. According to the aim of the article the upper theoretical limit has to be considered. Shockley & Queisser (1961) studied this limit for or a standard solar cell. An optimal cell with a band gap of 1.3 eV is limited by transmission losses of photons to 31% (310 W_pm^{-2}).
According to thermodynamic law, by considering the sun as a black body at 5760 K and a solar cell as another black body at 300 K the conversion

efficiency is related to the Carnot efficiency limit, which is nearly 95% (Green, 2002). Notice that the Carnot limit is only a theoretical limit and cannot be built in practice.
The large difference between the Shockley-Queisser and thermodynamic (Carnot) limits arises from the fact that a single material is characterized by a single energy gap, whereas the solar spectrum contains photons with a wide range of energies.
Several methods have been studied to increase the power conversion efficiency of solar cells, Razykov et al. (2011) review the future prospect of photovoltaic devices.

2.3 Wind-power

The mean power production for a wind turbine assuming 100 percent availability is equal to:

$$E=\int_0^{\infty} P(v)\varphi(v)dv$$

where $P(v)$ is the power curve of a wind turbines, $\varphi(v)$ is the statistical distribution of the wind velocity v. The wind velocity depends on the elevation that we consider for the hub. The potential is then strictly related to this assumptions.
The distribution of the wind velocity is usually a Weibull (Celik, 2003):

$$\varphi(v)=\frac{k}{\lambda^k}v^{k-1}e^{-(\frac{v}{k})^k}$$

where the scale parameter λ can be calculated from the following equation:

$$\bar{v}=\lambda\Gamma(1+\frac{1}{k})$$

with Γ gamma function and $\bar{v}$ mean yearly velocity.
The power $P(v)$ of a wind turbine is a function of the wind velocity v:

$$P(v)=\frac{1}{2}c_p\rho A v^3$$

where c_p is the power coefficient, ρ is the air density and *A* the swept area.
The power coefficient has a theoretical maximum value of 0.593, the so-called Betz limit (Bergey, 1979). It represents the maximum limit to the amount of energy that can be converted in usable power.
We consider the Betz limit has the theoretical maximum value for the estimation of the wind potential. The wind turbine converts 70% of the Betz limit into electricity. Consequently, the wind turbine converts the 41% of the available wind energy into electricity. This module has been developed using PyGrass.

3 Estimation of RE sources in the Alps

The availability of high quality GIS data played the strongest role in determining the results of this research. GIS data is often inconsistent across large regions due to the different policies in each alpine countries. It is not the aim of this article to analyzed the input data but a critical review it is necessary in order to know the uncertainty and the limits of input data.
We summarize the input data-set for the physical parameters in table 2.

RES	Variable	Input Source	Output Unit
Hydro-power	ΔH Q	STRM GRDC, MeteoSwiss	Basin (area ~ 150 km^2)
Wind-power	v	WINDHARVEST	300 m/px
PV	E_{sol}	PVGIS	1 km/px

Table 2: Physical parameters for theoretical potential estimation.

As shown in the previous section, the unit base for hydro potential maps is the river basin. By processing elevation data, rivers and watershed basins in the alpine area can be computed (Ehlschlaeger, 1989; Holmgren, 1994). The algorithm is available as a GRASS GIS software extension. We set the threshold parameter equal to 10000 and we use CIAT data-set derived from the USGS/NASA SRTM elevation data with a raster resolution of 60m. Jarvis, Reuter, Nelson, & Guevara (2008) have processed this data to provide seamless continuous topography surfaces.

The GRDC repository of river discharge data provides annual mean values for some gauging stations in the Alps. The database comprises discharge data of nearly 9.000 gauging stations from all over the world but does not obviously cover all the basins in the alpine region. For the basins without gauging stations, we use a rain data-set to compute the afferent discharge of each river basin as:

$$Q_{aff}=\sum_{i=1}^{N} A_i p_i c_d$$

where A_i is the cell area, p_i is the rain precipitation in the i-th cell, c_d is the runoff coefficient and N is the number of cells in the each basin. The rainfall dataset has been provided by the federal office of meteorology and climatology MeteoSwiss and was developed as part of EU project EURO4M (Isotta et al., 2014). The daily precipitation is derived from a pan-Alpine high resolution rain-gauge dataset of More than 8500 time series in total, approximately 5500 measurements each data. The precipitation considers rainfall plus snow water equivalent. The cover period is January-December, 1971-2008. The resolution in grid spacing is 5km/px, but the effective resolution is approximately 10-20 km, depending on local station density. Since we do not know the runoff coefficients of each basin in the Alps we set them equal to 0.5. We report in Figures 2, 3 and 4 the potential maps obtained by considering the whole potential energy and the 80% of the all available energy.

The Photovoltaic Geographical Information System (PVGIS) provides an inventory of solar energy resource in Europe (Šúri, Huld, Dunlop, & Ossenbrink, 2007). The map resolution is 1km/px. The data used to estimate solar potential are the global irradiation maps for horizontal surface and for surface inclined at an optimum angle with clear sky condition. The computation accounts for sky obstruction by local terrain features. As reported in the previous section, we consider the 95% of the solar radiation for the Carnot limit, the 30% for the Shockley-Queisser limit and 15% for the actual technology, Figure 3.

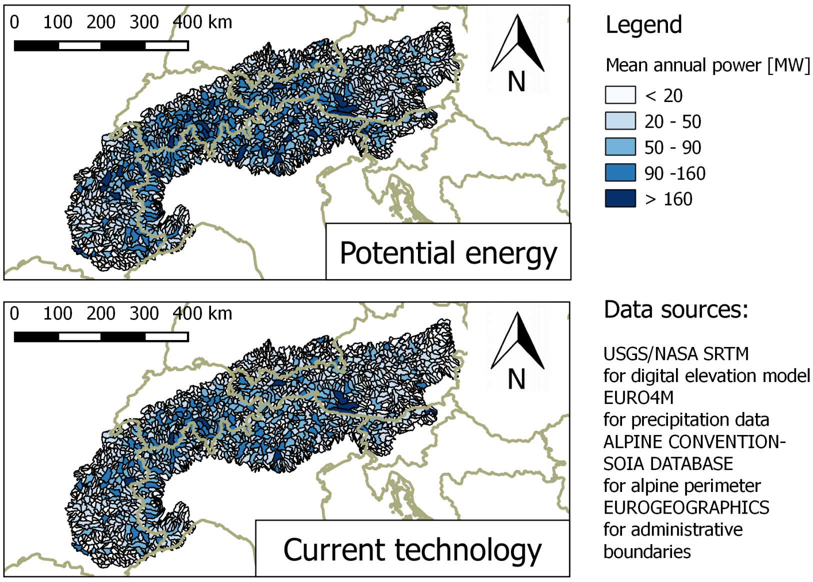


Figure 2: Hydro-power potential maps.

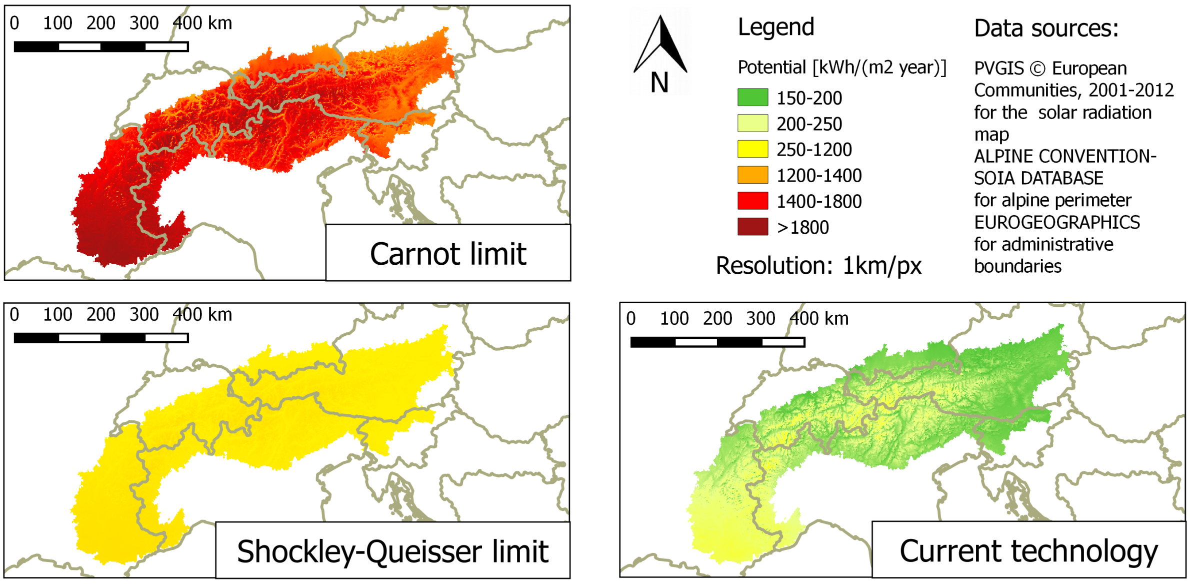


Figure 3: Solar-power potential maps.

Data of wind velocity at different elevation (50m, 70 m, 100m) are provided by the Alpine Space project Alpine Windharvest. Schaffner, Remund, Cattin, & Kunz (2006) use a statistical modeling methodology to compile wind map based on climatological data collected by 592 permanent or temporary meteorological stations. Regarding the accuracy of the data, they test the interpolation for 14 randomly chosen station and the standard deviation of this test is ±1.5ms-1.

The computed values can underestimate the wind speed but big over-estimations do not occur. The resolution is 250m/px.

We assume a Weibull distribution (Celik, 2003) with shape parameter k equal to 1.5 (Casale, 2009). In the Alpine area, difficulties of access limit the size of turbines and we consider only two different elevations, 50m and 70m.
We choose two sample wind turbines with about 800-900 kW rated wind generator (ENERCON). Figure 4 shows the potentials respectively based on the Betz limit and the current technology.

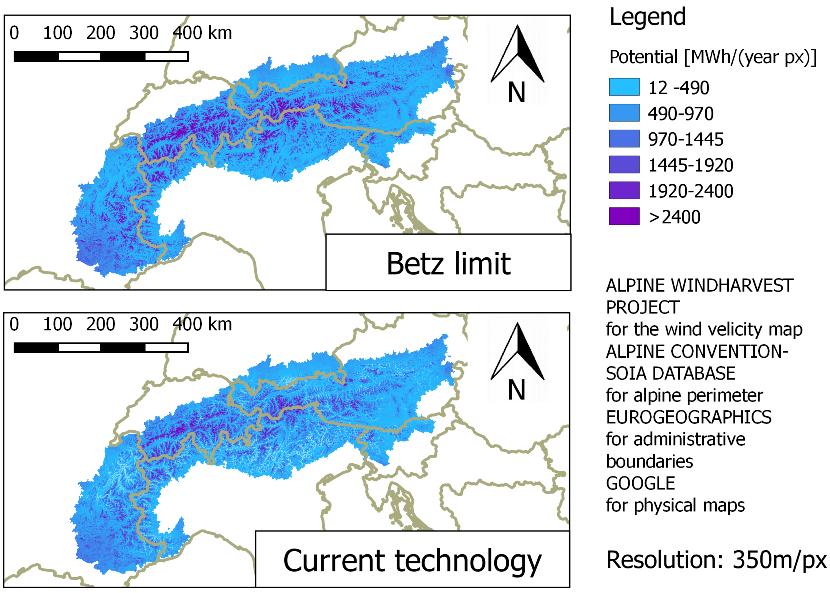


Figure 4: Wind-power potential maps.

4 Results and discussion

The algorithms can be applied to different scales. The quality and the resolution of the input data determine the accuracy of the output data. All the three potential definitions are the starting point to evaluate the residual potential in a region by considering existing plants and different constrains.
Considering wind and hydro potential, Figure 2 and Figure 4, we observe that the difference between the theoretical and the technical potential for wind and hydro-power is not so big. Consequently, an improvement of technology should not improve the energy potential.
Secondly, the elevation above the sea level in a mountain area is related to the population and energy consumption. For this reason, we compare the elevation with the energy potential by considering the current technology. Especially, for wind and solar the full-load equivalent working hours can be derived by knowing the installable power P_{pk} and the annual energy production E. In the case of hydro-power, data about the discharge annual distribution are not available and we plot the mean annual power P versus the elevation of the closure point of each basin.
In figures 5, 6 and 7, we exclude hypothetical plants behind a maximum elevation of 1500m. Above 1500m the costs to set up an energy plant increase considerably In the case of hydro-power we highlight the more productive basins (> 10MW), while for solar and wind-power we consider respectively capacity factor greater the 13% and 20%. The variables are plotted for each

alpine country.

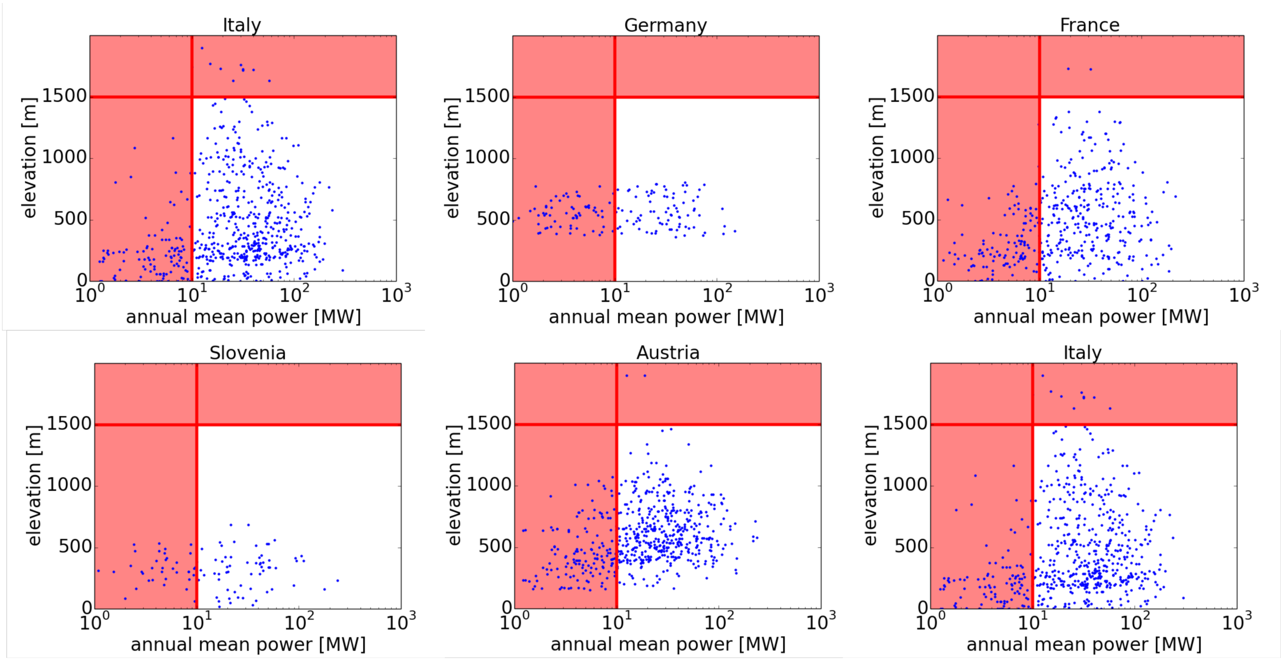


Figure 5: Annual mean power versus elevation in each unit catchment for hydro-power.

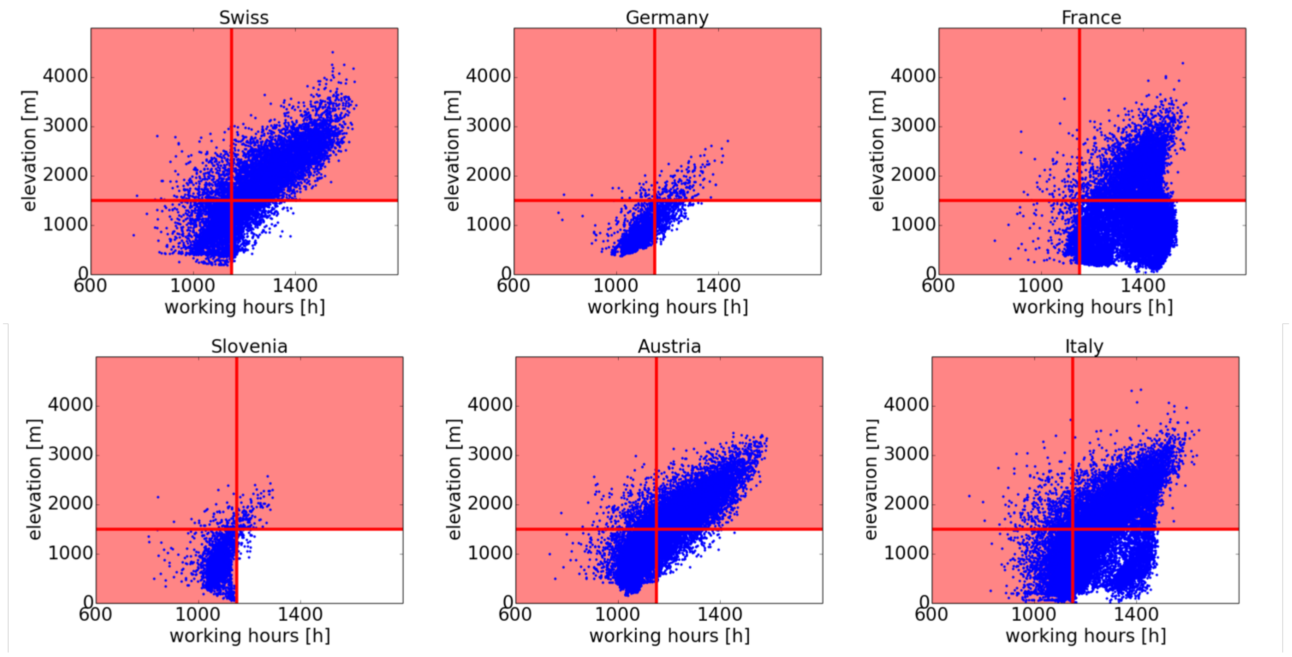


Figure 6: Full-load equivalent working hours versus elevation for solar-power.

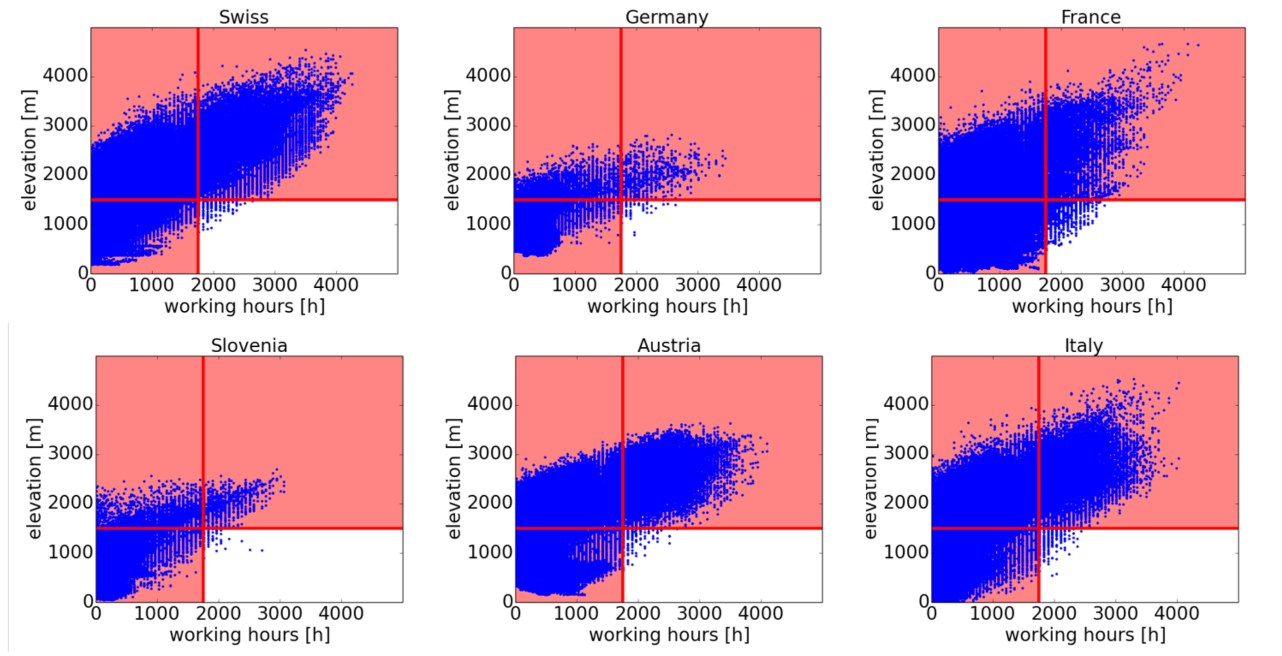


Figure 7: Full-load equivalent working hours versus elevation for wind-power.

5 Conclusion

A clear and coherent definition of the renewable energy potentials is necessary to compute energy potential maps. In this work, the theoretical limit of hydro, solar and wind-power is evaluated for the alpine region. The produced maps show the REs distribution and these data can be analyzed as shown in the previous section in order to find useful information for decision makers and stakeholders.

From the discussed analysis it can be noticed that all the alpine states have a great hydro-power potential also for basins with lower elevation, Figure 5, with no linear correlation between elevation and potential. This can be consideredone of the reasons why hydro power has been the most used renewable energy in the Alps.

In the case of photovoltaic ground system the linear correlation is observed in each alpine country except for Italy and French due to the south exposition of their valley. Indeed, these two countries have the greater solar potential. Finally, for wind a linear correlation between working hours and elevation is observed in each country. The distribution of the potential is then in disagreement with the distribution of the population and the energy needs.

After further testing the final source code used to produce the data and the graphs presented on this work will be available in the r.green module, available in the GRASS-addons repository. The final map of the energy potential for each source analyzed on this article will be published within the INSPIRE standard. To assure a higher visibility of the results to policy and decision makers at a supranational scale.

Next steps of this work will be to consider not only the renewable energy potential variables, but combine them to consider also economic, environmental and social variables.

Acknowledgments

This study was conducted in the frame of the Recharge.Green project “Balancing Alpine Energy and Nature” (http://www.recharge-green.eu), which is carried out within the Alpine Space Programme, and is co-financed by the European Regional Development Fund.The Recharge.Green project is focused on the analysis of how to reconcile biodiversity conservation of ecosystems and renewable energy production.

References

- ✔ Baban, S. M. J., & Parry, T. (2001). Developing and applying a GIS-assisted approach to locating wind farms in the UK. *Renewable Energy*, 24(1), 59–71. doi:10.1016/S0960-1481(00)00169-5.
- ✔ BERGEY, K. H. (1979). The Lanchester-Betz limit (energy conversion efficiency factor for windmills). *Journal of Energy*, 3(6), 382–384. doi:10.2514/3.48013.
- ✔ Bódis, K., Monforti, F., & Szabó, S. (2014). Could Europe have more mini hydro sites? A suitability analysis based on continentally harmonized geographical and hydrological data. *Renewable and Sustainable Energy Reviews*, 37, 794–808. doi:10.1016/j.rser.2014.05.071.
- ✔ Borga, M., Degli Esposti, S., & Norbiato, D. (2006). Influence of errors in radar rainfall estimates on hydrological modeling prediction uncertainty. *Water Resources Research*, 42(8), W08409. doi:10.1029/2005WR004559.
- ✔ Casale, C. A. (2009). Guida per l’utilizzo dell’Atlante eolico dell’Italia (Rapporto aggiuntivo 1 No. 9) (p. 83). Milano - Italia: ENEA – Ricerca sul Sistema Elettrico S.p.A. Retrieved from http://atlanteeolico.rse-web.it/guida_atlante_eolico.pdf.
- ✔ Celik, A. N. (2003). Weibull representative compressed wind speed data for energy and performance calculations of wind energy systems. *Energy Conversion and Management*, 44(19), 3057–3072. doi:10.1016/S0196-8904(03)00075-X.
- ✔ Chatzimouratidis, A. I., & Pilavachi, P. A. (2009). Technological, economic and sustainability evaluation of power plants using the Analytic Hierarchy Process. *Energy Policy*, 37(3), 778–787. doi:10.1016/j.enpol.2008.10.009.
- ✔ De Vries, B. J. M., van Vuuren, D. P., & Hoogwijk, M. M. (2007). Renewable energy sources: Their global potential for the first-half of the 21st century at a global level: An integrated approach. *Energy Policy*, 35(4), 2590–2610. doi:10.1016/j.enpol.2006.09.002.
- ✔ Draxl, C., & Mayr, G. J. (2011). Meteorological wind energy potential in the Alps using ERA40 and wind measurement sites in the Tyrolean Alps. *Wind Energy*, 14(4), 471–489. doi:10.1002/we.436.
- ✔ Ehlschlaeger, C. (1989). Using the AT Search Algorithm to Develop Hydrologic Models from Digital Elevation Data. In Proceedings of International Geographic Information Systems (IGIS) Symposium ’89 (pp.

275–281). Baltimore, MD.

- ✔ Formetta, G., Antonello, A., Franceschi, S., David, O., & Rigon, R. (2014). Hydrological modelling with components: A GIS-based open-source framework. *Environmental Modelling & Software*, 55, 190–200. doi:10.1016/j.envsoft.2014.01.019.
- ✔ Formetta, G., David, O., & Rigon, R. (2012). The OMS3 JGrass-NewAge Environmental Modelling System. AGU Fall Meeting Abstracts, -1, 1165.
- ✔ Green, M. A. (2002). Third generation photovoltaics: solar cells for 2020 and beyond. Physica E: Low-Dimensional Systems and Nanostructures, 14(1–2), 65–70. doi:10.1016/S1386-9477(02)00361-2.
- ✔ Hersch, P., & Zweibel, K. (1982). *Basic Photovoltaic Principles and Methods*. SERI/SP-290-1448.
- ✔ Holmgren, P. (1994). Multiple flow direction algorithms for runoff modelling in grid based elevation models: An empirical evaluation. *Hydrological Processes*, 8(4), 327–334. doi:10.1002/hyp.3360080405.
- ✔ Huld, T., Müller, R., & Gambardella, A. (2012). A new solar radiation database for estimating PV performance in Europe and Africa. *Solar Energy*, 86(6), 1803–1815. doi:10.1016/j.solener.2012.03.006.
- ✔ Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., … Vertačnik, G. (2014). The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *International Journal of Climatology*, 34(5), 1657–1675. doi:10.1002/joc.3794.
- ✔ Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe Version 4. Retrieved from http://srtm. csi. cgiar.org
- ✔ Mari, R., Bottai, L., Busillo, C., Calastrini, F., Gozzini, B., & Gualtieri, G. (2011). A GIS-based interactive web decision support system for planning wind farms in Tuscany (Italy). *Renewable Energy*, 36(2), 754–763. doi:10.1016/j.renene.2010.07.005.
- ✔ Norbiato, D., Borga, M., Merz, R., Blöschl, G., & Carton, A. (2009). Controls on event runoff coefficients in the eastern Italian Alps. J*ournal of Hydrology*, 375(3–4), 312–325. doi:10.1016/j.jhydrol.2009.06.044.
- ✔ Palomino Cuya, D. G., Brandimarte, L., Popescu, I., Alterach, J., & Peviani, M. (2013). A GIS-based assessment of maximum potential hydropower production in La Plata basin under global changes. *Renewable Energy*, 50, 103–114. doi:10.1016/j.renene.2012.06.019.
- ✔ Razykov, T. M., Ferekides, C. S., Morel, D., Stefanakos, E., Ullal, H. S., & Upadhyaya, H. M. (2011). Solar photovoltaic electricity: Current status and future prospects. *Solar Energy*, 85(8), 1580–1608. doi:10.1016/j.solener.2010.12.002.
- ✔ Resch, G., Held, A., Faber, T., Panzer, C., Toro, F., & Haas, R. (2008). Potentials and prospects for renewable energies at global scale. *Energy Policy*, 36(11), 4048–4056. doi:10.1016/j.enpol.2008.06.029.
- ✔ Sacchelli, S., Zambelli, P., Zatelli, P., & Ciolli, M. (2013). Biomasfor: an open-source holistic model for the assessment of sustainable forest bioenergy. i*Forest - Biogeosciences and Forestry*, 6(4), 285–293. doi:10.3832/ifor0897-006.

- ✔ Sauquet, E., & Leblois, E. (2001). Discharge analysis and runoff mapping applied to the evaluation of model performance. Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 26(5–6), 473–478. doi:10.1016/S1464-1909(01)00037-5
- ✔ Schaffner, B., Remund, R., Cattin, R., & Kunz, S. (2006). The Alpine Space Wind Map. Retrieved from http://www.meteotest.ch/fileadmin/user_upload/Windenergie/pdfs/paper_windmap_ewec_06.pdf
- ✔ Schmidt, K.-H., & Morche, D. (2006). Sediment output and effective discharge in two small high mountain catchments in the Bavarian Alps, Germany. *Geomorphology*, 80(1–2), 131–145. doi:10.1016/j.geomorph.2005.09.013.
- ✔ Shockley, W., & Queisser, H. J. (1961). Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *Journal of Applied Physics*, 32(3), 510–519. doi:10.1063/1.1736034.
- ✔ Šúri, M., Huld, T. A., Dunlop, E. D., & Ossenbrink, H. A. (2007). Potential of solar electricity generation in the European Union member states and candidate countries. *Solar Energy*, 81(10), 1295–1305. doi:10.1016/j.solener.2006.12.007.
- ✔ Van Haaren, R., & Fthenakis, V. (2011). GIS-based wind farm site selection using spatial multi-criteria analysis (SMCA): Evaluating the case for New York State. *Renewable and Sustainable Energy Reviews*, 15(7), 3332–3340. doi:10.1016/j.rser.2011.04.010.
- ✔ Zambelli, P., Gebbert, S., & Ciolli, M. (2013). Pygrass: An Object Oriented Python Application Programming Interface (API) for Geographic Resources Analysis Support System (GRASS) Geographic Information System (GIS). *ISPRS International Journal of Geo-Information*, 2(1), 201–219. doi:10.3390/ijgi2010201.

A new open source DSS for assessment and planning of renewable energy: r.green

Giulia Garegnani[1], Francesco Geri[2] , Pietro Zambelli [1-2] , Gianluca Grilli [1-2] , Sandro Sacchelli[1], Alessandro Paletto[3], Giorgio Curetti[1], Marco Ciolli[2], Daniele Vettorato[1]

[1] Eurac Research - Institute of Renewable Energy, Drususallee 1, 39100 Bolzano (Italy)

[2] Department of Civil, Environmental and Mechanical Engineering (DICAM-UNITN), University of Trento, via Mesiano 77, 38123 Trento (Italy)

[3] Agricultural Research Council e Forest Monitoring and Planning Research Unit (CRA-MPF), P.za Nicolini 6, I-38123 Villazzano, Trento, Italy Affiliation

Abstract

The spatially explicit Decision Support System (DSS) r.green, developed in the frame of the Alpine Space program recharge.green financed by EU, is able to identify the areas suitable for the installation of the main renewable energy systems based on criteria of sustainability and land conservation. The DSS is composed by a set of GRASS GIS modules r.green.wind, r.green.hydro, r.green.solar, r.green.biomassfor that deal with each specific renewable energy and r.green.impact that gives feedback on the impacts. Already available as a GRASS add-on, r.green can be used through Grass console or running the standard GUI within Grass or through QGIS.

Keywords

Renewable energy, wind, solar, forest biomass, hydropower, sustainability, Ecosystem services

1 Introduction

In the last decades, the importance of renewable energy is increasing in order to mitigate carbon dioxide emissions and to reduce the fossil fuel dependence of European Member States (Directive 2009/28/EC). Planning energy politics is one of the major task in present times, since energy availability is and will increasingly be crucial for the future well being and development of all the nations (Wirba et al. 2015, Martire et al. 2015, Michels et al. 2010). The growing demand for energy and in particular for renewable energy (RE) will increase the pressure on the environment and natural resources. The global economic crisis is exacerbating the problem especially in mountain areas that often guest fragile remnants of larger natural habitats. Both the demand for renewable energy and the need for the conservation of biodiversity, soil, and connectivity transcend national borders and must be addressed at a transnational level. The recharge.green project was financed by EU in the frame of the Alpine Space projects with the core objective to develop tools and an integrated strategy for renewable energy production, sustainable land use systems as well as the conservation of biodiversity and soil across the Alpine region, so as to support the implementation of relevant EU Directives (recharge.green site). In this frame EURAC, DICAM-UNITN, CRA-MPF developed a DSS model that aims to help decision makers to deal with renewable energy planning. A plethora of DSS (Laquaniti and Sala 2009, Damiano et al. 2010, Wolfslehner and Seidl, 2010, Mari et al. 2011, Brown and Reed, 2012; Schwenk et al., 2012; Herzig et al., 2013; Sánchez-Lozano et al. 2013; Sacchelli et al., 2014; Verkerk et al., 2014; Wang et al., 2014; Sporcic et al., 2011) that deal with different topics is available nowadays in literature. The GIS based DSSs include spatial dimension thus adding a large amount of information and making the results of the DSS more understandable (Tognetti and Cherubini, 2009; Tattoni et al., 2011; Ayanu et al., 2012; Zambelli et al., 2012). Nevertheless, many of these DSS, especially those that deal with environmental issues, are designed to represent in detail a very specific theme and they often lack a multifunctional approach (Sacchelli et al., 2013a). In other cases these DSSs are designed to produce results at a macro-scale (national or macro-regional level) and they represent the phenomena using large pixel units, (square kilometers or more) thus they are not easily usable by local decision-makers and communities (Zambelli et al., 2012, Verkerk et al., 2014). However, the greatest limitation of these DSS is probably due to the fact that very few of them are really publicly available and therefore usable due to license restrictions or to actual availability from the authors (Frombo et al. 2009; Tattoni et al. 2010, Emer et al., 2011; Sacchelli et al., 2013a). In this way only the original authors are capable to reproduce the research that are described in the scientific papers, and this fact limits enormously the possibility to improve the DSS models and to share them. Thus, we developed a model with a multifunctional approach, designed to be applied not only at a macro scale but also at a very local scale and using an Open Source platform to overtake the cited limitations. The r.green model we developed is based on the GIS GRASS (Neteler et al., 2012), and was also developed taking advantage of python libraries and pygrass (Zambelli et al., 2013) and of QGIS to facilitate the creation of an user friendly interface. The software r,green was released as an add on of GRASS and is already available to be tested and improved, respecting the open source licences and giving proper credit to the original authors (Steinigera and Hay, 2009).

2 The model

The holistic spatially explicit Decision Support System (DSS) r.green is able to identify and quantify the areas suitable for the installation of renewable energy systems based on criteria of sustainability and land conservation. The aim of the creators and of the developers is to highlight Alpine biodiversity, land use patterns and related ecosystem services, and model the carrying capacity of the Alpine ecosystems with respect to all aspects of RE production and consumption.

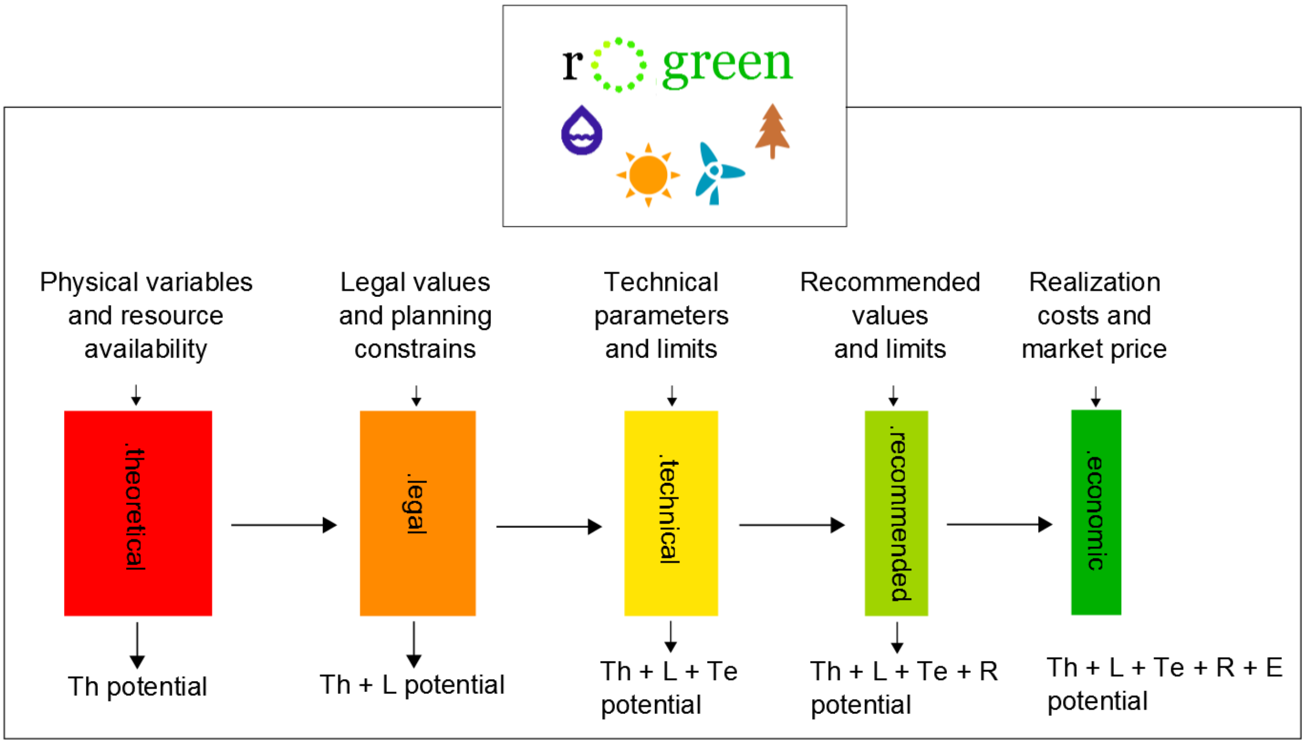


Fig. 1 : Sketch of the inputs and outputs of a possible scenario obtained with r.green. The logic steps that constitute the structure of the model are shown. Each sub-model follows the same logic structure.

The kernel of r.green resides within the GIS GRASS, and is composed by a set of add-on (additional modules) that can be run independently. Four main modules of the DSS, r.green.wind, r.green.hydro, r.green.solar and r.green.biomassfor deal with each specific renewable energy while the last one, r.green.impact, gives feedback on the impacts on the ecosystem services. Each module is structured by a set of sub-modules that represent a series of operational steps in an ideal flow of operations related to the level of exploitation considered. Each main module is composed by five sub-modules that were named theoretical, legal, technical, economic and recommended, as well as various modules specifically oriented to the analysis of impacts and assessment of the ecosystem services. The sub-module theoretical calculates the theoretical potential of the selected energy. The sub-module legal introduces legal constraints derived by plans or guidelines, the sub-module technical takes into account technical limits, the sub-module economic considers the economical convenience of the intervention and finally the sub-module recommended introduces the recommended parameters. The modules must be provided with a set of mandatory variables and a set of optional values. The more complete and accurate is the list of variables provided, the better the results.
All the run methodologies were used in four pilot areas: the Mis and Mae valley in the north east of Italy, the Triglav National Park in Slovenia, Leiblachtal in

Voralberg and the Gesso and Vermenagna Valley in the north west of Italy.

2.1 r.green.wind

Wind energy can help to cover the main peaks of energy consumption during the day. (Baban & Parry, 2001; Draxl & Mayr, 2011; Mari et al., 2011; van Haaren & Fthenakis, 2011) build a decision support system starting from wind velocity maps. The modules r.green.wind computes the energy potential starting from wind distribution functions and the power curve of available turbines (Draxl & Mayr, 2011 and Mari et al., 2011). The theoretical module considers the maximum limit to the amount of energy that can be converted in power, the so-called Betz limit (Bergey, 1979). The technical module needs the features of the wind-turbine (i.e. rated power, rotor diameter, hub heght, etc...). Output of the module are raster with the energy information (Mwh). The cost-benefit analysis, legal and environmental constrains can be considered in the other modules.

2.2 r.green.hydro

In the Alpine region, the main renewable energy actually used is hydro-power. Some authors ((Kusre et al., 2010; Palomino et al., 2013)) asses hydro-power potential using GIS. While Directive 2009/28/EC underlines the importance of renewable energies, one of the priorities of the Water Frame Directive (Directive 2000/60/EC) is the water protection. For this reason, the software r.green.hydro considers legal, technical, economic and sustainable principles of both the directives in order to evaluate the residual potential of rivers under different scenarios. Inputs of this module are raster files with discharge data along river network, digital terrain elevation model and vector files with existing plant position. The outputs of the model are two different vector files with theoretical available river segments, optimal position of the plants with their powers and their intakes and restitutions. In the legal module, we consider data on the minimum vital flows, the protected area, the maximum plant capacity, etc. Finally, we can derive the economic potential, i.e the market and the realization cost. This module let us to rank the feasibility of the plants on the bases of the cost. Notice that we consider all other geographical constrains (steepest area, distance from electricity grid, etc.) in the cost calculation.

2.3 r.green.solar

In the theoretical case we consider the Shockley & Queisser limit and the Carnot limit , while in the module r.green.solar.technical the efficiency of the solar cells can be added. The electricity is then computed according to european norm EN 15316-4-6. Mandatory input of the model can be the irradiation map, if not available it can be computed by using the GRASS module r.sun (Neteler and Mitasova 2008).
The different land use can be considered in the r.green.solar.reccomended in order to exclude area of particular interest. Also in this case, the economic module perform a cost-benefit analysis.

2.4 r.green.biomassfor

The characteristics of the Alpine region and the present forest landscape dynamics suggest that the exploitation of wood-energy sources can be extremely important for bioenergy production. The module r.green.biomassfor is able to quantify in MWh/y the potential bio-energy exploitable from wood biomass in forest ecosystems in the light of ecological and economic sustainability. It was developed as an evolution of Biomasfor, the model that also inspired the structure of the other r.green sub-models (Zambelli et al. 2011, Zambelli et al 2012, Sacchelli et al 2013). The multi-step approach and the model's internal structure permit the use of heterogenous input dataset. To run the model a series of mandatory variables is required and the results can be utterly refined inserting a series of optional variables. The r.green.biomassfor considers theoretical, legal, technical, economic and sustainable principles to evaluate the energetic potential. The model calculates spatially explicit scenarios represented as maps and tabular data that can be queried and exported to other GIS and DSS models. The user can interactively change input data and/or variables (like for example mechanization level or chip wood price) thus producing different scenarios. The model can produce an estimate of a CO2 emission and other multi-functionality parameters, such as fire risk and recreational evaluation.

2.4 r.green.impact

The present model contains also a module for estimating the economic impact of energy withdrawal on the ecosystem services (ESs). The overall objective of this module is to include the environmental costs and benefit into the computation of the exploitable energy quantity, in order to carry out cost-benefit analysis with environmental externalities. More precisely, the module considers the following ESs: timber and other wood products, carbon sequestration, hydrogeological protection and recreation. The fundamental input of the module is a vector file with the spatial definition of ESs economic values, typically a column for each different ES. The module automatically creates a raster map for each ES, based on the column´s name of the attribute table and calculates the change in the ES value considering a percentage of variation, which is different for each raster. This percentage can be edited by the user, based on the local characteristics and, if lacking, a default percentage is used.

2.5 Software interface

As declared in the software description, r.green is made by many modules and sub-modules, requires the input of many variables and its application may be very complex. Thus, the development of a graphic interface to open the software usability to non-specialist users was required.

To facilitate this step we developed three possible ways to launch r.green through different interfaces and relative GIS:

i) through the link command of Grass console, recommended for advanced users

ii) running the standard GUI within Grass, developed using the Wxpython graphic libraries

iii) through a plugin of QGIS developed using PyQt and the Qt graphics libraries

A Graphic User Interface has been developed as a plugin of QGIS, a very user friendly open-source GIS software characterized by a logical structure very

similar to the most widely used GIS proprietary software. The QGIS plugin processes and exports outputs in standard formats (shapefiles for vector data, ASCII and GeoTiff for raster data).
Anyway to obtain the best and more reliable results our advice is to read the manual and to read the reports and the papers cited in the software description.

3 Tests

All the models were run in all the pilot areas but the most accurate tests were carried out in each pilot area selecting only the suitable energies.
The model r.green.hydro was mainly used in two pilot areas: the Mis and Mae valleys in the north east of Italy and the Gesso and Vermenagna valleys in Alpi Marittime Natural park in the north west of Italy. Different scenarios were performed., outputs of the module r.green.hydro are maps with energy production and relative costs. The maps were discussed in several meetings with stakeholders in order to plan the area through a participative approach. The module is also under tesitng by ARPA and Eaux Valldotanes for the Regione Autonoma Valle d'Aosta.

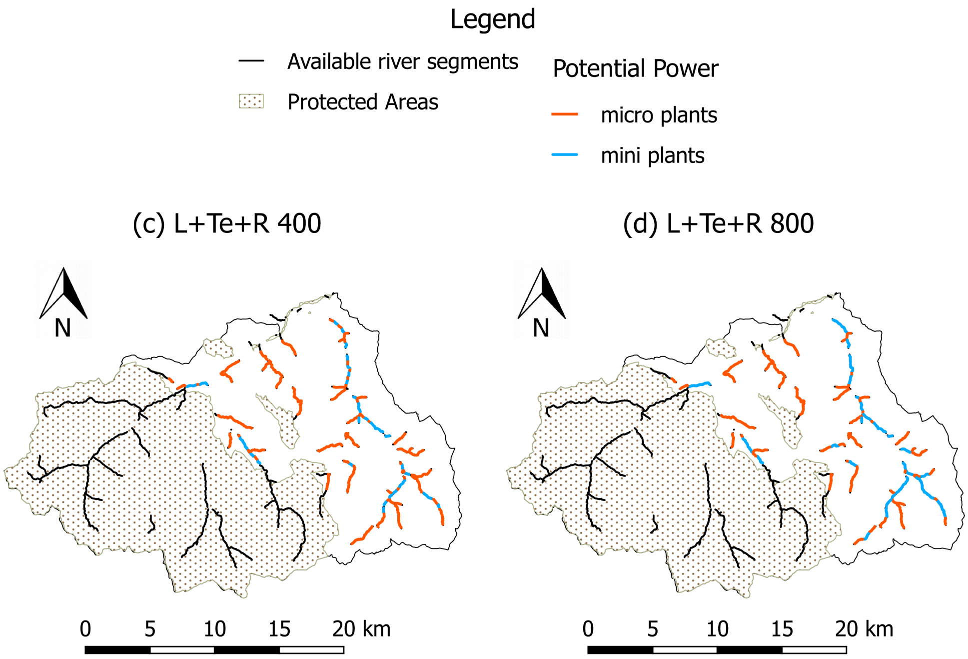


Fig. 2: Two different scenarios for hydro-power potential in Val di Gesso and Vermenagna (Piedmont, Italy) with different sets of plant lenghts

The models r.green.solar and r.green.wind were mainly used in pilot area of Vorarlberg/Leiblachtal, Austria. The region borders Germany in the North, Lake Constance in the West, and in the East and in the South the mountain Pfänder (1064 meters) forms a natural border.
The model r.green.biomassfor was used mainly in three pilot areas: the Mis and

Mae valley in north east Italy, The Alpi Marittime Natural park in north west Italy, and in the Triglav National Park in Slovenia. In Mis and Mae valley and in Alpi Marittime Natural park different scenarios with different wood chip prices combined with existing/planned bioenergy plants location were produced and showed to Regional administrators in different public events. This started a discussion on the optimization of bioenergy plants and in general on biomass opportunities. The model was used in meetings with local stakeholders of the Valleys. In Triglav National Park, different biomass production scenarios were created and a comparison with the Wisdom model previously used by the park managers to produce scenarios was carried out. Further details regarding the results of r.green processing in different Pilot areas are available in the documentation at the recharge.green site (http://www.recharge-green.eu/) and some of the results will be soon available for queries through the Jecami platform (http://www.jecami.eu/).

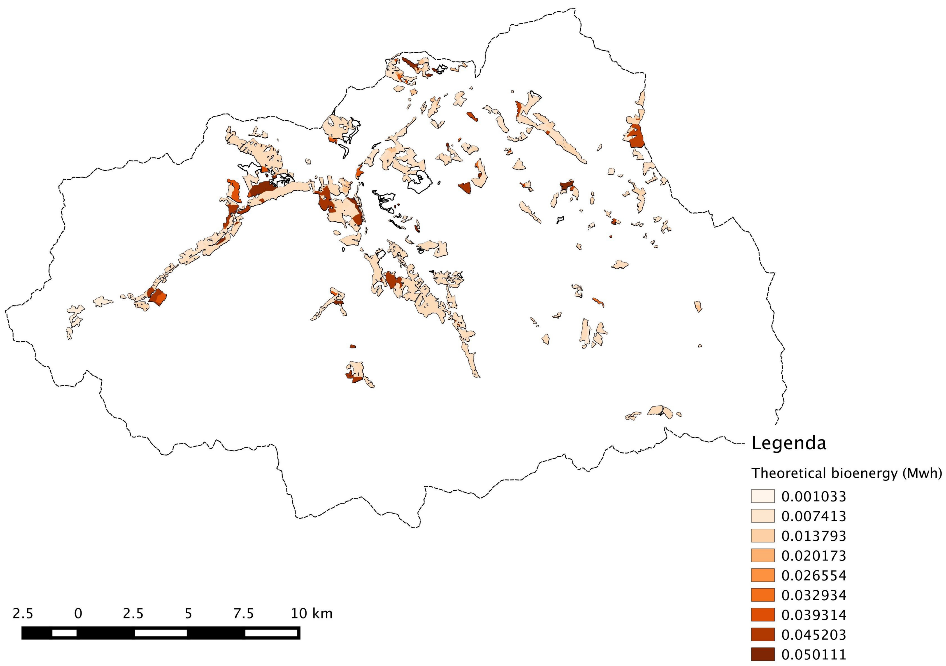


Fig. 3: A potential forest biomass scenario in in Val di Gesso and Vermenagna Piedmont (Italy), showing forest biomass residues potential in MWh per year. The potential is spatially explicit, this means that is calculated for each pixel.

4 Conclusions

The holistic spatially explicit Decision Support System (DSS) r.green is a very ambitious software. The first tests carried out in the frame of recharge.green project were more than encouraging, since the results of the model and of the submodels were used immediately with local stakeholders and decision makers. The models fostered discussions and produced scenarios that were judged plausible and useful. Thus the model is already usable as it is. However, the complexity of the topics treated by the model and consequently the complexity of the software suggest that there is a large space for improvement, both on the point of view of logical procedures and on the point of view of software performances. The module r.green is available as an easy to install GRASS add-on and moreover it can be used through Grass or through the more user friendly QGIS software, thus facilitating usability for less experienced GIS user. The software is released with an open-source license to encourage further development and to spread and share knowledge and science.

Aknowledgments

The study described in this work was carried out within the Recharge.Green project “Balancing Alpine Energy and Nature” (http://www.recharge-green.eu), which is part of the Alpine Space Programme, and is co-financed by the European Regional Development Fund. The Recharge.Green project is focused on the analysis of how to reconcile biodiversity conservation of ecosystems and renewable energy production. We want to thank all the Pilot areas that tested the models and helped us to tune it, our thanks to Erica Zangrando, Francesca Miotello, Simone Bertin, Ales Poljanec, Rok Pisek, Luca Giraudo. The first version of the model Biomasfor was developed in the BIOMASFOR project co-funded by the CARITRO Foundation through grant No. 101.

References

✔ Ayanu, Y.Z. , Conrad, C., Nauss, T., Wegmann, M., Koellner, T., 2012. Quantifying and Mapping Ecosystem Services Supplies and Demands: A Review of Remote Sensing Applications. Environmental Science & Technology 46, 16, 8529-8541. doi: 10.1021/es300157u.

✔ Baban, S. M. J., & Parry, T. (2001). Developing and applying a GIS-assisted approach to locating wind farms in the UK. Renewable Energy, 24(1), 59–71. doi:10.1016/S0960-1481(00)00169-5

✔ Bergey, K. H. (1979). The Lanchester-Betz limit (energy conversion efficiency factor for windmills). Journal of Energy, 3(6), 382–384. doi:10.2514/3.48013

✔ Brown, G., Reed, P., 2012. Values Compatibility Analysis: Using Public Participation Geographic Information Systems (PPGIS) for Decision Support in National Forest Management. Applied Spatial Analysis and Policy 5, 4, 317-332. doi: 10.1007/s12061-011-9072-x.

✔ Damiano, A., Gatto, G., Marongiu, I. 2010 Decision Support System for biomass energy exploitation in smart grid planning . SPEEDAM 2010 - International Symposium on Power Electronics, Electrical Drives, Automation and Motion, art. no. 5542256, pp. 1183-1188.

✔ Draxl, C., & Mayr, G. J. (2011). Meteorological wind energy potential in the Alps using ERA40 and wind measurement sites in the Tyrolean Alps. Wind Energy, 14(4), 471–489. doi:10.1002/we.436

✔ Emer, B., Grigolato, S., Lubello, D., Cavalli, R., 2011. Comparison of biomass feedstock supply and demand in Northeast Italy. Biomass & Bioenergy 35, 8, 3309-3317. doi: 10.1016/j.biombioe.2010.09.005.

- Frombo, F., Minciardi, R., Robba, M., Sacile, R. 2009. A decision support system for planning biomass-based energy production. Energy 34, 362-369. doi: 10.1016/j.energy.2008.10.012.
- Grilli G., Curetti G., De Meo I., Garegnani G., Miotello F., Poljanec A., Vettorato D., Paletto A. 2015. Experts' perceptions of the effects of forest biomass harvesting on sustainability in the Alpine region. South-East European Forestry 6(1): e1-e9.
- Grilli, Gianluca, Alessandro Paletto, and Isabella De Meo. 2014. "Economic Valuation of Forest Recreation in an Alpine Valley." Baltic Forestry 20(1): 167–175.
- Herzig, A., Ausseil Anne-Gaelle, E., Dymond, J.R.,2013. Spatial optimisation of ecosystem services. In: Dymond, J.R. (Ed.), Ecosystem Servicesin New Zealand: Conditions and Trends, pp 511-523.
- Kusre, B. C., Baruah, D. C., Bordoloi, P. K., & Patra, S. C. (2010). Assessment of hydropower potential using GIS and hydrological modeling technique in Kopili River basin in Assam (India). Applied Energy, 87(1), 298–309. http://doi.org/10.1016/j.apenergy.2009.07.019
- Lacquaniti, P., Sala, S. 2009 Energy from biomass: Decision support system for integrating sustainability into technology assessment. WIT Transactions on Ecology and the Environment, 121, pp. 291-302.
- Mari R., Bottai L., Busillo C., Calastrini F., Gozzini B., Gualtieri G., (2011) A GIS-based interactive web decision support system for planning wind farms in Tuscany (Italy) Renewable Energy 36, 754 e 763
- Martire, S., Castellani, V., Sala, S. 2015 Carrying capacity assessment of forest resources: Enhancing environmental sustainability in energy production at local scale. Resources, Conservation and Recycling, 94, pp. 11-20.
- Michels, L., El-Gayar, O., Wills, M. 2010 Solar Wind and Energy Resource Assessment (SWERA): A usability case study . Modelling for Environment's Sake: Proceedings of the 5th Biennial Conference of the International Environmental Modelling and Software Society, iEMSs 2010, 3, pp. 2541-2548.
- Neteler, M., Bowman, M.H., Landa, M., Metz, M., 2012. GRASS GIS: A multi-purpose open source GIS. Environmental Modelling and Software 31, 124-130. doi: 10.1016/j.envsoft.2011.11.014.
- Neteler, M.; Mitasova, H. (2008). Open Source GIS : a GRASS GIS approach, 3rd edition. New York: Springer. ISBN 978-0-387-35767-6.
- Palomino Cuya, D. G., Brandimarte, L., Popescu, I., Alterach, J., & Peviani, M. (2013). A GIS-based assessment of maximum potential hydropower production in La Plata basin under global changes. Renewable Energy, 50, 103–114. http://doi.org/10.1016/j.renene.2012.06.019
- Pang X., Mörtberg U., Brown N., 2014 Energy models from a strategic environmental assessment perspective in an EU context-What is missing concerning renewables?, Renewable and Sustainable Energy Reviews, Vol. 33, pp 353-362, http://dx.doi.org/10.1016/j.rser.2014.02.005.
- Sacchelli, S., Bernetti, I., De Meo, I., Fiori, L., Paletto, A., Zambelli, P., Ciolli, M., 2014. Matching socio-economic and environmental efficiency of wood-

residues energy chain: a partial equilibrium model for a case study in Alpine area. Journal of Cleaner Production 66, 431-442. doi:10.1016/j.jclepro.2013.11.059.

- Sánchez-Lozano, J.M., Jiménez-Pérez, J.A., García-Cascales, M.S., Lamata, M.T. 2013 Decision support systems to obtain decision criteria by fuzzy AHP for Location of Renewable Energy Facilities IJCCI 2013 - Proceedings of the 5th International Joint Conference on Computational Intelligence, pp. 300-308.
- Schwenk, W.S., Donovan, T.M., Keeton, W.S., Nunery, J.S., 2012. Carbon storage, timber production, and biodiversity: comparing ecosystem services with multi-criteria decision analysis. Ecological Applications 22, 5, 1612-1627.
- Sporcic, M., Landekic, M., Lovric, M., Martinic, I., 2011. Planning and Decision Making Models in Forestry. Croatian Journal of Forest Engineering 32, 1, 443-456.
- Steinigera, S., Hay, G.J., 2009. Free and open sourcegeographic information tools for landscape ecology. Ecological Informatics 4, 183-195. doi:10.1016/j.ecoinf.2009.07.004.
- Tattoni, C., Ciolli, M., Ferretti, F., 2011. The fate of priority areas for conservation in protected areas: a fine-scale Markov chain approach. Environmental Management 47, 2, 263-278. doi: 10.1007/s00267-010-9601-4.
- Tattoni, C., Ciolli, M., Ferretti, F., Cantiani, M.G., 2010. Monitoring spatial and temporal pattern of Paneveggio forest (northern Italy) from 1859 to 2006. iForest 3, 72-80. doi: 10.3832/ifor0530-003.
- Tognetti, R.; Cherubini, P., 2009. Change is in the air: future challenges for applied forest research. iForest 2, 56-58. doi: 10.3832/ifor0492-002.
- Verkerk, P.J., Mavsar, R., Giergiczny, M., Lindner, M., Edwards, D., Schelhaas, M.J., 2014. Assessing impacts of intensified biomass production and biodiversity protection on ecosystem services provided by European forests. Ecosystem Services 9, 155-165. doi: 10.1016/j.ecoser.2014.06.004.
- Wang, Q., M'Ikiugu, M.M., Kinoshita, I., 2014. A GIS-based approach in support of spatial planning for renewable energy: A case study of Fukushima, Japan. Sustainability 6, 4, 2087-2117. doi: 10.3390/su6042087.
- Wirba A.V., Mas'ud A.A., Muhammad-Sukki F., Ahmad S., Mat Tahar R., Abdul Rahim R., Munir A.B., Ershadul Karim M.D., 2015 Renewable energy potentials in Cameroon: Prospects and challenges, Renewable Energy, Vol 76, pp 560-565, http://dx.doi.org/10.1016/j.renene.2014.11.083.
- Wolfslehner, B., Seidl, R., 2010. Harnessing Ecosystem Models and Multi-Criteria Decision Analysis for the Support of Forest Management. Environmental Management 46, 6, 850-861. doi: 10.1007/s00267-009-9414-5.
- Zambelli, P., Lora, C., Ciolli, M., Spinelli, R., Tattoni, C., Vitti, A., Zatelli, P., 2010. A foss4g model to estimate forest exploitation methods and biomass availability for renewable energy production. In: FOSS4G 2010 Selected Presentations. Proceedings of FOSS4G 2010, Barcelona, 6e9 September 2010, pp. 1e17. URL: http://2010.foss4g.org/papers/3405.pdf.

- ✔ Zambelli P, Lora C, Spinelli R, Tattoni C, Vitti A, Zatelli P, Ciolli M (2012). A GIS decision support system for regional forest management to assess biomass availability for renewable energy production. Environmental Modelling and Software 38: 203-213.
- ✔ Zambelli, P., Gebbert, S., Ciolli, M., 2013. Pygrass: An Object Oriented Python Application Programming Interface (API) for Geographic Resources Analysis Support System (GRASS) Geographic Information System (GIS). ISPRS International Journal of Geo-Information 2, 201-219. doi:10.3390/ijgi2010201 www.mdpi.com/journal/ijgi.

How far do Dutch people live from attractive nature? An assessment using parallel computing with Python and FOSS4G libraries

Bob Dröge, Leon van der Meulen, Govert Schoof

Research and Innovation Support, Centre for Information Technology, University of Groningen

Abstract

How valuable is living nearby nature? Does nature have a positive effect on nearby residential property prices? How much are we willing to pay for nature in our living environment, and does this amount decay with distance to nature? Increasing urbanization and stress on natural landscapes makes such questions more and more important in spatial planning. However, quantifying the value of public green space is challenging, especially for large study areas, because of the required high computing power. In a recent conference paper by Daams et al. (2014), over 200.000 (!) individual properties across the Netherlands were analyzed to give insight into the Dutch people’s willingness to pay for living near highly attractive public nature. Unlike existing studies of such kind, not only the relation between property prices and the most nearby nature, e.g. within 1 or 2 kilometer, was analyzed, as effects from the quantity of *attractive* nature up to 10 kilometers away were evaluated in the initial research process. That analysis required comprehensive and highly detailed spatial data, as the areas of the all natural land use polygons, with many vertices per feature, needed to be summed for each of the 200,000 properties separately. The required resources to do so far exceeded those that a single computer, even with heavy specifications, could provide. In this paper we discuss our solution to this problem that Daams et al. (in prep.) encountered: parallel computing with Python and FOSS4G libraries. More specific, we describe how we supported this project by specifying and applying several python scripts and libraries, and running these on our high performance cluster.

Keywords

Parallel computing, Nature, Property value, Big Data

1 Data

In the research by Daams et al. (in prep.), attractive natural spaces were designated with point data from the Hotspotmonitor (see also Sijtsma & Daams, 2014). The Hotspotmonitor (http://en.hotspotmonitor.eu) is an online survey using which over 13.000 people in the Netherlands, Germany, and Denmark have ’marked’ attractive natural places and explained why (University of Groningen, 2014) However, these point-data do not comprise

information about the quantity of attractive natural space, so these were combined with land use data. Our important contribution to these data preparations was the measurement of the land use area of both attractive and non-attractive nature surrounding each of the 200,000 observed properties. As it was initially not clear across which distance attractive nature would be found to impact on property prices, the total area of (attractive) natural space was measured in distinct rings with different radii (1km, 2km, ..., 10km).Computing the total area of natural space within multiple rings surrounding each property involves a huge number of calculations. In order to speed this up and be able to run this computation in an efficient way on our high-performance computing cluster consisting of over 3000 cores, we implemented a Python application that can do the computation in parallel: there are no dependencies between the computations for different properties, which makes it an embarrassingly parallel problem.
Our implementation makes use of the Shapely module (based on the GEOS and JTS libraries) for the geometrical functionality and the R-tree module (based on libspatialindex) for spatial indexing functionality. The following versions of software and libraries have been used:

- Python 2.7.4
- Rtree 0.7.0
- libspatialindex 1.8.1
- Shapely 1.2.18
- GEOS 3.4.2
- pyshp 1.2.0

The implementation of the Python application can be described in pseudocode as follows:

```
    Read the shapefile containing the natural spaces
    Read the shapefile containing the residential properties
    Build an R-tree based on the bounding boxes enclosing the natural spaces
    Start a pool of parallel worker threads
    For each property to be processed, let one worker thread do the following:
        Create a buffer of a given radius around the property
        Query the R-tree for natural spaces that possibly intersect with the buffer
        For each of these natural spaces:
            Calculate the area of the intersection with the buffer
            Add the calculated area to the sum for this particular property
    Write the calculated sums for all properties to a CSV file
```

The way of parallelizing the problem allows the program to make use of all the cores on a single machine in an efficient way. In order to scale even beyond a single machine and make even better use of our high-performance computing facilities, we added extra input parameters to the application for specifying a range of properties to be processed. Each range is independent of all the others and can therefore be processed on a different machine, significantly decreasing the runtime of the application even further.

References

- ✔ Daams, M.N., Sijtsma, F.J., Van der Vlist, A.J. (in.prep.) The effect of natural space on property prices: accounting for perceived attractiveness. Presented at the 2nd Workshop on Non-Market Valuation, Aix-Marseille, France.
- ✔ Sijtsma, F., & Daams, M. (2014). How near are urban inhabitants to appreciated natural areas? An exploration of Hotspotmonitor based well-being indicators. Results for the Netherlands, Germany, and Denmark.: Report at the request of the OECD; supporting the How's life in your region? study. (348 ed.) Groningen: Urban and Regional Studies Institute / University of Groningen.
- ✔ University of Groningen (15-10-2014) *Unifocus 19:De Hotspotmonitor* [Video file] Retrieved from https://www.youtube.com/watch?v=0M9lMFQwrgg

Open Web Services: new tools for Medievalist Historians to manage and share their research work

Daniela Carrion[1], Federica Migliaccio[1], Guido Minini[1], Cynthia Zambrano[2]

[1] Department of Civil and Environmental Engineering (DICA), Politecnico di Milano,

Piazza Leonardo Da Vinci 32, Milano, Italy

[2] E-VIA S.p.a, Retelit Group, Milano, Italy

Abstract

The research project presented in this paper regards a joint cooperation between Medievalist Historians and Geomatics experts. The idea was to study if new tools based on Geomatics technologies could provide Historians with new methods to develop their studies, archiving digital data in a geo-database and “spatializing” the information sources to produce maps in a GIS environment. So, in the first part of the project, a GIS was designed and implemented based on data collected from documents preserved in the Italian National Archives. Afterwards, the problem of data integration and sharing among research groups working on Medieval History was dealt with, in order to make data available for the consultation and query by several research groups. Three approaches based on a client-server architecture have been explored: one is typical of the WebGIS architecture; the second one is based on OGC Web services and the third one exploits a Web page while the GIS tools are provided by a Desktop GIS installed locally on a PC. In the paper, the different approaches will be described, in order to underline advantages and disadvantages with respect to the Historians' requirements.

Keywords

Historical GIS, Medieval Geodata, Web Service, FOSS, QGIS

1 Introduction

The research project presented in this paper is the result of many years of research and cooperation between Geomatics experts and Medievalist Historians. The main aim of the cooperation was to provide the Historians with a modern tool to manage, visualize and share their sources exploiting Geographic Information Systems. The first phase of the work consisted in understanding the characteristics of the Medieval data to be organized in a GIS environment, in order to design a suitable database structure (Carrion, Migliaccio, Minini, & Zambrano, in press) for the GIS. This project has been dealt with by using a cognitive approach, aimed at identifying the key issues and the appropriate tools to help the Historians in satisfying their research

needs.
At first, data collected from historical sources have been organized into table archives; these data have been associated with spatial information (point coordinates), corresponding to the place names found in historical sources: knowing the spatial reference it is possible to import data into a Geographic Information System (GIS), to display their location on the area of interest and to overlay them with existing base maps. The data used for the project date back to the late Middle Ages (mid-Fifteenth Century) and represent fiscal information, related to various kinds of taxes paid to different Kings or Princes in the Kingdom of Naples. Then, a relational database has been designed and populated with the data. The relational database design has been achieved through the definition of the Entity-Relationship Diagram (ERD).
In the second phase of the project, that is the main focus of this paper, the problem to be resolved was how to allow Historians to share and query their data exploiting the Web. Three kinds of approaches have been explored. The first one is typical of the WebGIS architecture, in which all the data and the GIS analysis tools are stored in a server and published on the Web, so that the only client-side software requested is a Web Browser. The second one is a Web service, compliant with the OGC standards (http://www.opengeospatial.org/standards), namely a Web Mapping Service, (WMS), a Web Feature Service (WFS) or a Web Processing Service (WPS). The third approach is again based on a client-server architecture, but in this case only the data, collected into a DB and managed through a Database Management System (DBMS), are stored in a server, while the GIS tools are provided by a Desktop GIS installed locally on a PC.
In the next chapters these approaches will be discussed in more detail, with the purpose of understanding which of them can be more effectively applied in the case of the proposed historical application, taking into account the specific requirements of this project.
Regarding the GIS technologies to be exploited, it has been decided to use only Free Open Source Software: this goal raises from the need to develop and maintain the system for a long period with limited economical resources.
Another important objective pertains the ease of use of the GIS, because in this instance the end users are not GIS experts, so the system must be complete but also simple and intuitive to use.

2 State of the art

GIS potentiality is spreading to many disciplines, including History (Schlichting, 2008). At the international level, many works have to be mentioned, see e.g. (Gregory, 2002), (Gregory, Bennet, Gilham, & Southall, 2002), (Berman, 2005) and (Gregory & Healey, 2007), being the main references for this kind of research.
Many examples of historical data organized into national atlases can be found, see e.g. (Pawson, 1997) and (Black, 2003); in some cases data are structured into relational databases, to be managed into a GIS environment, see e.g. (Ardissone & Rinaudo, 2005); in most cases they are related to census data and rarely they date before the XVI century: (Boonstra, Collenteur, & van Elderen, 1995), (De Moor & Wiedemann, 2001), (Fitch & Ruggles, 2003).

The possibility to share georeferenced data on the Internet (Tait, 2004) is precious to improve the cooperation among scientists. It is possible to find some examples of historical data organized into a GIS and published on the Internet, such as the China Historical GIS (http://www.fas.harvard.edu/~chgis/) or the northern Italy cadastral map WebGIS (Brovelli, Minghini, Giori, & Beretta, 2012). Another interesting example published on the Web is constituted by the “E 179 database”, (http://www.nationalarchives.gov.uk/e179/) which contains records relating to lay and clerical taxation and which is included into the United Kingdom National Archives. However, in this case the geographic component is not made explicit into a map.
It must be also mentioned that, when publishing geodata on the Web, the attention to interoperability is essential (see Maguire and Longley, 2005; Johnson et al., 2011).

3 The historical data

One of the aims of the project is to create a GIS that can be useful for Historians to draw maps from historical data in order to study the territorial dynamics emerging from Medieval sources, achieving a cartographic representation of the spatial distribution of historical information.
Currently three historical sources have been collected into the DBMS:

- the *Liber Focorum Regni Neapolis*, produced by the financial offices of Alfonso V of Aragon in the 1440s, after his conquest of the Neapolitan kingdom;
- the *Quaterni declaracionum*, produced between 1446 and 1463 by the leading Treasury officials (the *magistri rationarum*) of the last Prince of Taranto, Giovanni Antonio Orsini del Balzo, the greatest feudal vassal of the kingdom (Pizzuto, 2009);
- the *Quaternus decimarum*, drawn up in 1478 by Paolo Vassallo, bishop of Aversa, that registers the payments of the *Decima*, a tax related to ecclesiastical benefits (Mangia, 2013).

From the data collected by Historians, several entities have been defined following the rules of relational databases. A fundamental entity was the one referring to the historical place names, which could then be connected with present-day place names, allowing for the geo-referencing of information by means of map coordinates in UTM WGS84 mapping and reference system. The geo-referencing of the location has been performed by the Historians with the support of the Geomatics experts. Another crucial entity for Historians is the source, which must be always linked to the data stored into the database. Besides, entities regarding census data, types and amount of taxes due and collected, historical characters names and functions, and other, have been defined for the purposes of the project.
Historians' know how was fundamental during the phase of extraction of data from historical archives and translation of the information into digital format (typically into a spreadsheet). Also the conceptual model of the database has been planned and designed in cooperation with Historians, in order to understand their needs and the meaning of the collected information (in particular the preparation of a data dictionary was necessary).

The possibility to easily visualize on a GIS-based map the locations linked to the data collected from the historical sources is a remarkable achievement for Medievalist Historians. From the Geomatics point of view the spatial component of the information is really simple: once the Historians have defined the location corresponding to the present placement, it translates into point coordinates. What is challenging is to allow the scholars with an efficient tool to manage the complexity of tables which are linked to those coordinates with the relationships designed in the ERD.

4 Data publication on the Web

Also during the discussion on issues related to the publication of data the participation of Historians was again fundamental, because they represent the end users of the GIS, so it is important to understand what tools can be useful for them. The database design and the historical sources organization into a GIS has been presented in Carrion et al. (in press).

For easy data sharing among the Medievalists, the choice of the Web architecture is very important, considering the elements which are crucial in this peculiar case:

i) reduced costs for software acquisition, system implementation and maintenance;

ii) ease of use for non expert users;

iii) ease of maintenance over time;

iv) scalability, ease of integration of new data.

As mentioned in the previous paragraph, the data considered in this project show their complexity in the database structure. So, another key point is to allow the management of the relationships among the database tables.

We can then summarize the system requirements as follows:

- exploitation of Open Source Software, in order to comply with the low budget available;
- access to the database tables, including the ones not containing georeferenced data;
- access to the relationships between entities;
- possibility to display the results of the table queries on maps and vice-versa;
- user friendliness;
- simplicity (to reduce the resources needed both for the system development and its maintenance).

In the following three possible approaches for Medieval data sharing through the Web will be presented and discussed, showing their advantages and disadvantages.

4.1 Historical WebGIS

The first type of GIS architecture approach adopted and implemented to publish the Medieval data has been a Historical WebGIS. The WebGIS (Carrion, Migliaccio, Minini, & Zambrano, 2014) is based only on Free Open Source Software; its architecture is shown in Figure 1.

The server side is stored in an Ubuntu Linux virtual machine and is composed

by three modules: a DBMS, a GIS Server and a Web Server.
The DBMS used is PostgreSQL, an Open Source Software that can be integrated with a PostGIS extension, which manages the geographical information stored in the database and allows to interact with QGIS Mapserver, a WMS server that takes advantage of the QGIS libraries and uses the .qgs projects created with QGIS Desktop. The GIS server is invoked by the Apache Web Server, whose task is to publish web pages.
The client side is entrusted to a Web Browser, which interprets the web pages written in HTML5 language. The three components of this language are HTML, CSS and JavaScript: they manage contents, customization and behavior of web pages. The JavaScript toolkit GeoExt was used: it is composed by the ExtJS library, which allows to improve the graphical user interface and to integrate them with grids, buttons and toolbars, and the OpenLayers library, which allows to include web-mapping functionalities.
The Historical WebGIS, named *Geografie Medievali* (Medieval Geographies) and published into a Website, can be accessed through a "WebGIS" mode or a "Show table" mode (see Figure 2 and 3).

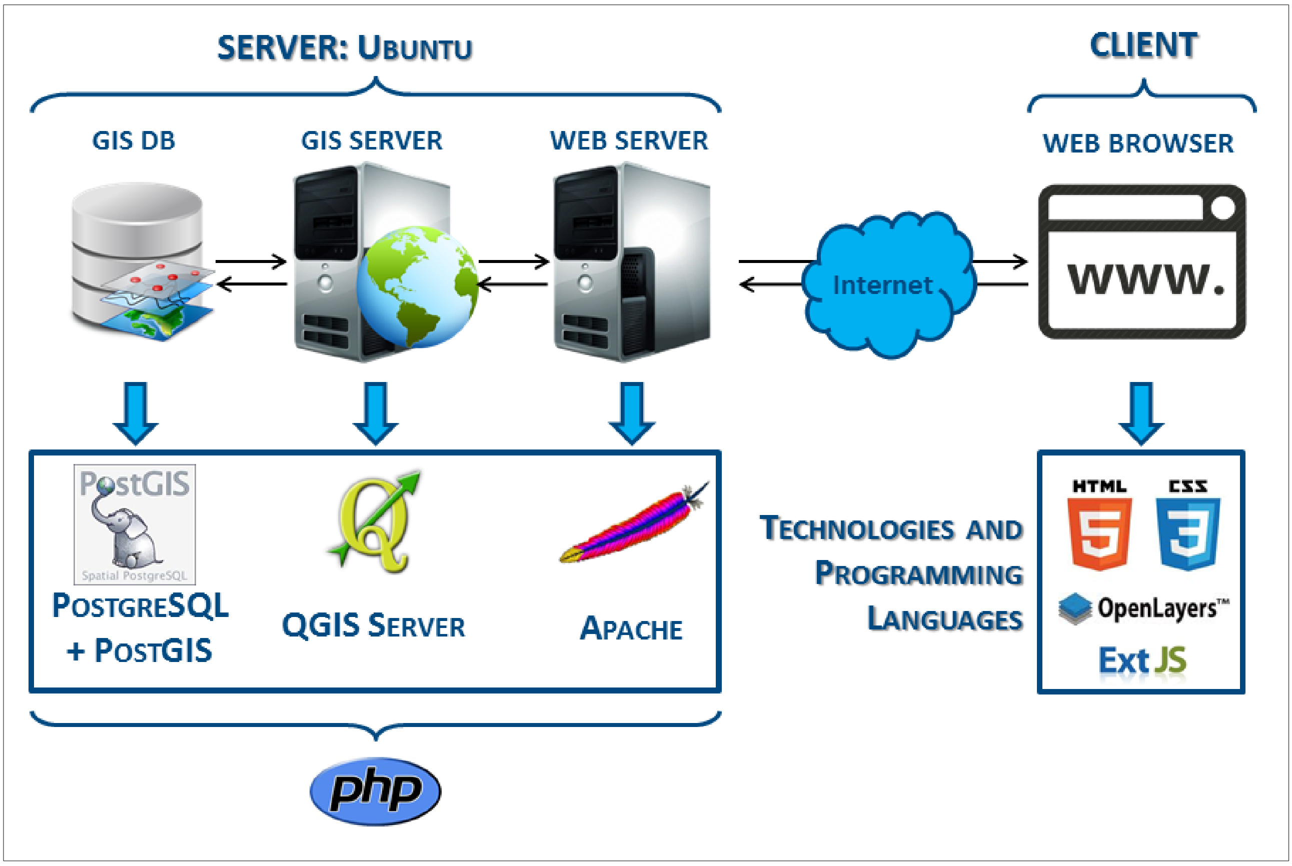


Figure 1 - Hardware and software architecture of the Historical WebGIS.

However, it must be underlined that the WebGIS only represents the georeferenced entities of the database and is based on the architecture just described, while the "Show Table" page displays all the entities, georeferenced and not, and some views that combine the information of two or more entities in one table, exploiting the relationships between them; this page is realized again with ExtJS libraries and allows to query tables and views with a query builder tool.

The typical advantage of this kind of architecture is that the installation of GIS software is not required on the client-side, because all the GIS functionalities are implemented into the server. On the other hand, in this case, the GIS tools provided by a GIS server are less advanced than those provided by a Desktop GIS, since every tool must be integrated or developed on purpose and the resources available for this project are very limited. Then, the main disadvantage of this solution is that every change must be programmed *ad hoc.*

The idea is to integrate over time other historical sources into the database, allowing Historians to explore the connections of the spatial information over the past. In this perspective a system where every upgrade requires complex programming is not feasible, and for this reason new approaches are being studied.

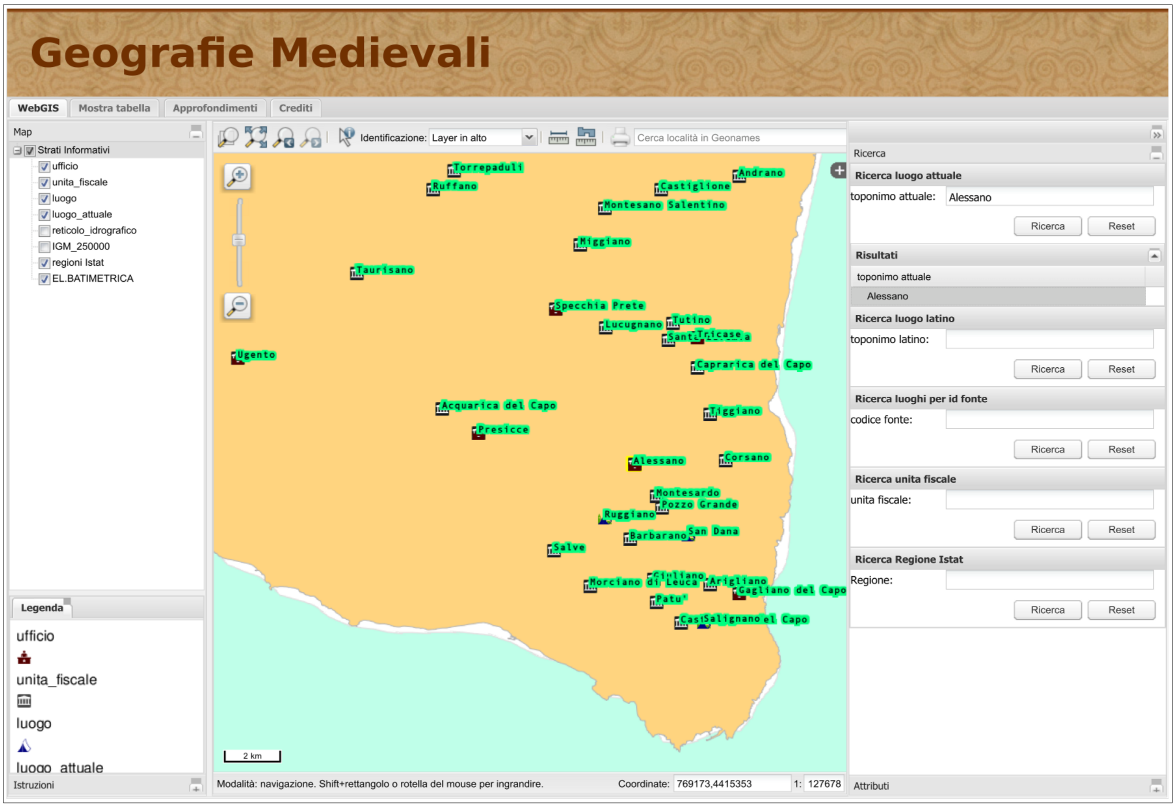


Figure 2 - The "Geografie Medievali" Website: "WebGIS" mode.

Figure 3 - The “Geografie Medievali” Website: “Show table” mode.

4.2 OGC Web services

The second approach which has been explored, however not yet implemented, is the possibility to publish the database through OGC Web services. The services which correspond to the OGC standards that have been considered are WMS, WFS and WPS. WMS is too simple for our case: in fact it only allows showing a map. To take into account the complexity of the designed database a WFS or a WPS would be necessary. However, a strong implementation work is needed also in this case, in particular if Free Open Source Software is used. A strategy is suggested in Peng and Zhang (2004). It is important to underline that what is crucial for the Medieval data considered in this work, is the possibility to be able to exploit the complexity of the database structure. From this point of view, a significant drawback of the OGC services approach for this project is the impossibility to display tables corresponding to not-georeferenced entities.

The advantage of this solution is again the fact that no software is needed to be installed on the client-side, apart possibly for some plug-ins. Moreover, the interface can be designed on purpose, allowing to make it as simple as possible, for non expert users.

The main disadvantage is still very similar to that of the WebGIS approach: the heavy programming (Rautenbach, Coetzee, & Iwaniak, 2013) which is needed, both during the implementation and the maintenance phases of the project, which makes this solution less appealing, for this low-budget project. In addition, usually a Web Service provides much less functionalities with respect to GIS Desktop software.

4.3 Web page and QGIS client architecture

The third approach presents an architecture (shown in Figure 4) that is much simpler than the one of the WebGIS: in this case only the GIS database is stored into the Server; the Client must be provided by GIS software e.g., QGIS Desktop, that natively supports the connection with a PostgreSQL database through the Internet.

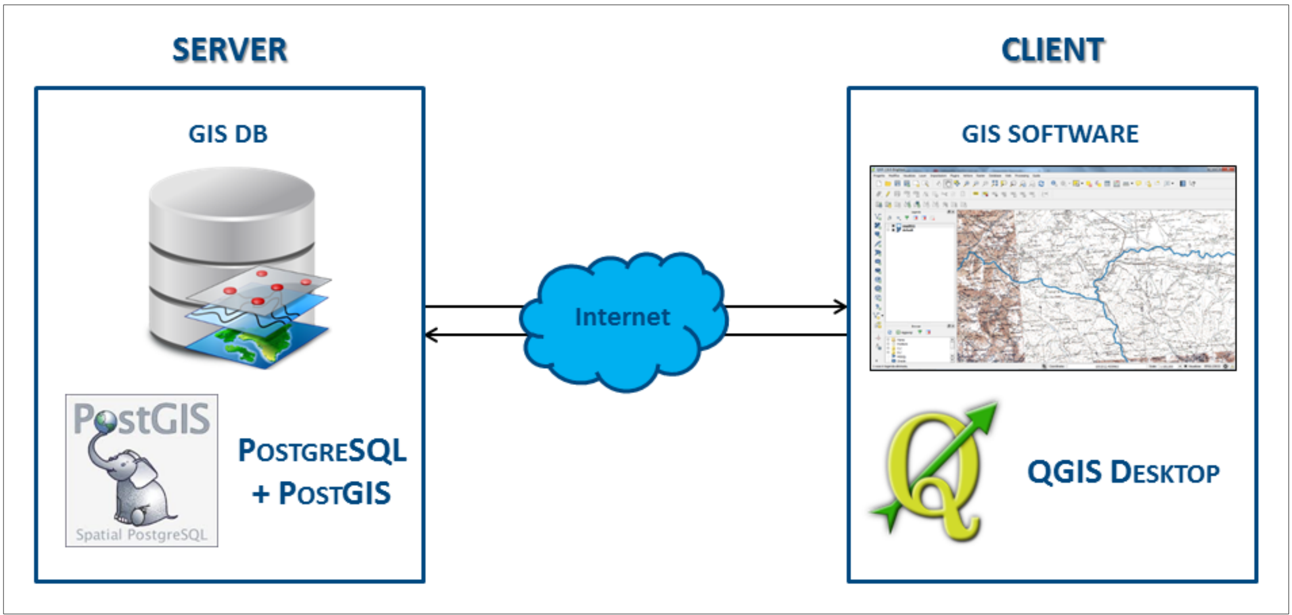


Figure 4 - Client-Server architecture for Medieval Geo-resources sharing through the Web.

Of course, GIS Desktop software offers many tools that are not usually available in a WebGIS architecture (like the one proposed in section 4.1), unless they are specifically designed; moreover, QGIS Desktop is an Open Source Software that does not require licensing costs and, in recent years, is catching up with commercial software like ESRI ArcGIS from the point of view of completeness and performances.

The main problem that remains in this approach is that the relationships implemented into the database are not recognized by QGIS, so when the connection with the DBMS is realized, the entities imported into QGIS, shown as layers or attribute tables, are not associated through relationships. This is a significant limitation when working with a database in which the non-spatial tables play an important role, like in our case.

One possible solution is the creation of views into the database, that contain the information stored in two or more entities. The main disadvantage is its rigidity and the increasing in number of attributes contained in a single table, when joining many entities, which makes the consultation of data much more difficult and less intuitive.

The solution which is being investigated at this moment, since it could prove to be more advantageous, is to exploit the possibility of creating some scripts that work into the QGIS environment and allow the software to take into account the relationships among entities established in the database, as it has been proposed by Cho, Bellemans, Janssens, and Wets (2014).

The entities imported into QGIS from the Historical Database and their cartographic representation are shown in Figure 5. They are overlaid onto a basemap of the Italian Military Geographic Institute (IGM) made available as WMS service on the Italian National Geoportal.

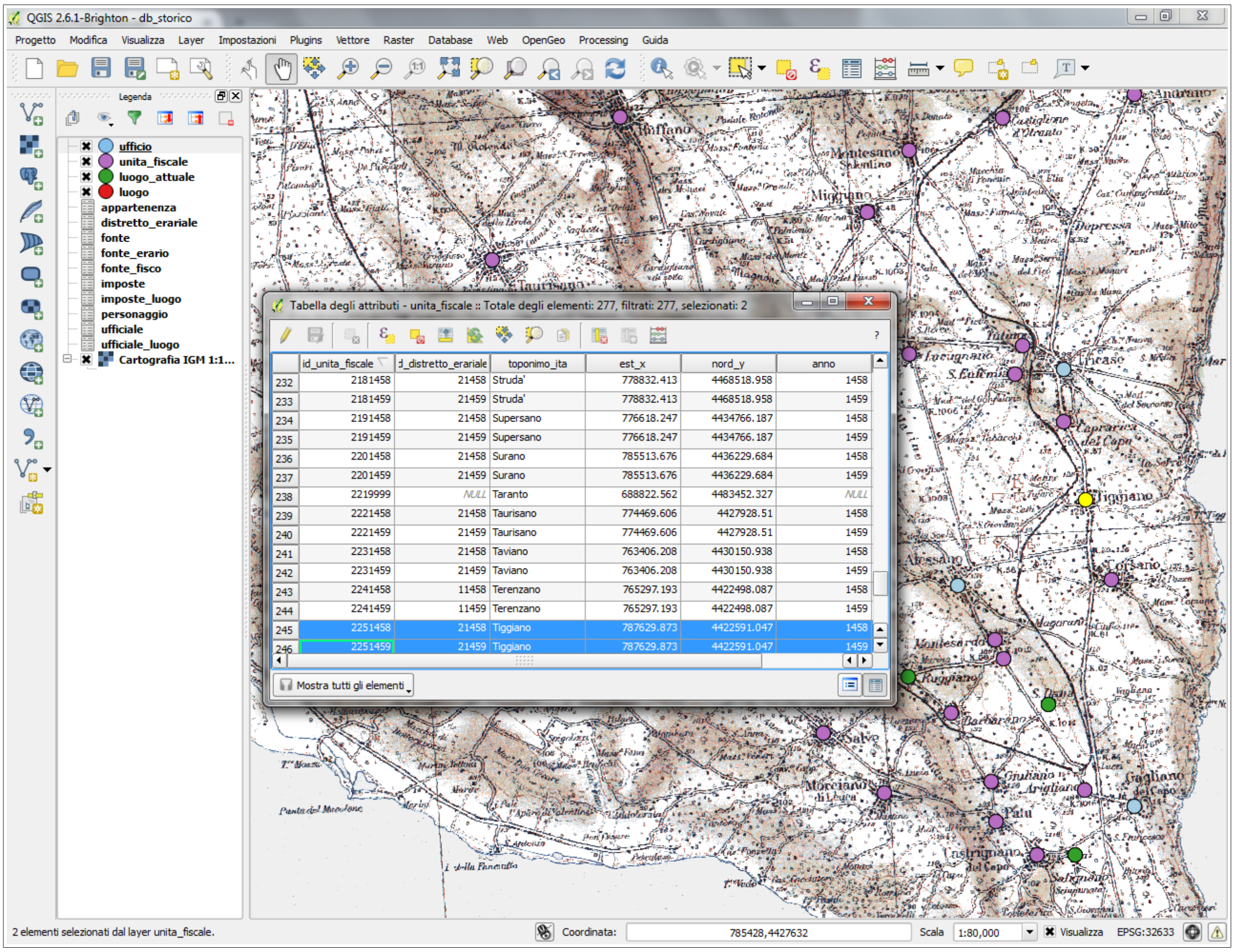


Figure 5 - The Historical Database imported into QGIS.

5 Conclusions

The research presented in this paper has the final objective to implement tools allowing to share a Historical Medieval database on the Web and to manage it through a GIS environment, exploiting all its potentiality and completeness (all the entities related with each other, as represented in the ERD design). The goal is to supply Historians participating to the project with an instrument that makes it possible to realize complex queries on Historical data and to show the outcome of spatial queries on maps.

Different approaches have been explored, namely: a WebGIS, which has been implemented and tested by Historians, OGC Web Services (such as WMS, WFS and WPS) and a Web page – QGIS Desktop combination with *ad hoc* implemented tools. In Table 1 the compliance of the proposed approaches with respect to the system requirements listed in section 4 is shown.

	WebGIS	Web Services	Web page +Desktop GIS
Open Source Software	Yes	Yes	Yes
Access to the database tables	Yes	No (WMS) Partial (WFS or WPS)	Yes
Access to the database relationships	Yes, but it needs ad-hoc implementation	No	Yes, but it needs ad-hoc implementation
Possibility to see the results of the table queries on the map and viceversa	Yes, but it needs ad-hoc implementation	Yes (WFS or WPS)	Yes
User friendliness	Yes	Yes	Yes
Maintenance simplicity	No	Yes	Yes

Table 1 – The proposed approaches against the requirements of the project.

According to the evaluation presented in Table 1 the approach based on the Web page – Desktop GIS combination with *ad hoc* implemented tools has been considered as the most convenient with respect to the Historians' needs and is being implemented. In the foreseen scenario, geodata will be stored on a server and published on a Website, so that the data sharing features of the system will be maintained. Moreover, thanks to the user friendliness of the last versions of QGIS, a very large set of functionalities will be available to Historians with a limited training effort on their side.
Many studies based on the application of GIS technologies to Historical studies have been and are being carried on, however the case discussed in this paper presents quite a unique character, since it deals with data from the Medieval period, for which such applications are still rather uncommon.

References

✔ Ardissone, P., & Rinaudo F. (2005). A GIS for the management of historical and archaeological data. In *Proceeding of the CIPA 2005 XX International Symposium*. Torino.

✔ Berman, L. M. (2005). Boundaries or networks in Historical GIS: Concepts of measuring space and administrative geography in Chinese History. *Historical Geography, vol. 33, Special issue: Emerging trends in Historical GIS*. Baton Rouge.

✔ Black, J. (2003). Mapping the past: Historical atlases. *Orbis, 47*, 277-293.

✔ Boonstra, O., Collenteur, G., & van Elderen, B., eds (1995). Structures and contingencies in computerized historical research. In *Proceedings of the IX International Conference of the Association for History and Computing*. Nijmegen, 1994 (Hilversum, 1995).

✔ Brovelli, M. A., Minghini, M., Giori, G., & Beretta, M. (2012). Web

Geoservices and ancient cadastral maps: The Web C.A.R.T.E. project. *Transactions in GIS 16*, 125-142.

- ✔ Carrion, D., Migliaccio, F., Minini, G., & Zambrano, C. (2014). Sharing historical geodata through a WebGIS. A Web application for History researchers. *Conference of Informatics and Management Sciences March, 24. - 28. 2014*, 331-334.
- ✔ Carrion, D., Migliaccio, F., Minini, G., & Zambrano, C. (in press). From historical documents to GIS: A spatial database for medieval fiscal data in Southern Italy, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*
- ✔ Cho, S., Bellemans, T., Janssens, D., & Wets, G. (2014). Integrating GIS and FEATHERS: A conceptual design. *Procedia Computer Science 32*, 405-412.
- ✔ De Moor, M., & Wiedemann, T. (2001). Reconstructing territorial units and hierarchies: a belgian example. *History & Computing 13*, 71-97.
- ✔ Fitch, C. A., & Ruggles, S. (2003). Building the National Historical Geographic Information System. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 36*, 41-51.
- ✔ Gregory, I. N. (2002). A Place in History: A guide to using GIS in historical research. *Oxbow*. Oxford.
- ✔ Gregory, I. N., Bennet, C., Gilham, V. L., & Southall, H. R. (2002). The Great Britain Historical GIS: From maps to changing human geography. *The Cartographic Journal*.
- ✔ Gregory, I., & Healey, R. (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography 31*, 638-653.
- ✔ Johnson, G., Gaylord, A., Franco, J., Cody, R., Brady, J., Manley, W., Dover, M., Garcia-Lavigne, D., Score, R., & Tweedie, C. (2011). Development of the Arctic Research Mapping Application (ARMAP): Interoperability challenges and solutions. *Computers & Geosciences 37*, 1735-1742.
- ✔ Maguire, D., & Longley, P. (2005). The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems 29*, 3-14.
- ✔ Mangia, M. A. (2013). *La diocesi di Benevento nel Quaternus decimarum del 1478: un contributo alla realizzazione di un GIS di cartografia storica sul Mezzogiorno bassomedievale*. Master’s degree thesis, Università del Salento.
- ✔ Pawson, E. (1997). The New Zealand Historical atlas. *Journal of Historical Geography 23*, 496-499.
- ✔ Peng, Z., & Zhang, C. (2004). The roles of geography markup language (GML), scalable vector graphics (SVG), and Web feature service (WFS) specifications in the development of Internet geographic information systems (GIS). *Journal of Geographical Systems 6*, 95-116.
- ✔ Pizzuto, S. (2009). Il Quaternus declaracionum di Francesco de Agello (1450-1461). Un contributo allo studio della geografia politica del Principato di Taranto in età orsinana. In *I domini del principe di Taranto in età orsiniana (1399-1463). Somaini, F. and Vetere, B. eds,*

Congedo Editore, 61-76.

- Rautenbach, V., Coetzee, S., & Iwaniak, A. (2013). Orchestrating OGC web services to produce thematic maps in a spatial information infrastructure. *Computers, Environment and Urban Systems 37*, 107-120.
- Schlichting, K. (2008). Historical GIS: New ways of doing history. *Historical Methods 41*, 191-196.
- Tait, M. (2005). Implementing geoportals: Applications of distributed GIS. *Computers, Environment and Urban Systems 29*, 33-47.

Teaching R as a GIS: problems, solutions and lessons learned

Robin Lovelace, Rachel Oldroyd
Centre for Spatial Analysis and Policy, University of Leeds

Abstract
R is a fast and flexible programming language and data environment which is widely accepted among the statistical community. It provides an open-source alternative to proprietary software such as SPSS and Stata. Despite R's acceptance among the quantitative disciplines, it still remains unpopular in the social sciences, especially geography. This unpopularity is somewhat surprising considering the number of advanced spatial packages that have emerged in recent years. Factors contributing to R's unpopularity may be its Command Line Interface (CLI), the steep learning curve associated with the language or the popularity of conventional Geographical Information Systems (GIS). In an attempt to raise awareness of R's spatial capabilities and its use as a GIS, a teaching team from the University of Leeds repeatedly delivered a day-long workshop over the period of 18 months. Based on teaching experiences and formal feedback, the workshop materials were incrementally improved based on an understanding that R skills are developed through 'experiential learning'. A number of measures were introduced to facilitate the learning process and the findings of the teaching activity should be of interest to anyone teaching computer code to beginners. The resulting materials and guidance will be of practical use for others teaching R for spatial applications.

Keywords
GIS, R, Spatial Analysis, Education

Introduction
The features, performance and graphics provided by R's spatial packages now mean that R can rival market-leading GIS software suites such as QGIS and ArcGIS (Bivand, Pebesma & Gómez-Rubio, 2013). Despite the recent progression of R's spatial capabilities, the software is still not frequently used by geographers and social scientists. Factors which may contribute to R's unpopularity include; the command line-based interface, the dominance and familiarity of proprietary GIS products and R's reputation as being difficult to learn. To the majority of computer users, the Command Line Interface (CLI) is an alien concept. The prospect of using a tool which only responds to typed commands is unfamiliar and most social scientists prefer to 'click and point' their way to a pretty map. However, once familiarity with the CLI is gained, the benefits of using R compared to other Graphical User Interface (GUI) based programs become clear, these include:

- Ease of reproducibility, including modifying and building on the work of others.
- Speed of undertaking batch processes: a single script can perform the same process on many files with little or no modification.
- Advanced and highly customisable graphical capabilities, empowering the user to visualise data in new ways (e.g. faceted plots and animations).

- Highly extensible environment with over 6800 additional libraries (Comprehensive R Archive Network [CRAN], 2014).

In short, R has proved itself as an integrated tool for processing, modelling and mapping geographical data (Ihaka & Gentleman, 1996), but this has yet to translate into widespread uptake by social scientists and geographers. In an attempt to dispel the negative image associated with R, a 1 day workshop was developed and delivered by staff at the University of Leeds. This workshop was delivered through the Geospatial Analysis and Simulation (TALISMAN) project, funded by the UK's Economic and Social Research Council (ESRC).

Case Study

The aim of the Introductory R for Spatial Analysis (IRSA) workshop was to raise awareness of R's spatial capabilities by teaching new and emerging methods for processing, analysing and visualising spatial data in R. 6 courses were delivered to approximately 130 participants between the period of November 2013 and March 2015 and informed changes were made to the teaching material and delivery of the workshop throughout this period. These changes were based on formal and informal feedback. Formal feedback was collected from 75 participants during 4 of the courses. The remaining 2 courses were delivered to non-native English speakers as part of international conferences and formal feedback was not collected.

The Teaching Model

Prior to development of the teaching materials, the target audience was identified. It was thought that the majority of attendees would be academics working in a social science related field with some experience of GIS. Measurable objectives were also defined to provide context to the teaching practice and provide a template for the teaching materials. It was hoped that, by the end of the workshop participants would:

- Have an understanding of R's underlying syntax and structure
- Have an understanding of loading, manipulating, processing and transforming spatial data in R
- Have used R's graphics packages to create aesthetic maps in R
- Have an awareness of R's more complex spatial functionality

A teaching model consisting of short lecture style talks alongside independent work was adopted for the workshop. It was thought that this approach would facilitate experiential learning, 'learning by doing', and would allow participants to work at a pace which was comfortable for them. This would also provide the opportunity and time for individuals to explore concepts which were not familiar to them or which were particularly interesting. It was decided that an existing tutorial named 'A short Introduction to R' by Harris (2012) would be used in the first half of the workshop and a more spatially-inclined tutorial would be used in the second half to introduce more complex functionality. The latter tutorial was named 'Creating Maps in R' (www.github.com/rlovelave/Creating-Maps-in-R) and was purposefully written for use in this workshop.

In order to provide a quantitative measure of course effectiveness, feedback forms were collected after each workshop. The feedback form consisted of a number of Likert scale questions. The results were summarised whereby each category was given an average score out of 100 based upon participant feedback. The categories were; the content was interesting; the course materials were useful; speakers knowledge; clarity of explanations; difficulty of the course; prior knowledge assumed; speed of presentation; time for questions. This quantitative reflection process provided an objective way of highlighting which parts of the course were sucessful and which were not.

Problems and Solutions

A well-defined problem associated with developing training workshops, is pitching the material at the appropriate level. It was decided that the workshop would be advertised as entry level, suitable for beginners who had little or no experience with R. In reality, a range of participants with varying levels of experience registered for the event. These ranged from proficient R users with limited experience of spatial data, to participants who regularly used spatial data but had no experience of R and very infrequently, participants who had limited experience of both R and spatial data. The extreme range of knowledge introduced a number of problems. Not only were the materials required to cover R’s basic syntax and capabilities, they also had to cover general programming techniques and basic geographical concepts whilst maintaining the interest of the higher-level users.

The majority of the problems experienced over the 18 month period stemmed from the range of participant knowledge. It soon became clear that completing two separate tutorials was not conducive to effective learning. The participants felt pressured into completing the tutorial within the limit of the one day period; ‘*I acted more like a typewriter*’, ‘*I was typing code without understanding what I was doing*’. To address this problem, the ‘Creating Maps in R’ tutorial was extended and adapted to form an entry level tutorial on which the IRSA course would be solely based. Although this meant that less material was covered, this approach allowed attendees more time to comprehend the code and attain a greater level of understanding. This was reflected in the feedback as the average score for ‘course difficulty’ rose from 93 to 97 and the score for ‘prior knowledge assumed’ rose from 89 to 96 out of 100, where a higher score reflected higher suitability.

Reducing the amount of material introduced another problem. Participants who were more proficient in R skipped over the basics covered at the beginning of the tutorial and finished earlier than expected. In an attempt to preserve the introductory level of the ‘Creating Maps in R tutorial’ but also maintain the attention of these participants, a number of optional extension exercises were introduced. These were added to the end of each section allowing higher-level users to further explore R’s spatial functionality. A conscious effort was made to clearly signpost these exercises as optional, as it was found that other participants felt pressured into completing all sections if not. Over the duration of the 18 months the score for the ‘Content was interesting’ and ‘Materials were useful’ categories increased from 84 to 89 and 85 to 91 respectively. The feedback also reflected upon the success of the optional exercises; ‘*Although I started the course with considerable R experience, and the course had to go* *through the basics. The pace was just enough and there were enough extra bits of information to keep me interested'*, *'I think the course strikes a good balance between being accessible for an R novice and interesting for more experiences R users'*.

Originally it was assumed that participants would have some familiarity with geographical datasets, however this was not true of all the attendees. A number of attendees were unfamiliar with the shapefile format used in the tutorial and struggled to understand the differentiation between geometric data, attribute data and the associated R 'slots'. In order to address this problem, it was decided that a traditional GIS would be used to supplement the teaching. QGIS provided a quick and easy way to visualise the geometry and attribute data of a shapefile for participants who were not familiar with geographical formats. This initial understanding proved pivotal to the comprehension of the concepts that followed and was extremely successful.
The aim of the IRSA workshop was to raise awareness of the spatial capabilities of R and to dispel the negative image associated with the language. The majority of participants had not used R before and it was therefore important to create a good first impression. The benefits of using Integrated Development Environments (IDE) for beginners are well defined (Debuse & Lawley, 2012), therefore the R Studio IDE was used to support the learning process. The auto-completion function in R studio was particularly useful as it meant that the complexities of the language did not impair learning. The help window allowed for quick and easy access to documentation and the workspace pane creates a list of active data, values and functions which helps the debugging process. It is thought that these devices all helped to ease the steep learning curve associated with the language.

Lessons Learned

In conclusion, the IRSA course has been extremely successful and continues to be in high demand. The materials have been improved incrementally over an 18 month period and a number of lessons have been learned. When communicating computer code to beginners it is important to provide a supportive learning environment and introduce devices to ease the learning process. The use of an IDE is one example, provided it is not over complicated (Reis & Cartwright, 2004). Quality always rules quantity in terms of materials and it is best to cater for the needs of the less experienced users as optional tasks can always be provided. Having a high number of demonstrators is essential when teaching an entry-level course and participants appreciate individual interaction; *'Explanations given by the tutors in practical sessions was excellent'.* A basic understanding of geographical data formats was essential to the success of the course and it should not be assumed that participants will be familiar with these formats. Traditional techniques have proved successful in providing familiarisation with the datasets. Lastly, it is important to mention the appropriateness of the course. This workshop was aimed at academics who had some experience of geographic data. In reality, a range of partipants attended. A small minority of attendees had used neither R or a GIS and the IRSA course was unsuitable for their needs. Training in traditional GIS techniques may have been more suitable. Defining the target audience and clearly signposting any prerequisities is essential in providing participant satisfaction.

References

Bivand, R. S., Pebesma, E. J., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*. New York, Springer.

Comprehensive R Archive Network. (2014, November 21). Contributed Packages. Retrieved from: http://cran.r-project.org/web/packages/

Debuse, J., & Lawley, M. (2012). The learning and productivity benefits to student programmers from real-world development environments. *Information Systems Education Journal*, 10(5), 61-81.

Harris, R. (2012, November 2014). *A short introduction to R*. Retrieved from: http://www.social-statistics.org/wp-content/uploads/2012/12/intro_to_R1.pdf

Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5, 299-314.

Reis, C. & Cartwright, R. (2004). Taming a professional IDE for the classroom. Proceedings of the 35th SIGCSE technical symposium on Computer science education, 35, 156-160.

GIS-based multivariate statistical analysis for landslide susceptibility zoning: a first validation on different areas of Liguria region (Italy)

Roberto Marzocchi[1,2], Agnese Rovegno[2], Bianca Federici[2], Rossella Bovolenta[2], Riccardo Berardi[2]

[1] Gter srl Innovazione in Geomatica Gnss e Gis, (University of Genoa Spin-Off company) – Via Greto di Cornigliano 6r, 16152 Genova (Italy)

[2] Dipartimento di Ingegneria Civile Chimica ed ambientale (DICCA), Università degli studi di Genova, Via Montallegro 1, 16145 Genova

Abstract

Several approaches have been developed worldwide in order to assess landslide hazards. The most diffuse methods are heuristic analysis, statistical or process-based methods (Fell et al., 2008). The present research consists in applying univariate and multivariate statistical analysis for landslide susceptibility zoning, considering the influence of various factors on the occurrence and triggering of slides and flows, and critically reviewing the choice of the calibration area in terms of extension and characteristic of influencing factors. The analysis was applied to different areas of Liguria region (Italy).

Keywords

Landslide susceptibility zoning, statistical methods, GRASS GIS

1. Introduction

Almost 10% of the Italian territory is affected by hydraulic and geological risks (floods and landslides) and the 6.9% of the territory is catalogued as landslide by the IFFI national inventory of landslide (ISPRA , 2013). Recently, various procedures have been developed for the analysis of spatial distribution of future landslides, at different scales and with different approaches.

As well described in literature (Varnes and IAEG 1984; Hutchinson 1995, Fell et al., 2008) the hazard analysis is based on the assumption that landslides occur in the same geological, geomorphological, hydrogeological and climatic conditions as in the past and are normally controlled by these identifiable physical factors. To assess landslide hazard several approaches have been developed worldwide; the main diffuse methods are heuristic analysis, statistical or process-based methods (Fell et al., 2008). Each method has its advantages and disadvantages, but in general the choice of the method depends on objective, scale of the study and finally also on the available data.

Heuristic methods are based on the judgement of the expert carrying out the

analysis, both for the individuation of the perimeter of occurred landslides and for the creation of landslide susceptibility maps. Often these methods are based on geological and geomorphological criteria and use local observations which are collected in the studied area.
Instead, statistical methods usually correlate an inventory of landslides occurred in the past with factors which are supposed to be responsible of slope failure (Cascini, 2008). The statistical methods are considered more suitable to be applied to wide and differentiated zones and usually this approach uses GIS environment to integrate spatial variables; nevertheless statistical approaches sometimes require complex calculation procedures (Guzzetti et al., 1999).
Concerning the landslide risk management, in authors’ opinion, the availability of an instrument able to perform hazard analysis on a wide area, quickly and with low resources, is really appreciable. Within the activities of a national research program, entitled “Mitigation of landslide risk using sustainable actions”, we are analysing landslide susceptibility zoning through statistical methods, focusing the attention on the influence of various factors on the occurrence and triggering of landslides, and critically reviewing the choice of the calibration area in terms of extension and characteristic of influencing factors.
The type of landslides involved in this study are slides and flows, according to the UNESCO WP/WLI (1993). GRASS GIS 7.0 software is used.
On the basis of field experience and our previous research (Natali et al., 2010; Bovolenta et al, 2011) and on the basis of an accurate analysis of the pertinent literature, we have deduced that geomorphological, geological, anthropic and climatic parameters are the most important factors. Consequently, we have considered and analysed them from a statistical point of view in order to understand their influence on the landslide susceptibility and their spatial variability, calibrating the statistical model on the entire Liguria region (in North-western Italy) or on its four administrative provinces (i.e. Imperia, Savona, Genova, La Spezia).
Final purpose of the mentioned research is to develop a guideline for the landslide susceptibility zoning, giving an instrument to non-GIS and non-statistical expert users for the choice of factors to be considered.

2 Methods

Statistical methods usually correlate an inventory of landslides occurred in the past with factors which are supposed to be responsible of slope failure. Hence, an accurate analysis of the pertinent literature and field experiences were necessary to identify the main predisposing factors to landslides. Then a logistic multiple regression was chosen to analyse them.

2.1 Predisposing factors to landslide

Starting from a literature analysis (Dai & Lee,2002; Pauditš & Bednárik, 2002; Ayalew & Yamagishi, 2005; Lee, 2007; Dahal et al., 2008; Cencetti et al., 2010; Natali et al., 2010; Pradhan & Lee, 2010) we identified several predisposing factors to landslide, as summarized in table 1. Some of them, as lithology, slope and land use, are considered by the majority of the references, because they are proven factors of instability. Other factors are taken into account depending on the features of the study area.

	Geological					Geomorphological						Anthropogenic			Climatic	
	Lithology	stratigraphy	Distance from tectonic alignments	Soil type	Soil depth	Slope	Morphological parameters	Aspect	Elevation above sea level	Distance from the drainage network	Accumulation	Land use	Vegetation	Distance from the road network	Climatic aggressivity F_{FAO}	Mean monthly rain
Natali et al. (2010)	x					x						x				
Pradhan & Lee (2010)	x		x	x		x		x		x		x	x			x
Pauditš & Bed-nárik (2002)	x					x			x							
Lee (2007)	(x)		x	x		x		x		x		x	x			
Dahal et al. (2008)	x			x	x	x		x	x		x	x		x		
Cencetti et al. (2010)	x	x				x						x			x	
Dai & Lee (2002)	x					x	(x)	x	x	(x)		x				
Ayalew & Yamagis hi (2005)	x	x	x			x		x	x					x		

Table 1: Literature analysis of predisposing factor to landslide [(x) indicates a weak suggestion of such factor]

In our work, combining literature analysis with some considerations on cartographic data availability, we considered the following factors:

- *Geological factor*: we use a *geo-lithological* vector map (1:100'000), for a lithological description of soils.
- *Geomorpological factors*: we use digital elevation models (DEM), which describe the surface topography, to obtain *elevation*, *slope*, *aspect* maps.

Moreover surface topography controls flow sources, flow direction and soil moisture concentration, that are important factors correlated to the density and spatial extent of landslides. Hence we use the DEM to calculate the *accumulation* map.

- *Anthropogenic factors*: several studies demonstrated that the cuts of the slope caused by the roads are usually causes of anthropologically induced instability; this information could be extracted from either a *land-use cover* map or a *road* map.
- *Climatic factor*: to take into account the rain occurrence and its temporal distribution, which may affect the landslide occurrence; in particular we use the *climatic aggressivity* (F_{FAO}) named also *"Fournier index"*. It was defined by Arnoldus (1977) as:

$$F_{FAO}=\frac{\sum_{i=1}^{12} p_i^2}{P} \qquad (1)$$

where p is the mean monthly precipitation and P the mean annual precipitation.

2.2 Logistic multiple regression

Logistic regression is an example of generalized linear model (GLM). It is a flexible generalization of ordinary linear regression, which is appropriate for response variables that have error distribution models other than a normal distribution (i.e. landslide occurrence). It is a well-working method for dichotomous data, which means that the dependent variable can have only two values (event occurring or not occurring), and predicted values can be interpreted as probabilities since they are constrained to fall in the interval between 0 and 1. In the present study, the dependent variable is a binary variable representing the presence or absence of landslides. In particular logistic regression is known to be more flexible than the alternative methods for dichotomous data, such as the linear regression.
The technique of logistic multiple regression yields coefficients for each variable based on data derived from samples taken across a study area. These coefficients serve as weights in an algorithm which can be used in the GIS environment to produce maps depicting the probability of landslide occurrence.
Using the multiple regression model, the relationship between the occurrence and its dependency on several variables can be expressed as following:

$$logit(p)=\ln\left(\frac{p}{1-p}\right)=b_0+\sum_{j=1}^{n} b_j X_j \qquad (2)$$

where X are the variables, b the coefficients and p the landslide occurrence probability obtained by a landslide inventory map.
In the present work the calculation of b regression coefficients has been done using the *r.regression.multi* GRASS GIS command. This tool is implemented for the linear regression model, hence the inputs of the model are the maps of independent variables X_j and the logit of the landslide map.
The outputs are:

- a map with estimate of logit(p)

- a map with the errors (optional)
- a file with coefficient values, the Akaike Information Criterion (AIC) index and the R^2 correlation coefficient for each explaining variable.

The AIC is a measure of the relative quality of a statistical model for a given set of data. If the adding of a variable has a positive effect on the model, the values of AIC have to decrease. Adopting a stepwise method we tried to understand the influence of adding a variable or not.

The R^2 correlation coefficient for each explaining variable represents the additional amount of variance when including the variable compared to when excluding that. It gave us a measure of the influence of a single factor in each model.

Preliminary to the application of the logistic multiple regression, we have to analyse each variable singularly through a bivariate statistical approach. In this way we are able to understand the real influence of each variable on the physical phenomenon and to correctly classify the variable map in order to have a direct or inverse proportionality with landslide occurrence. This preliminary step is especially necessary in case of no-numerical variable, as the geolithological and the land-use map, or in case of numerical variable not characterized by a regular proportionality with landslide occurrence (i.e. aspect map). In fact, the adoption of a logistic regression model involves the calculation of the value of logit through a linear function of the independent variables X; therefore a monotonic (increasing or decreasing) behaviour needs to be observed also in the classification of the X variables.

Using a bivariate statistical approach the landslide occurrence is considered as dependent on a single variable and the conditional probability is obtained using the Bayes's theorem:

$$P(\text{landslide}|X) = \frac{P(\text{landslide} \cap X)}{P(X)} \quad (3)$$

On the basis of bivariate analysis it is possible to correctly reclassify the geolithological, land-use and aspect maps using an increasing or decreasing rule, that otherwise is not possible.

3. Study area and data sources

The susceptibility statistical analysis described in the previous section was calibrated on the whole Liguria region (north-western Italy) or on its four administrative provinces (Imperia, Savona, Genova, La Spezia in figure 1).

The following are the data used:

- The landslide inventory map, available from Liguria region geo-portal. Its development starts with the IFFI project (Inventory of Landslide Phenomena in Italy). IFFI project was carried out by ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) and the Regions and Autonomous Provinces since 1998. It supplies a detailed picture of the distribution of landslide phenomena within Italy. Actually the inventory is periodically updated from each region and the version used for calibration is updated to 2013. The type of landslides involved in this study are slides and flows, according to the UNESCO WP/WLI (1993).

- A raster DEM of the entire Italy downloaded from ISPRA, with a cell size of 20x20m.
- A geo-lithological vector map (scale 1:100'000) coming from the Liguria Region Environmental Service.
- A land-use vector map (scale 1:10'000) updated to 2012.
- A road network vector map (scale 1:10'000) updated to 2003.
- A map with the ARPAL rain gauges with daily rain data available for at least 10 years (ARPAL, 2012). From the daily rain data available in several rain gauges on the Ligurian territory (figure 1), we calculated the monthly and annual mean precipitation and consequently the F_{FAO} as expressed by the equation 1. The F_{FAO} values in correspondence of the rain gauges were interpolated, after having tested various interpolation methods, as IDW (Inverse Distance Weighted) with 3 or 4 points, RST (Regularized Splines with Tension), Voronoi and Natural Neighbours. We have chosen the Watson algorithm for Sibson natural neighbour interpolation (Sibson, 1980, Sibson, 1981; Watson, 1992) implemented in the GRASS GIS *r.surf.nnbathy* add-on command, because this algorithm is reliable for large dataset with an irregular sampling. Nevertheless the result is available only inside a convex polygon which includes the known data. In order to obtain a complete result on all the Liguria region the Voronoi or Thiessen polygons are used in the boundary region, where the natural neighbour interpolation does not return results. F_{FAO} interpolation is shown in Figure 2.

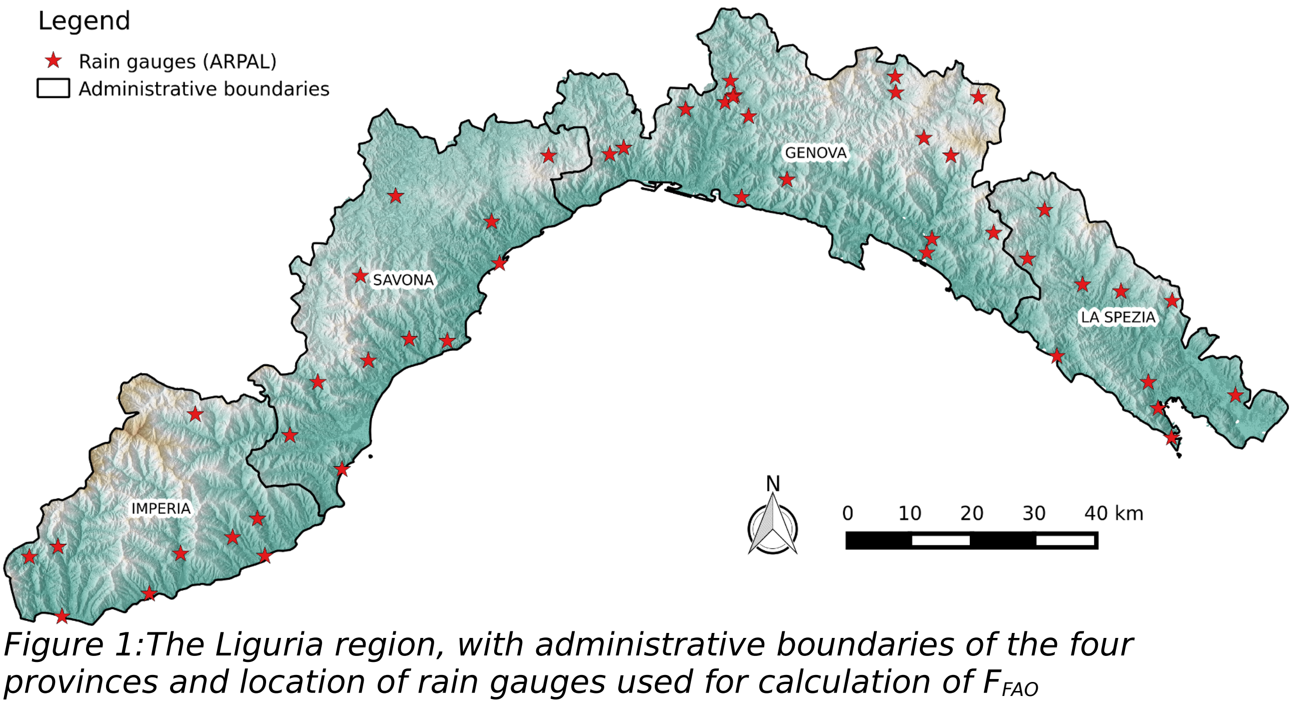


Figure 1:The Liguria region, with administrative boundaries of the four provinces and location of rain gauges used for calculation of F_{FAO}

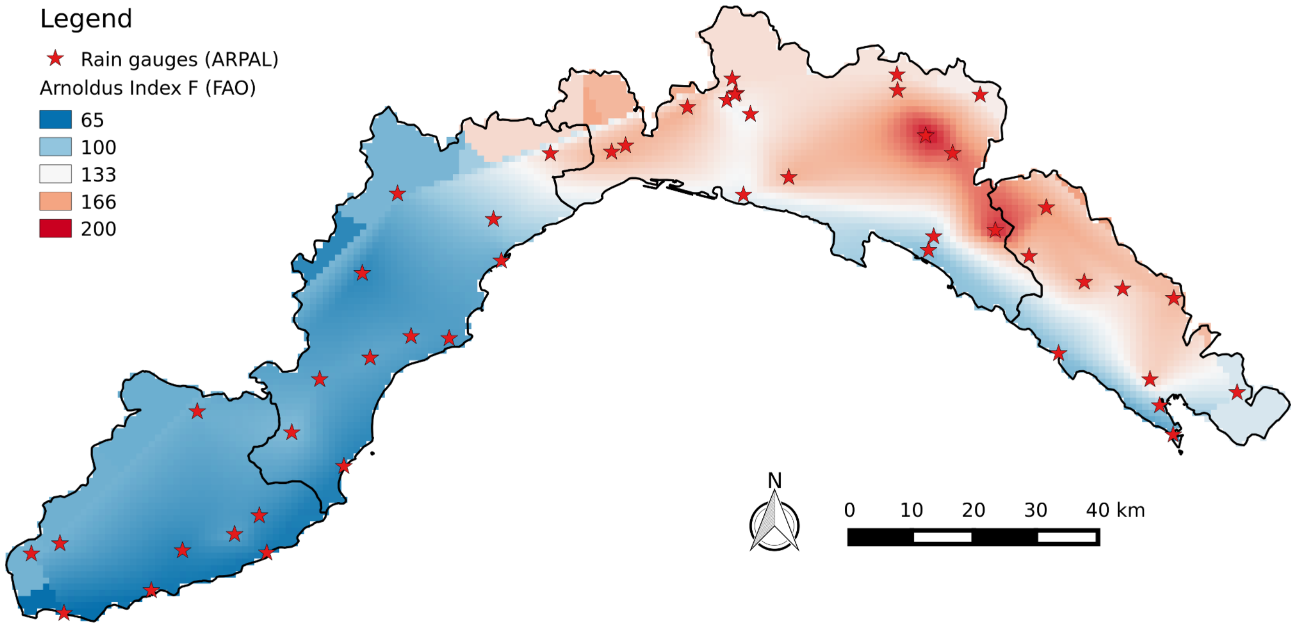


Figure 2: Result of interpolation of Arnoldus or Fournier Index (F_{FAO}).

3. Results and discussion

The 5 maps previously listed were used to calculate the 8 maps of the analysed predisposing factors to landslide: slope, land use, lithology, accumulation, distance from road, elevation, aspect, F_{FAO} . Such 8 maps were conveniently classified in 3 to 9 classes, in function of the results of the bivariate statistical analysis of each variable singularly considered and the landslide inventory map, applied on all the Liguria territory (figure 3). The meaning associated to the classes for the non-numerical variables, i.e. lithology and land cover, or for numerical maps as log(accumulation) and F_{FAO} are reported in table 2.

F_{FAO}		Log(accumulation)		Lithology		Land cover	
1	<100	1	≤ 1 ridge	1	Deposits	1	other
2	100-200	2	1 – 2	2	Magmatic rocks	2	buildings
3	>200	3	2 - 3	3	Metamorphic rocks	3	Agricultural areas
		4	3 – 4	4	Sedimentary rocks	4	wood
		5	4 – 5	5	Other	5	sparse vegetation
		6	> 5 humid zone			6	Open space with little or no vegetation

Table 2: Classification of explaining variables (see figure 3)

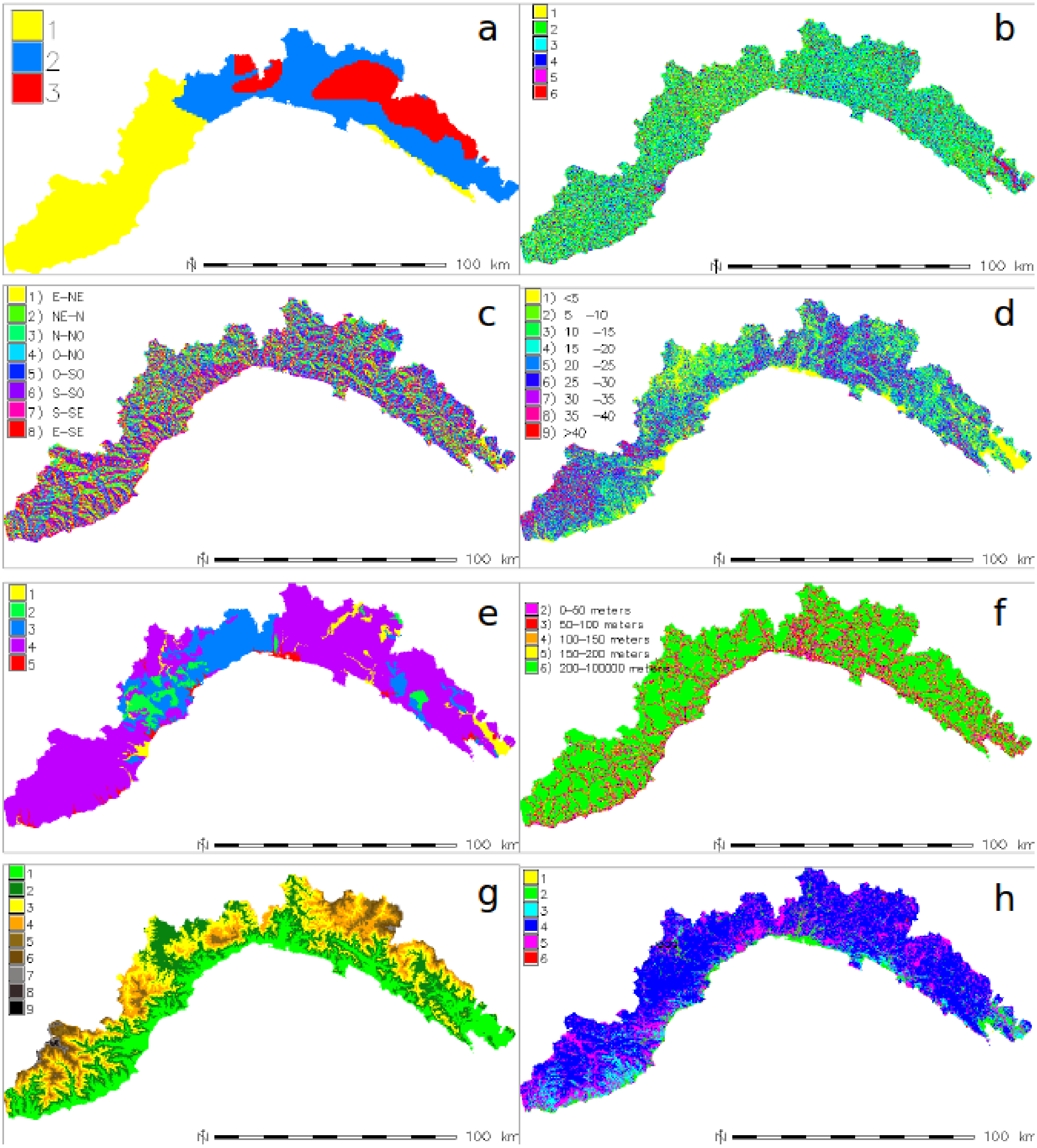


Figure 3: The reclassified maps of explaining variables: a) F_{FAO}; b) Log(accumulation); c) aspect; d) slope; e) lithology; f) distance from road; g) elevation [a class every 250 m a.s.l.]; h) land cover.

Then we have applied the logistic multiple regression and analysed the variability either of AIC index related to the number of analysed explaining variables, and the b and R^2 correlation coefficients for each best model. The calibration has been done using all the Liguria region and the four provinces separately shown in figure 1.

	Liguria		Imperia		Savona		Genova		La Spezia	
Number of variable with lower AIC index	8		7		8		8		7	
Parameter	**b**	**R^2 (10^{-3})**	**b**	**R^2 (10^{-3})**	**b**	**R^2 (10^{-3})**	**b**	**R^2 (10^{-3})**	**b**	**R^2 (10^{-3})**
slope	0.029	0.68	0.040	1.40	0.015	0.19	0.028	0.41	0.07	0.12
land use	0.059	0.32	0.042	0.19	0.030	0.08	0.104	0.71	0.013	0.07
lithology	-0.044	0.22	0.020	0.00	-0.045	0.29	-0.079	0.47	0.010	0.05
accumulation	0.021	0.14	0.033	0.36	0.016	0.09	0.019	0.08	0.004	0.02
aspect	0.006	0.03	0.010	0.11	-0.002	0.00	0.009	0.05	0.003	0.05
elevation	-0.012	0.05	0.015	0.12	-0.005	0.01	-0.086	1.39	-0.002	0.00
distance from road	-0.022	0.19	0.001	0.00	-0.026	0.25	-0.028	0.23	-0.001	0.00
F_{FAO}	0.019	0.04	-	-	-0.038	0.05	0.141	0.61	-	-

Table 3: Results of the logistic multiple regression.

For the Liguria region the AIC value decreases until the eighth explaining variable and the same behaviour was observed for the Savona and Genova provinces (figure 4). For Imperia and La Spezia provinces the F_{FAO} variable has a quasi-uniform class in the calibration regions, therefore its influence is negative on the model as well demonstrated by the increase of the AIC parameter for the model with 8 variables.

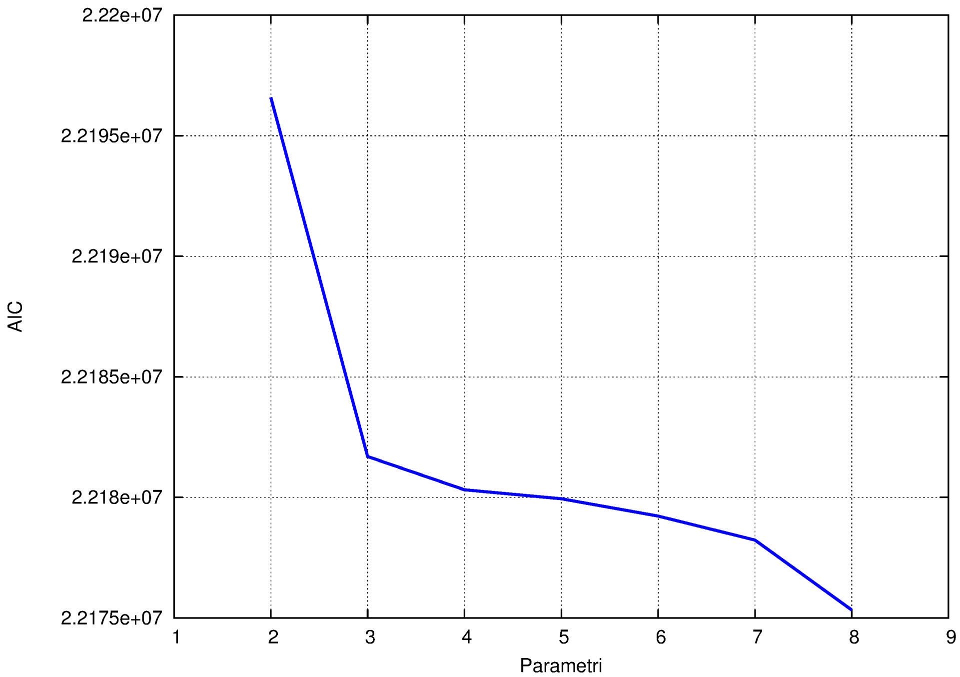


Figure 4: AIC index for the calibration on the entire region.

R^2 gives an idea of the influence of each explaining variable on the landslide susceptibility: higher the value higher the influence. Also the coefficient b is an

indicator of the influence of each variable on the landslide occurrence: if it is positive the variable is a predisposing factor, otherwise it is not-predisposing.
A coherence of behaviour in the various calibration areas is an indicator of a greater reliability of the factor, as it happens for example for slope, land use and accumulation. But some factors, as elevation, aspect, F_{FAO} and lithology, show a very variable influence on the occurrence of landslides, with positive and negative b values if obtained from different calibration areas, and hence lower reliability.
In particular, the coefficients applied to slope, land use and accumulation maps are always positive for all the different calibration areas, meaning that their increasing have an effect also on the increasing of landslide occurrence probability. The coefficient applied to the map “distance from road” assumes values almost always negative, even if very low in absolute terms, because lower distances to the road have higher effect on landslide occurrence.
The elevation variable, seems to have, at least for the Liguria region, a rather low influence on the landslide occurrence, even if it can be strongly related to other physical phenomena hence statistically significant also on the landslide occurrence.
Concerning aspect variable, we observed a different behaviour for the West coast (Imperia and Savona provinces) and East coast (La Spezia province) and a different classification is probably more suitable and will be tested in further analyses.
With regard to F_{FAO} it seems to be a good parameter in analysis on large areas with different climatic conditions while it is not a good parameter when applied to small calibration area characterized by similar climatic conditions.
Finally, lithology map seems to be poorly or highly related to landslide occurrence in the different areas of calibration, as demonstrated from the high variability of R^2 correlation coefficient in the different provinces. This is probably due to the low map scale affecting also the spatial variability, which, in some provinces is very low (e.g. Imperia). For this reason, in our opinion, the use of this lithology map is appropriate only for wide areas (entire region). Obviously the availability of a more accurate vector geo-lithological maps (e.g. 1:25'000) could solve the problem, and the use of this map could become more appropriate from a statistical point of view.
For all these last four parameters, this first validation (conducted on different calibration areas) highlights the importance of bivariate analysis prior to the logistic multiple regression in order to understand if there is a correlation. In positive case, the next step is to correctly classify the map, or exclude the map from the multiple regression analysis in negative case.

5. Conclusions

This first validation performed in the Liguria region (considering different areas for extension, morphological and climatic characteristics) was surely interesting to understand few critical issues of the multivariate statistical analysis.
The next steps of our research are:

- a more accurate analysis of aspect and elevation maps comparing the Liguria region with other areas;

- an extension of the validation area to the Piemonte region in order to add different morphological conditions to the statistical analysis (a more extended alpine area, the Po valley, the Langhe and Monferrato hills, etc.).

The final aim of our research consists in providing guidelines for not expert users which may clearly explain the requested steps to obtain a susceptibility map with a statistical multivariate approach in order to avoid any mistakes.

Acknowledgement

This work is financed by the national research program (PRIN 2010-2011) “*Landslide risk mitigation through sustainable countermeasures*" coordinated by the University of Salerno.

References

✔ Arnoldus, H.M.J. (1977). Methodology used to determine the maximum potential average annual soil loss due to sheet and rill erosion in Morocco. *FAO Soil bulletin,* 34, 39-48.

✔ ARPAL (2012). Atlante climatico della Liguria (http://www.arpal.gov.it/homepage/meteo/analisi-climatologiche/atlante-climatico-della-liguria.html)

✔ Ayalew, L., & Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan . *Geomorphology,* 65, 15-31.

✔ Bovolenta, R., Federici, B., & Natali, D. (2011). Realizzazione automatica di carte di pericolosità da frana superficiale utilizzando il GIS open souce GRASS. - X Giornata Mondiale dell’Acqua. Frane e Dissesto Idrogeologico: consuntivo. *Atti dei Convegni Lincei*, Accademia Nazionale dei Lincei, Roma, ISSN 0391805; ISBN 9788821810367, 169-177.

✔ Cascini, L. (2008). Applicability of landslide susceptibility and hazard zoning at different scales. *Engineering Geology,* 102, 164-177.

✔ Cencetti, C., De Rosa, P., Fredduzzi, A., & Minelli, A. (2010). La valutazione della suscettibilità da frana tramite i metodi dell'Analisi Condizionale e degli Indici di Franosità: un'applicazione tramite GRASS al bacino del T. Ventia (Umbria settentrionale). *Geomatic Workbooks*, 9, 41-54.

✔ Dahal, R.K., Hasegawa, S., Nonomura A., Yamanaka, M., Masuda, T., & Nishino K. (2008). GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology,* 54(2), 314-324, doi:10.1007/s00254-007-0818-3

✔ Dai, FC, & Lee, CF. (2002). Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology,* 42, 213-22.

✔ Fell, R., Corominas, J., Bonnard, CH., Cascini, L., Leroi, E., & Savage, B. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102, 85-98, ISSN 0013-7952

✔ Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology,* 31, 181–216.

- ✔ Hutchinson, J.N., (1995). Landslide hazard assessment. In: *Proc VI Int Symp on Landslides*, Christchurch, 1, 1805-1842.
- ✔ ISPRA (2013). Linee guida per la valutazione del dissesto idrogeologico e la sua mitigazione attraverso misure e interventi in campo agricolo e forestale. ISPRA, Manuali e Linee Guida 85 ISBN 978-88-448-0586-9
- ✔ Lee, S. (2007). Application and verification of fuzzy algebraic operators to landslide susceptibility mapping. *Environmental Geology.* 52, 615-623.
- ✔ Natali, D., & Federici, B. (2010). Realizzazione di carte della pericolosità da frana su aree estese: una procedura automatica in GRASS. *Geomatic Workbooks,* 9, 121-126.
- ✔ Pauditš, P, & Bednárik, M. (2002). Using GRASS in evaluation of landslide susceptibility in Handlovská kotlina Basin. *Proceedings of the Open source GIS - GRASS users conference 2002,* Trento, Italy, 11-13 September.
- ✔ Pradhan, B., & Lee, S. (2010). Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia. *Landslides,* 7, 13–30.
- ✔ Sibson, R. (1980). A vector identity for the Dirichlet tesselation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, 87, 151–155.
- ✔ Sibson, R. (1981). A brief description of natural neighbour interpolation. In Barnett V. (ed.), *Interpreting Multivariate Data*, Wiley, New York, USA, 21–36.
- ✔ Varnes, D.J. and the International Association of Engineering Geology Commission on Landslides and Other Mass Movements (1984). Landslide hazard zonation: A review of principles and practice. *Natural Hazards,* 3, Paris, France. UNESCO, 63p.
- ✔ Watson, D.F. (1992). *Contouring: A Guide to the Analysis and Display of Spatial Data*. Pergamon Press, Oxford, UK.
- ✔ WP/WLI-International Geotechnical Societies UNESCO Working Party on World Landslide Inventory (1993). *Multilingual Landslide Glossary*. Bitech Publisher, Richmond, British Columbia, Canada, 59 pp.

Effective tools for supporting energetic policies: 3d urban models generation and analysis

Filiberto Chiabrando[1], Chiara Danna[1], Andrea Lingua[1], Francesca Noardo[1], Anna Osello[2]

[1] Politecnico di Torino - DIATI, c.so Duca degli Abruzzi, 24 10129 Torino

[2] Politecnico di Torino - DISEG, c.so Duca degli Abruzzi, 24 10129 Torino

Abstract

Following the growing importance of the energy policies, automatic tools become an arising necessity for administrations. The necessary 3D data can be effectively generated by photogrammetry techniques and Structure-from-Motion (SfM) software, and meaningfully managed and analyzed in Geographical Information System (GIS). In this case, for the whole workflow, both commercial and free open source software (OSS) were used. The known advantage of OSS, besides their cheapness, is the possibility to control all the processing steps and to exploit the offered interoperability. Finally in the paper results, problems and limitations of the achieved products were investigated in order to evaluate and confirm the reliability and convenience of OSS.

Keywords

Photogrammetry, Open source, renewable energy, GIS, surface analysis, DSM.

1 Introduction

Energy saving is a very challenging topic in the society and the use of energy from renewable sources is growing. One of the basic needs, for both private and public investors, is the identification and analysis of surfaces suitable for the installation of photovoltaic or solar system for effectively planning a part of the energetic policies. At first, the analysis allows to evaluate the convenience and the feasibility of the investment; secondly, in this way it's possible to plan the use of energy, choosing easily the areas which are best suited for the installation of solar systems. For this reason a system providing reliable results in an automatic or semi-automatic way could play a key role.
For the mapping of the solar radiation, the use of GIS tools is consolidated (Agugiaro et al., 2012; Borfecchia et al., 2013; Nex et al., 2013), as demonstrated by the solar cadastres of some cities (for example New York, Boston, Bolzano, Morgex) available on the web. However, these services are limited in the level of detail of the roof analysis because they don’t evaluate the real available surface and the possible presence of obstacles on the roofs. This deficiency is linked to the fact that these applications consider rather large territories and exploit in some cases existing, not best suited, information. Even when LIDAR information is available, it often doesn’t allow to extract dense

three-dimensional representation, since generally the used resolution is too low to acquire all the necessary details (Bonesso et al., 2013; Borfecchia et al., 2013). For example, the maximum density for a scale 1:2000 point cloud is usually 1 point every 0,80m, which is not enough to capture every useful detail. On the other hand, the use of OSS is going to increase also in the energy application field: they offer suitable tools in order to promote the sharing and exchange of data and information. They have some advantages compared to commercial software, which is important issue both for the public administration, compelled to respect the instructions contained for example in the italian Circular 63/2013 (Agenzia per l'Italia Digitale, 2013), and for the private societies. First of all, OSS permit to save expensive licensing and maintenance costs, moreover they promote the data interoperability by using open formats and implementing international standards. Finally, the knowledge of the source code is a key element to effectively monitor and control the operations of the software. This enables the user to adapt the processing to specific needs and, where necessary, to correct, change or enhance it, contributing to its improvement (Larosa et al., 2008; Neteler et al., 2012). Nevertheless, OSS sometimes present some difficulties to beginners or new users, since they are often less user-friendly than commercial software.
According to this introduction, a workflow for building and analyzing 3D city models has been developed during the DIMMER Project (District Information Modeling and Management for Energy Reduction). DIMMER is a EEUU project of the Seventh Framework Programme for research and technological development realized in Manchester and Turin. In this project (2013/2016) several partners such as industries, research centers and universities are involved, including the Politecnico di Torino. The aim of DIMMER, is the realization of an open platform for real time data processing, visualization and monitoring of energy consumption and production from renewable sources. In this paper the proposed semiautomatic methods for the generation of a 3D model of a portion of Torino for energy purposes is described. Moreover, the approach of data analysis in a GIS platform is described as well. For complying with the italian Circular 63/2013 and for exploiting the OSS advantages, all the process should be performed using open source tools. Consequently, as a first step, the reliablity of OSS products has been tested through the comparison with the similar ones produced by well known commercial tools. Then, a repeatable workflow has been developed for simplifying the use of OSS to inexpert users.

2 Test area and data sources

The test area (approximately 0,2 km^2) is a portion of the "Crocetta" district in Torino. It is quite flat, but the buildings present different size and geometry.
A photogrammetric approach was followed for the realization of the 3D model. The initial data were 10 aerial images acquired in 2007 for the realization of the orthophoto of the Piemonte Region, from a flight height of about 900 m, aiming at a nominal scale of 1:2000: 0,40m for both planimetric (tolerance 0,69 m) and heights (tolerance 0,80m) accuracy (Brovelli et al., 2010). The data were acquired by a Zeiss Imaging (Z/I) Digital Mapping Camera (DMC): a turnkey, frame-based digital aerial camera system designed to support aerial

photogrammetric missions demanding high-resolution and accuracy. The DMC uses eight individual lenses rather than a single lens. The lenses operate simultaneously and collect color imagery (RGB), CIR (color infrared), and black-and-white panchromatic imagery. After the flight the captured image data are normalized, verified, rectified, colour coded, formatted and made available for digital photogrammetric production. The first post-processing step is the generation of Level 1 images that include the elimination of defect pixels and a normalization (Diener et al., 2000). Next, the image data are geometrically converted and mosaiked. Finally the virtual image is generated, the lens distortions are geometrically corrected based on the camera calibration and the images created in this way have a new, virtual camera constant and can be considered ideal central perspective photogrammetric images. The employed images were supplied with the interior calibration and the external orientation parameters: pose coordinates (X,Y,Z) and attitude (ϕ,ω,κ). The exterior parameters were calculated during the bundle block adjustment in a traditional digital photogrammetric software combining the information derived from the on board sensors (GNSS and IMU) and several GCPs previously measured on the terrain.

3 Data processing

As is mentioned before, the data processing steps were divided in two parts: the realization of three-dimensional models and its analysis, taking in account the real morphology of the buildings and the weather conditions. Since the solar radiation received by the surfaces is strongly influenced by their three-dimensional conformation, a DSM (Digital Surface Model) of the area is needed in order to perform correct analysis.
For having a reference product useful for checking the reliability of the results obtained through the image-matching approach, a traditional stereoscopic plotting was realized using Z-Map software (www.mencisoftware.it) in order to obtain a correct 3D drawing of the area. It was then converted in a Dense DSM and in a dense Point Cloud using an home made software “DDSM to dxf” (Lingua et al., 2010).

4 Image-matching

The image-matching methodology was firstly developed and tested for Remote Sensed data. At first it has been planned to meet orientation solutions and then to perform DTM/DSM (Digital Terrain Model / Digital Surface Model) extraction from aerial or satellite stripes (Marr & Hildreth, 1980; Torre & Poggio, 1980; Ackermann, 1984; Grun, 1985; Baltsavias, 1991). Recently it is extensively used in close-range application concerning architectural and archaeological survey (Chiabrando et al., 2014; Bastonero et al., 2014; Toldo et al., 2013; Kersten & Lindstaedt, 2012; Remondino & Menna, 2008).
The developed technique were tested here in a urban context. The images have been processed automatically using two software tools: the commercial low-cost software Agisoft PhotoScan (www.agisoft.ru) and the open source suite Apero–MicMac implemented by IGN (Institut National de l’Information Géographique et Forestière, Paris) (Pierrot-Deseilligny & Cléry, 2011). In both

cases, the DMC aerial images and their external and interior orientation parameters were used in order to orient the images, to generate the point cloud (Figure 1a,1b), the dense DSM (Figure 1c,1d) and, finally, the orthophoto.

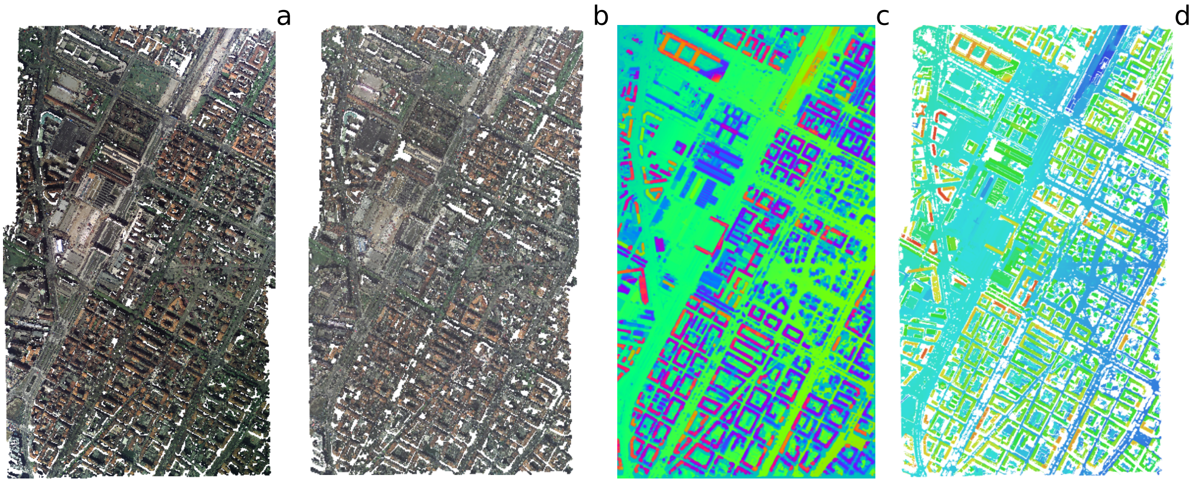


Figure 1: Colored 3D point cloud (left) by MicMac (a) and by PhotoScan (b) and dense DSM (right) by MicMac (c) and by PhotoScan (d).

The data-processing has been performed on the total area covered by the images and it required a few hours to generate the final results (two hours with MicMac and three hours with PhotoScan).

In order to perform the image matching process the internal orientation parameters (provided in the calibration file) were fixed. This was necessary since the DMC image are provided without any exif file, and their “virtual” geometry doesn’t permit the automatic evaluation of the correct focal length. Therefore, the employed software can’t calculate the calibration parameters automatically. In particular Agisoft doesn’t work and MicMac calculates a totally wrong focal length (7.5mm instead of 120mm).

For the point cloud and DSM generation the typical data processing workflow of each software was used. In PhotoScan high resolution Tie Point Extraction and Dense Matching was set up; as a consequence a final 65.838.239 point cloud was achieved and subsequently the DSM was generated. In MicMac using the Ortho option in Malt (simplified MicMac tool), a dense 50.664.687 point cloud was realized. The DSM (memorized as integer values for optimizing the memory use) were converted (section 5.2) in float values useful for data comparison. Analyzing the graphical results some noisy areas are visible in correspondence to shadowed zones and dark roofs (Figure 2).

Figure 2: Portion of an aerial image (a) with a dark roof (red square) and the lack of corresponding information in the generated point cloud (b).

This noisiness is quite common in similar cases due to the various factors conditioning the matching proccess, such as radiometric quality of the images, image overlapping, shadows and texture (or external features) of the buildings.

5 Results analysis and comparisons

The DSMs and 3D point clouds obtained from image matching algorithms (MicMac and Photoscan) were compared to the achieved roof shape generated by the manual plotting, in order to evaluate their precision and accuracy.
Since these last data, used as reliable reference, are related to a restricted area of interest, this one was considered on the image matching results, in order to exclude the external parts from the analysis, avoiding misleading errors.

5.1 Point clouds vs point clouds

The comparison between the point clouds has been performed using CloudCompare, OSS that allows to analyze and manage 3D models and point clouds by means of several tools (Girardeau-Montaut, 2011).
The distances (and thus the discrepancies) between the clouds generated by the matching software and the reference deriving from the manual plotting were automatically computed, obtaining both numerical (statistical data in table 1) and graphical results (a thematic point cloud were differences are mapped). The gross errors were removed and not considered for the analysis.

	MicMac	*PhotoScan*
Minimum distance [m]	0	0
Maximum distance [m]	2.960	3.687
Mean distance [m]	0.815	0.565
Standard deviation [m]	0.486	0.597

Table 1: Statistic values related to the comparison between the 3D point cloud produced by MicMac and Photoscan and compared to manually extracted data.

As final consideration, according to the standard deviation reported in Table 1, it is possible to state that the point cloud computed by MicMac (OSS) appears even more accurate than the one obtained by PhotoScan.

5.2 DSM vs DSM

The other analyzed products were the DSMs. Photoscan can easily give this data (using an unknown algorithm), while for MicMac some intermediate passages are needed. One way is to interpolate in a GIS software (in this case QGIS with a linear interpolation algorithm has been used) the dense point cloud, whose correctness and accuracy were verified before. Alternatively, a DSM can be built by MicMac (which use a linear interpolation as well). This last opportunity presents thus a problem: the raster values must be corrected, since in the optimization process MicMac calculates the height using the voxel cell (height resolution) and establishes an arbitrary height origin, when it produces the TIFF file. The transformation has been applied through the raster calculator tool in QGIS, following the relation “(raster value * height resolution) + height origin”. The two DSMs have been compared to the reference one

(deriving from the manual plotting) using QGIS raster calculator tools.
QGIS (www.qgis.org) is a free and OS desktop GIS (subject to the GNU General Public License), which presents a user-friendly interface and can perform spatial analysis and processing typical of GIS. It presents some advantages also for the huge number of tools exploitable in the software, fulfilling customized processing needs. The typical QGIS functionalities can be enhanced by other specific python programmed plug-ins, or integrated tool sets from different platforms, such as GRASS (Geographic Resources Analysis Support System), SAGA (System for Automated Geoscientific Analysis), Orfeo Toolbox, and so on.
In order to focus on the surface of the roofs, which are our main interest, some preliminary operations are performed. Only the roofs are extracted by masking the building with the areas defined in the Turin digital map (scale 1:2000).

	MicMac	*MicMac + QGIS*	*PhotoScan*
Minimum distance [m]	-21.518	-20.990	-21.440
Maximum distance [m]	35.095	37.155	35.204
Mean distance [m]	0.593	0.666	0.245
Standard deviation [m]	1.825	3.290	1.228

Table 2: Statistic values related to the comparison between the DSM

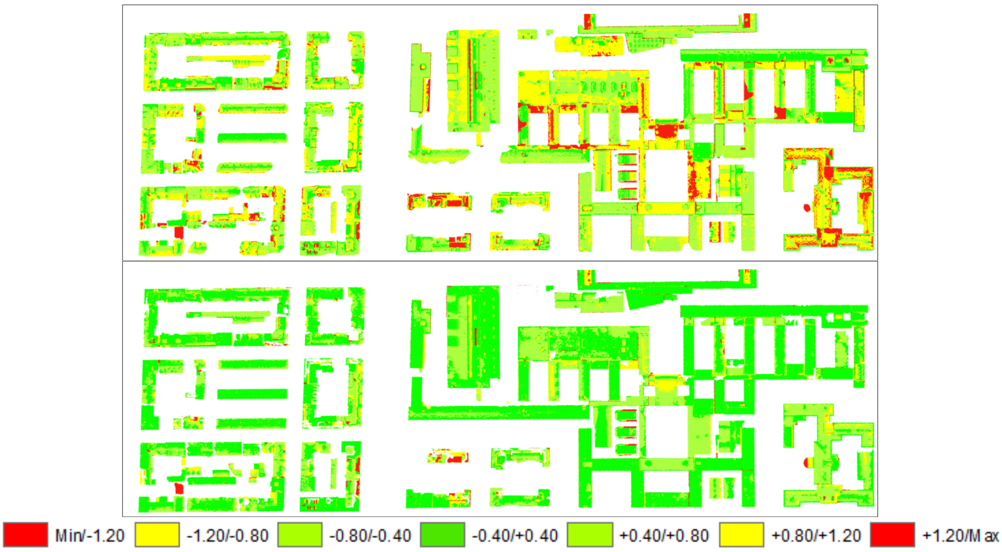


Figure 3: Comparison between DSMs: Manual/MicMac (upper) and Manual/PhotoScan (lower)

The discrepancies (Figure 3) have been calculated through the “raster calculator” tool, able to make calculations using raster values. Analyzing the values of standard deviation (Table 2), it is possible to state that the PhotoScan DSM is the nearest to the reference one. Checking the two DSM deriving from MicMac data, we can observe that on most part of the surfaces the results are quite similar, but some roofs (in particular some shadowed or dark ones) reports some problems. In fact, these are not included in the MicMac DSM, while the DSM interpolated in QGIS from the MicMac point cloud presents big errors in that areas (up to 20 m). For this reason we chose the DSM produced directly by MicMac, in order to prevent that the wrong areas could affect the other analysis.

6 Analysis of roof surface and extraction of information in GIS

Starting from the digital models, in GIS software, it is possible to extract information about the aspect (Figure 4), the slope (Figure 5) and the solar radiation (Figure 6) using different tools. These are important and useful information for planning energy saving policies and interventions. The mentioned processing tools are typical in GIS, both commercial and OS. In this case QGIS has been used again, in particular by exploiting the GRASS GIS algorithms. The reliability of the results of the different modules available in GRASS GIS were used and tested in a variety of studies (Hofierka & Šuri, 2002; Šuri & Hofierka, 2004). Obviously, the data extracted from the manual DSM are less noisy than the other ones, but following some punctual checks, we can conclude that the data extracted by the image matching approach are quite similar, and, thus, equally reliable.

To perform the analysis of solar irradiation the command *r.sun - Solar irradiance and irradiation model* has been used (Agugiaro et al.,2012). This module calculates five different maps (*Irradiation layer, Insolation time, Diffuse irradiation layer, Ground reflected irradiation layer and Global irradiance / irradiation layer*) in clear-sky conditions, so it does not take into consideration the presence of clouds. The analysis is not only based on the DSM, but it is a function of many factors such as aspect, slope, latitude, longitude, declination value, air turbidity coefficients (Linke data), the percent of diffuse component of the radiation, and so on. This information (except for the shadows mask) must be manually introduced in the shell for running the algorithm, so that it could be more difficult than a similar process run by other commercial software, but it permits to monitor each step of the processing. Studies also prove how this software offers more correct and precise results than other commercial ones (Nex et al., 2013). This is probably due to the ever-improvement by the community. In the analyzed case the process were performed in the days of solstice and equinox and the shadows are incorporated within the results.

On the other hand, some parts could be difficult to non-expert user, also because a problem of the available plug-in is that they often need specific formats in input data. For example, the plug-in which computes the solar radiation require the raster of the latitudes and longitudes expressed in decimal degrees, but a specific tool for this task doesn’t exist yet. This requires several steps, each of whom needs specific parameters to be set. Therefore, a graphical model has been defined through the expressed module in QGIS. This allows to create complex models using a simple and easy-to-use interface: one can define the sequence of plug-in to run on input data, with pre-set parameters to reach the needed result. This permits to save time and efforts for repeating the same process different times. Furthermore, the model could be shared with different users. The only condition is that the used plug-in must be loaded and activated in the software which uses it. After some attempts using similar plug-ins that could be considered as alternatives to a first analysis, the model was defined. It could be enhanced for including the full processing (from the DSM to the generation of slope and aspect maps, to the computing of the solar radiation map). Surely it could be useful if these analysis would become a “good practice” for Public Administrations, that could rapidly obtain useful data.

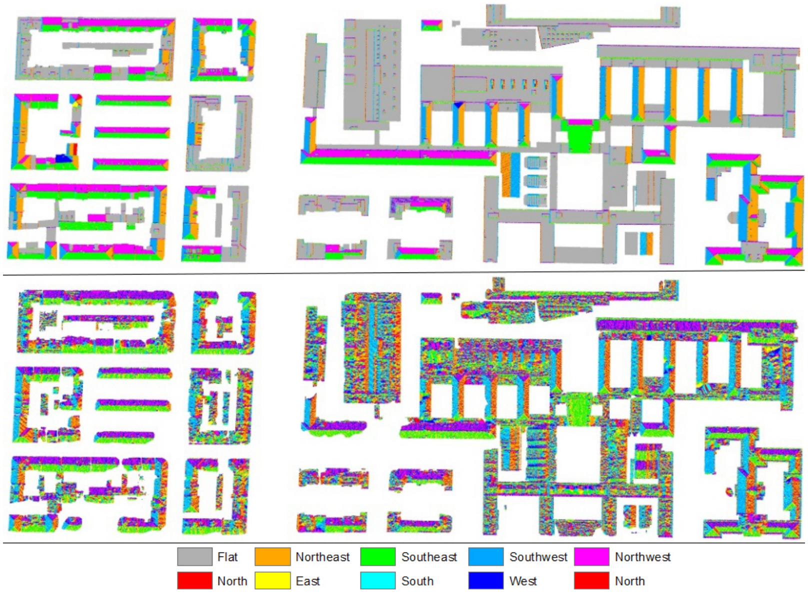


Figure 4: Analysis Aspect with manual DSM (upper) and MicMac DSM (lower)

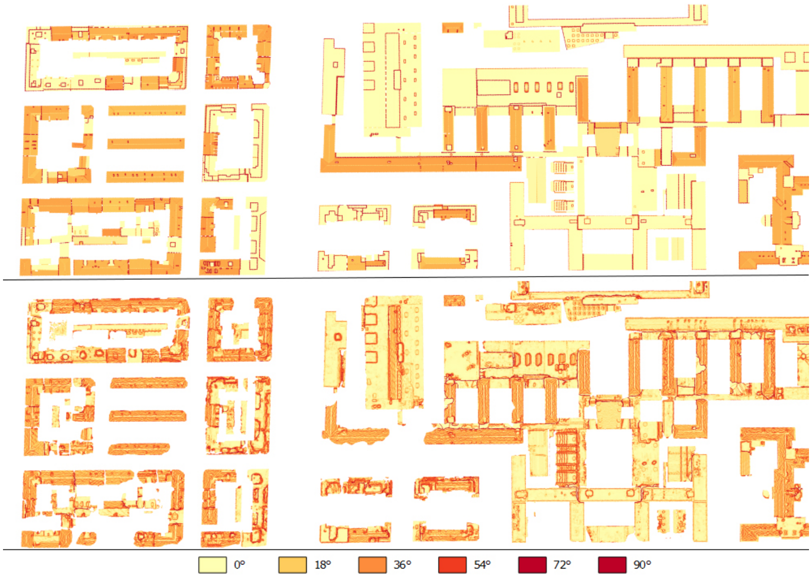


Figure 5: Analysis Slope with manual DSM (upper) and MicMac DSM (lower)

Figure 6: Analysis solar radiation. From this maps it is possible to see that in this analysis they have succeeded in considering the shades related to the chimney or the other three- dimensional volumes. The precision is related to the DSM resolution.

Finally, in order to verify if the obtained values of solar radiation were reasonable, in terms of numbers and trends, a comparison with the values established by the UNI 10349 for the city of Turin has been performed. For this test four samples of points having the same aspect values has been queried; the average of the obtained values were compared to the norm ones, as shown in Figure 7 left . Here it is clear that the solar radiation values computed by QGIS can be considered reliable.

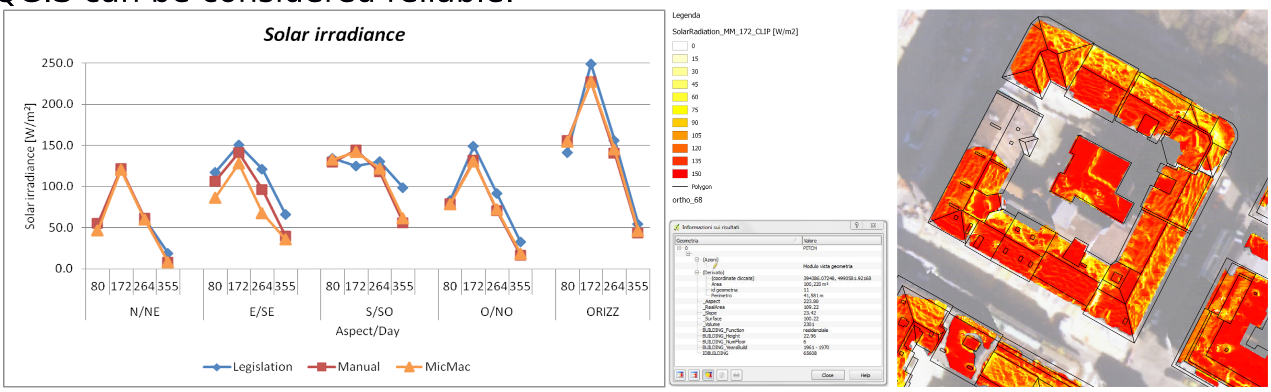


Figure 7: Values of the solar radiation (left) and Geographical Information System (right).

7 The management of the information in GIS

Public Administration or private users would use the resulting data to plan interventions. For effectively manage them in a system which permits queries, analyses and automatic maps, a part of a GIS has been built, which could be integrated in wider databases. The results of our processing have been stored in. In particular, the raster maps concerning terrain elevation, aspect, slope and solar radiation, have been inserted within the GIS. The DSM processed in MicMac was used as final products in order to adopt a workflow that can be re-realized by anyone, with the advantages stated before.

In order to describe the information related to the whole roof pitches or buildings, vector layers were inserted (Pitch and Building). In this case this have been obtained by manual plotting on the base of processed products.

Through the command Zonal Statistics in QGIS it has been possible to import as attribute of the pitches the values of the raster maps (e.g. slope and aspect). Moreover, other attributes regarding the geometry of the polygons are computable with a Field Calculator tool. The Pitch entity can be queried and immediately provides a set of information such as slope, aspect, surface and real area (function of the projected area and slope). The Pitch entity is connected to the Building entity, which provides attributes concerning, for example: year of construction, destination of use and number of floor, etc. In addition, the raster maps always give punctual information (Figure 7, right) which can be superimposed.

Conclusions

The analysis of a portion of Torino was performed with the aim of extracting useful information for planning interventions towards the improvement of renewable energy exploitation.

In the research, the complete workflow was considered, from the generation of the 3D models to the data management in a GIS. Some OSS were used and verified through the comparison with other commercial software and the data manually produced used as reference.

We can conclude that the 3D model produced by the OSS MicMac is similar to both the PhotoScan one and the one deriving from the manual plotting. Also the results of the analysis in QGIS/GRASS are validated in comparison to the official values and the one produced by other commercial software, so that it’s possible to affirm that OSS are suitable at all for this kind of application.

Actually, one of the main limitations of OSS is the major difficulty in their use, which is due sometimes in the missing of graphical interface (as is the case of MicMac) and in the necessity of inputting specific parameters without admitting different formats or semantics in the operations. Moreover, for example in QGIS, some plug-in often in beta version could sometimes give problems. On the other hand, this could also be a strong point considering the number of plug-ins performing the same analysis but with some slight differences, which permit to adapt the processing to one’s specific needs.

In future work it would be interesting to test other less expensive methods for data acquisition, such as UAV with low cost camera which would even simplify the photogrammetric process, even if some topic regarding the calibration would be opened. Furthermore, the next necessary step will be the publication of the GIS on the web. This could be done directly through QGIS Cloud, which is the WebGIS connected to QGIS, or using different tools, such as the open source platform Geonode, Geoserver, and so on. It would be extremely useful to manage all these data together with monitoring data measured on the interested areas, for checking or adjusting the processed data. These next step could effectively integrate this first part, which is, anyway, the indispensable starting point.

References

✔Ackermann, F. (1984). Digital image correlation: performance and potential application in Photogrammetry. *The Photogrammetric Record*, 11(64), 429–439. doi:10.1111/j.1477-9730.1984.tb00505.x

✔Agenzia per l'Italia Digitale – Presidenza del Consiglio dei Ministri, Circolare 6 dicembre 2013 n.63 - Linee guida per la valutazione comparativa prevista dall'art. 68 del D.Lgs. 7 marzo 2005 n. 82 "Codice dell'amministrazione digitale"

✔Agugiaro, G., Nex, F., Remondino, F.,De Filippi, R., Droghetti, S., & Furlanello, C. (2012). Solar radiation estimation on building roofs and web-based solar cadastre. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-2, 177-182. XXII ISPRS Congress. doi:10.5194/isprsannals-I-2-177-2012

✔Baltsavias, E. P. (1991). Geometrically constrained multiphoto matching, Mitteilungen Nr. 49, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland. Retrieved from http://www.igp-data.ethz.ch/berichte/Blaue_Berichte_PDF/49.pdf

✔Bonesso, S., Mutani, G., Hubina, T., & Ramassotto, A. (2013). L'applicazione delle tecnologie fotovoltaiche integrate sulle coperture degli edifici con software GIS. *GEOmedia*, n. 2/2013, 6 – 11. Retrieved from http://mediageo.it/ojs/index.php/GEOmedia/article/viewFile/268/241

✔Borfecchia, F., Pollino, M., De Cecco, L., Martini, S., La Porta, L., Marucci, A., & Caiaffa, E. (2013). Integrated GIS and Remote Sensing Techniques to Support PV Potential Assessment of Roofs in Urban Areas. *Computational Science and Its Applications–ICCSA 2013*, 422-437. doi:10.1007/978-3-642-39646-5_31

✔Brovelli M.A., Cilloccu F., Crespi M., Dequal S., Lingua A.M. (2010). *Ortoimmagini 10K e modelli altimetrici – Specifiche tecniche*, CISIS.

✔Brovelli M.A., Cina A., Crespi M., Lingua A.M., Manzino A. (2010). *Ortoimmagini e modelli altimetrici a grande scala – Linee guida*, CISIS.

✔Bastonero, P., Donadio, E., Chiabrando, F., & Spanò, A. (2014). Fusion of 3D models derived from TLS and image-based techniques for CH enhanced documentation. *ISPRS ANNALS OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES*, *2*, 73-80.

✔Chiabrando, F., Lingua, A., Noardo, F., & Spanò, A. (2014). 3D MODELLING OF TROMPE L'OEIL DECORATED VAULTS USING DENSE MATCHING TECHNIQUES. *ISPRS ANNALS OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES*, *2*, 97-104.

✔Diener, S., Kiefner, M., & Dörstel, C. (2000). Radiometric normalisation and colour composite generation of the DMC. *International Archives of Photogrammetry and Remote Sensing,* XXXIII, Part B1, 82-88. Retrieved from http://www.isprs.org/proceedings/XXXIII/congress/part1/82_XXXIII-part1.pdf

✔Girardeau-Montaut, D. (2011). CloudCompare-Open Source project. OpenSource Project. Retrieved from http://www.danielgm.net/cc/

✔Grüen, A. (1985), Adaptive least square correlation: a powerful image matching technique. *South African Journal of PRS and Cartography*, 14(3), 175-187. Retrieved from http://www.idb.arch.ethz.ch/files/alsm_awgruen.pdf

✔Hofierka, J., Kaòuk J., & Gallay, M. (2014). The spatial distribution of photovoltaic power plants in relation to solar resource potential: the case of the czech republic and Slovakia. *Moravian Geographical Reports*, n. 2/2014, 22, 26-33. doi:10.2478/mgr-2014-0009

✔Hofierka, J., & Šuri, M. (2002). The solar radiation model for Open source GIS: implementation and applications. *Proceedings of the Open source GIS - GRASS users conference (pp. 1-19). Retrieved from* http://www.ing.unitn.it/~Grass/conferences/GRASS2002/proceedings/proceedings/pdfs/Hofierka_Jaroslav.pdf

✔Kersten, T. P., & Lindstaedt, M. (2012). Automatic 3D object reconstruction from multiple images for architectural, cultural heritage and archaeological applications using open-source software and web services. *Photogrammetrie, Fernerkundung Geoinformation* (PFG), 2012(6), 727–740. doi:10.1127/1432-8364/2012/0152

✔Lingua, A., Nex, F., & Rinaudo, F. (2010, October). Integration of airborne laser scanner and multi-image techniques for map production. In Remote Sensing (pp. 783107-783107). International Society for Optics and Photonics. Retrieved from http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=725867

✔Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences.* 207(1167), 187-217. Retrieved from http://rspb.royalsocietypublishing.org/content/207/1167/187

✔Neteler, M., Bowman, M.H., Landa, M., Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Enviromental Modelling & Software*, 31, Maggio 2012, 124-130. doi:10.1016/j.envsoft.2011.11.014

✔Nex, F., Remondino, F., Agugiaro, G., De Filippi, R., Poletti, M., Furlanello, C., Menegon, S., Dallago, G., & Fontanari, S. (2013). 3D SolarWeb: a solar cadaster in the italian alpine landscape. I*nternational Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XL-7/W2, 2013. doi:10.5194/isprsarchives-XL-7-W2-173-2013

✔Deseilligny, M. P., & Clery, I. (2011). Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38*, 5. Retrieved from http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XXXVIII-5-W16/269/2011/isprsarchives-XXXVIII-5-W16-269-2011.pdf

✔Remondino, F., & El-Hakim, S. (2006). Image-based 3D modelling: a review. *The Photogrammetric Record*, 21(115), 269–291. doi:10.1111/j.1477-9730.2006.00383.x

✔Remondino, F., & Menna, F. (2008). Image-Based surface measurement for close-range heritage documentation. I*nternational Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. XXXVII, part B5-1, 199-206. Retrieved from

http://3dom.fbk.eu/sites/3dom.fbk.eu/files/pdf/remondino_menna_ISPRS08.pdf
- ✔Šuri, M., & Hofierka, J. (2004). A New GIS-based solar radiation model and its application to photovoltaic assessments. *Transactions in GIS*, 8, 175-190. doi:10.1111/j.1467-9671.2004.00174.x
- ✔Toldo, R., Fantini, F., Giona, L., Fantoni, S., & Fusiello, A. (2013). Accurate multiview stereo reconstruction with fast visibility integration and tight disparity bounding. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(5/W1), 243–249. doi:10.5194/isprsarchives-XL-5-W1-243-2013
- ✔Torre, V., & Poggio, T. A. (1986). On edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2), 147-163. Retrived from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4767769&tag=1
- ✔UNI 10349, Riscaldamento e raffrescamento degli edifici – Dati climatici, Aprile 1994.

Promoting slow tourism through FOSS4G Web Mapping: an Italian-Swiss case study

Milan Antonovic[1], Maria Antonia Brovelli[2], Massimiliano Cannata[1], Mirko Cardoso[1], Candan Eylül Kilsedar[2], Marco Minghini[2], Giorgio Zamboni[2]

[1] Institute of Earth Sciences, SUPSI, Campus Trevano, 6952 Canobbio, Switzerland

[2] Politecnico di Milano, Como Campus, DICA, Via Valleggio 11, 22100 Como, Italy

Abstract

Slow tourism defines a sustainable way of experiencing a territory based on environmentally-friendly forms of transportation, the appreciation of nature and the rediscovery of local cultural traditions. Advancements in geo-information technology have opened new possibilities for promoting this practice. A slow tourism ongoing project focused on the cross-border area between Italy and Switzerland is presented. Several FOSS4G-based Web Mapping applications are developed which address different users and feature different functionalities. The applications include: a mobile app allowing tourists to report Points of Interest (POIs) along the paths; a mobile app enabling professionals to survey new paths; a traditional 2D Web viewer providing access to and interaction with the project data; an application offering a virtual tour along the paths; and a virtual-globe based 3D viewer with participative functionalities.

Keywords

FOSS4G, mobile, slow tourism, virtual, Web Mapping.

1 Introduction

Slow tourism has recently acquired great interest. Based on environmentally-friendly forms of transportation like hiking, cycling and horse-riding, it pursues the enjoyment of the territory coupled with a (re)discovery of its local culture (e.g. people lifestyle, history, art, architecture, religion and cuisine). “The Paths of Regina” (http://www.viaregina.eu) is an INTERREG project (Cross-border Cooperation Operational Programme Italy–Switzerland 2007-2013) aiming at the valorisation of slow tourism in the beautiful, hilly area at the boundary between Italy and Switzerland. Thanks to the synergy between Swiss and Italian experts in geomatics, landscape designers, administrations and cultural associations, the project seeks to create new tools to rediscover the European cultural identity of the area.

2 FOSS4G Web Mapping

Traditional tourism tools providing knowledge of the territory (e.g. paper maps and signage along the paths) are coupled with innovative, ad hoc Web Mapping applications developed within the project. They are fully built with FOSS4G and take into account the new trends of mobile visualization, user participation and crowdsourcing, virtual reality and multi-dimensional view. The applications, that are partially still under development, are separately described in the following Subsections. It is worth to state that the geospatial data used were derived from the administrations involved in the project, ad hoc field surveys by hikers associations (especially for paths and POIs along them) and the OpenStreetMap (OSM) database.

2.1 2D Web viewer

FOSS4G server-side technologies deployed among Italy and Switzerland for the 2D Web viewer take advantage of the latest implementations of GeoServer and ZOO WPS. The client application benefits instead of the latest OpenLayers 3 JS library. Besides simple display of data and more advanced functionalities like geospatial routing and terrain profile computation, it enables navigation to two other applications, i.e. virtual tour and 3D Web viewer detailed in Subsections 2.4 and 2.5, respectively. Routing applied to the cross-border path networks makes use of Dijkstra's algorithm and is performed by pgRouting. Optimal path between two user-selected points is computed based on its cost, which in turn depends on the user input about the types of POIs (e.g. religious buildings, museums, etc.) and their relative weights. This way, the returned path is the most suitable to the user interests. The returned path is the input for terrain profile computation which is based on ZOO GdalExtractProfile service. The display of the chart is then supported by external JS libraries such as Highcharts and jQuery. Before walking on the field users can print a customized page including the returned path, its terrain profile and some additional related information like the total distance, the average slope and the height difference between the start and end points. Lastly, POI data are enriched with TripAdvisor information retrieved by its official API. When clicking a certain position on the map the attractions, hotels and restaurants from TripAdvisor located within a predefined radius are returned to the user together with the related information which is accessible inside a popup.

2.2 Mobile app for slow tourists

A second application involves people in contributing the local knowledge they acquire during slow tourism experiences. Open Data Kit (ODK) is exploited to manage field data collection. This suite is mainly composed of a server module (ODK Aggregate), running under Tomcat and synchronized to a PostgreSQL database to store the data; and an Android app (ODK Collect) allowing users to report POIs (ranging from historical/cultural points of interest to tourism services and morphological elements) through their mobile devices (see Figure 1). Users have to enter the details of the POI (e.g. name, type and description), register its position (e.g. using the device GPS), and take a picture (and optionally also a video and audio) of it. Thanks to the PostGIS extension of the PostgreSQL database, GeoServer can then read and publish this crowdsourced information which is made available as a separate layer in the Web viewer described in Subsection 2.1. Besides being a typical feature of GeoWeb 2.0 (Maguire, 2007) the bottom-up approach where citizens are not only consumers but also active producers of geospatial information, or Volunteered Geographic Information (Goodchild, 2007), represents a major innovation in slow tourism projects. The app was successfully tested during a number of mapping parties performed all along the project area (see e.g. the data collected during a mapping party in Cernobbio at http://viaregina.como.polimi.it/mapparty).

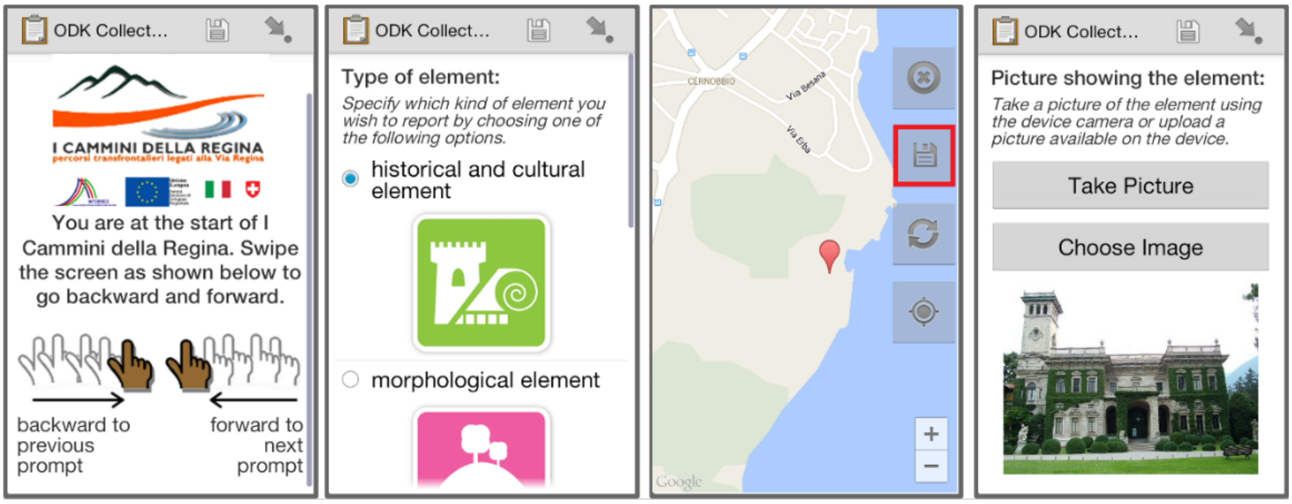


Figure 1: Sample screenshots of the mobile app for slow tourists.

2.3 Mobile app for professionals

During the last decades Switzerland has developed a well-defined procedure to map historical routes. However its application relies only on the manual work of professionals without a specific supporting instrument. An application named GeoIVS was thus developed to fill this gap. Before entering the survey mode, it asks for the definition of the path and its initial surface type. The survey mode interface, which also works offline, is organized into three columns split into two horizontal sections (see Figure 2).

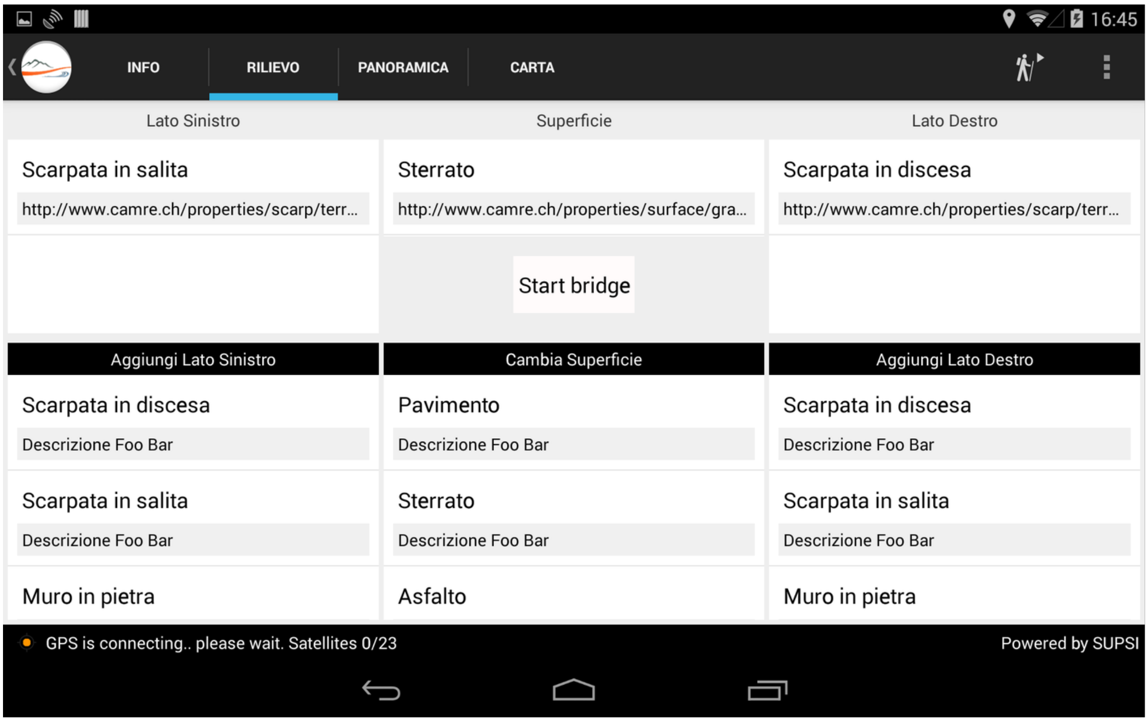


Figure 2: Sample screenshot of the Android virtual tour application.

Left column represents the elements on the left side of the path, right column represents the elements on the right side of the path, while central column represents the path surface. For each column, the lower section shows all the possible elements the user can enter while the upper section represents the elements being registered. Tapping on an element starts its registration by acquiring the starting point; sliding of an element under registration acquires its ending poi1t and stops its registration. Starting and ending points are recorded using the device GPS and specific algorithms which search for the best accuracy in a user-defined period. Linear referencing is then applied to determine the element geometry. Textual and vocal notes as well as pictures and videos can be also associated to each surveyed element. The survey results in a JSON file, which encapsulates all the information and – as soon as a connection is available – is sent to a server for storage, processing and publication. From the technological point of view the application relies on FOSS software like OpenLayers 3, Android SDK, JTS Topology Suite and proj4js.

2.4 Virtual tour application

The objective of this application is the provision of cultural information in an attractive context in order to transform a path without outstanding attractions in an emotional experience. To achieve this goal, an integrated framework was adopted which combines geospatial technologies, the concept of virtual tour, the use of multimedia and the location-based approach. Thanks to the collaboration with designers and experts of the territory (historians, geologists, ethnographists and biologists), the two test paths selected for the project (Sagno-Bisbino-Cernobbio and Bellinzona-San Jorio-Menaggio) were studied and represented in two storyboards. They highlight the cultural aspects of the areas traversed by the two paths through the use of thematic POIs. From a design perspective, the application represents the virtual location of the tourist with a red circle that moves along the path and that is synchronized with the altitude profile at the bottom of the page (see Figure 3).

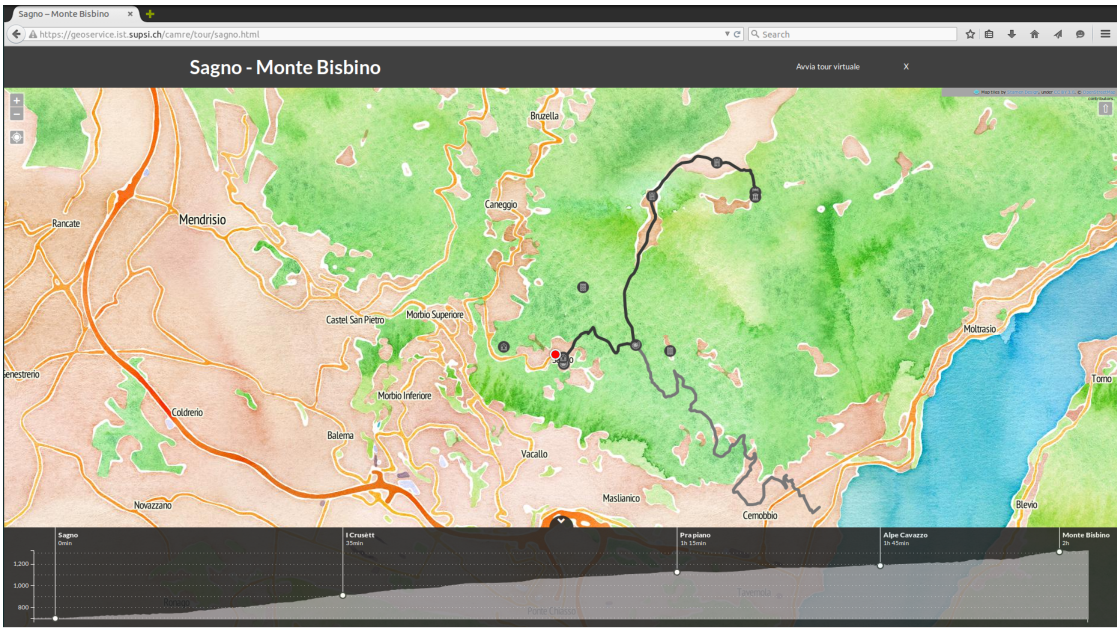


Figure 3: Sample screenshot of the virtual tour application.

When a POI is intercepted, instead of a classical popup a full-screen animation showing sliding bands over a series of background images is activated. Each band has a title, a subtitle and a description which are narrated in an audio. Bands are enriched by multimedia, which currently consist of images, slideshows, videos and photospheres. Notification for users exploring the path on the field is currently under development. The application relies on open source technologies, open formats and open data such as OpenLayers, Bootstrap, JQuery, D3.js, Less.js, GeoJSON, and OpenStreetMap.

2.5 3D Web viewer

Unlike the classic 2D visualization, multi-dimensional viewers like virtual globes offer fully-realistic content visualization that is particularly suitable for tourism purposes (Brovelli, Hogan, Minghini, and Zamboni, 2013). A 3D Web viewer based on the NASA World Wind virtual globe was developed which aims at the valorisation of tourism paths on the project area (see Figure 4). Moreover, the new trend of participation was exploited within the viewer allowing users to add relevant contents (photos, videos, documents and notes) which become heritage of the whole community. A first level of information is the map of paths, which is superimposed to one among several user-selected base maps. In addition to the servers specifically deployed within the project, the application can transparently connect to any WMS compliant server available on the network. A second level of information consists of the crowdsourced data collected by professional hikers as well as occasional walkers, tourists, citizens, etc. The platform has in fact the innovative capability to connect to the ODK server of the project, and to publish the related POIs reported by users through their mobile devices (see Subsection 2.2). When clicking on a POI, its field-collected information is displayed. Visualization of the field-reported POIs can be also filtered in time through the use of a slider.

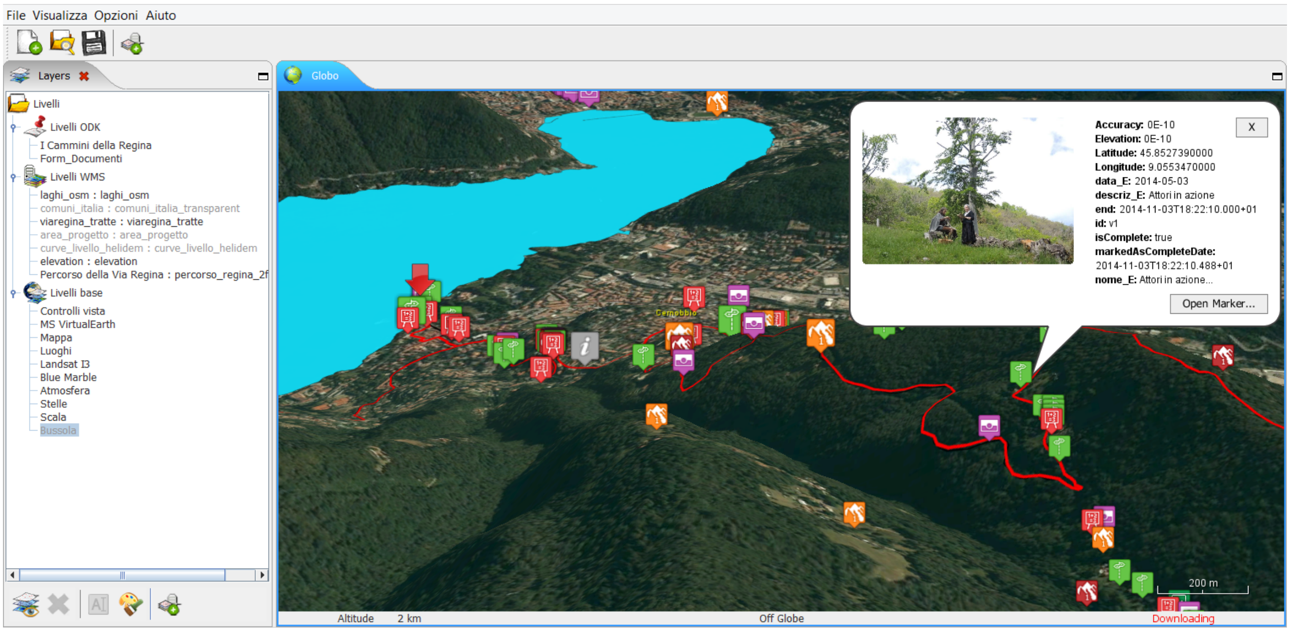


Figure 4: Sample screenshot of the 3D Web viewer.

3 Conclusions

In contrast to the current mass-process phenomenon tourism has largely become, the practice of slow tourism represents a challenging model based on sustainability, environmental friendliness and discovery of local knowledge. In the frame of a slow tourism project focused on a cross-border area among Italy and Switzerland, a set of Web Mapping applications are developed to valorize the local cultural richness through the exploitation of modern technological trends including mobile data collection, virtual reality and 3D visualization. This set of applications greatly extends the geospatial tools typically developed for slow tourism projects and mainly consisting of simple 2D Web viewers. A major point of innovation is represented by the involvement of users, who are not just consumers but also contributors of new data. From the technical perspective, the use of FOSS4G technologies was fundamental as they allowed to develop customized products according to the project needs.

Acknowledgements

This study has been funded by the project “I Cammini della Regina – Percorsi transfrontalieri legati alla Via Regina” (“The paths of Regina – Cross-border paths linked to Via Regina”) within the INTERREG Cross-border Cooperation Operational Programme Italy–Switzerland 2007-2013 (ID 33829732).

References

- ✔ Brovelli, M. A., Hogan, P., Minghini, M., & Zamboni, G. (2013). The power of Virtual Globes for valorising cultural heritage and enabling sustainable tourism: NASA World Wind Applications . *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-4/W2, 115-120. doi: 10.5194/isprsarchives-XL-4-W2-115-2013.
- ✔ Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal*, *69*(4), 211-221. doi: 10.1007/s10708-007-9111-y.
- ✔ Maguire, D. J. (2007). GeoWeb 2.0 and volunteered GI . In *Workshop on Volunteered Geographic Information, University of California, Santa Barbara*, 104-106.

Reproducible computing with Tzar and Docker

Lucy Bastin[1,3], Javier Martínez-López[1], Ascelin Gordon[2], Bill Langford[2], River Satya[2]

[1] Joint Research Centre of the European Commission, Ispra, Italy

[2] Royal Melbourne Institute of Technology, Melbourne, Australia

[3] Aston University, UK

Abstract

The potential for sharing environmental data and models is huge, but can be challenging for experts without specific programming expertise. We built an open-source, cross-platform framework ('Tzar') to run models across distributed machines. Tzar is simple to set up and use, allows dynamic parameter generation and enhances reproducibility by accessing versioned data and code. Combining Tzar with Docker helps us lower the entry barrier further by versioning and bundling all required modules and dependencies, together with the database needed to schedule work.

Keywords

Reproducible computing, model interoperability, distributed computing

1 Introduction and context

The growing availability of open data and models, shared through portals and commonly agreed standards, offers huge potential for environmental decision-making. The explosion in quantity and variety of discoverable resources has been accompanied by a welcome increase in tools which can allow users to share, chain and reproduce analyses; (for example OpenMI [1], Taverna[2] and Kepler[3]) and tools to assist in documenting the exact steps by which results were achieved – i.e., the provenance of a result (e.g., PROV [4] and Sumatra[5]). While they are extremely valuable, many of these toolkits have a steep learning curve for scientists who may be expert in a specific domain but are not experienced enough in programming to, for example, write a wrapper for legacy code or translate an API.

2 Motivation for the work

The Tzar project[6] developed from a need to share versioned models among a group of ecological researchers and to reliably document the execution of those models – often over large numbers of runs with varying parameters. Typical challenges were: simulating agents' behaviour under hypothetical landscape change or regulation (e.g., Gordon et al., 2013); benchmarking systematic

[1] http://www.openmi.org/
[2] http://www.taverna.org.uk/
[3] https://kepler-project.org/
[4] http://www.w3.org/TR/prov-overview/
[5] http://neuralensemble.org/trac/sumatra
[6] https://tzar-framework.atlassian.net/wiki/display

conservation planning algorithms to assess their sensitivity to different types of uncertainty in species and habitat data (e.g., Langford et al., 2011); and investigating multiple scenarios such as variations in biodiversity offsetting policy (Gordon et al., 2011). The nature of the investigations, by definition, required numerous model runs with carefully documented variations in parameters, and it was important to be able to distribute the work over multiple computers and, ultimately, nodes in a research cloud. We found that while many researchers had significant expertise in using a specific programming language or set of algorithms, the challenges of sharing models and running them in a reproducible way were significant. In developing Tzar, therefore, we aimed to:

1. Lower the barrier to entry for different domain experts who might not necessarily have good programming skills;
2. Allow easy scaling from one to many machines (real or virtual); and
3. Increase reproducibility, with the capacity to branch and compare workflows and scenarios.

2 The Tzar application

2.1 Architecture and technology choices

Tzar is an open-source project[7], written in Java, which allows a computer to acquire and run models written in R, Python and Java, and potentially to act as a 'slave' which can be assigned 'jobs' involving those models. In order to run Tzar itself, the user needs only have a Java Virtual Machine capable of executing the compiled code, which is downloadable as a single jar file. However, since Tzar is not itself an execution environment, until recently it required local installation of the relevant Python / R libraries and other supporting dlls or executables on every machine running the computations. This required installation expertise and admin rights, and could interfere with users' everyday work by requiring them to install different versions of libraries and packages from those they already used. However, by using Docker containers (fully documented in section 2.4), a more straightforward and reproducible means of model sharing was achieved.

2.2 Sharing a model in Tzar

2.2.1 Configuration and versioning

Models shared using Tzar are versioned using standard approaches such as subversion or github. A model run in Tzar is flagged with the exact version of the code which was used, permitting replication. Input data may also be versioned in a repository, or it may be downloaded from a range of web sources – depending on the input format requirements of the specific model, these could be file servers, Sensor Observation Services[8] or OGC Web Feature[9] / Coverage[10] Services. Tzar does not provide a standardised data structure or interface to represent models: therefore each model must encapsulate resolvable references to its data sources and must contain the necessary code to open, transform or manipulate the downloaded data. While Tzar aims to enable model sharing, it is beyond the scope of the Tzar project to

[7] https://code.google.com/p/tzar-framework
[8] http://www.opengeospatial.org/standards/sos
[9] http://www.opengeospatial.org/standards/wfs
[10] http://www.opengeospatial.org/standards/wcs

ensure semantic interoperability between *multiple* models in terms of data types and methods. Therefore, if a Tzar model was to be included in an automated workflow, the model execution should be wrapped in a standardized interface, following a standard such as OpenMI.

To share a model in Tzar, the source code should be checked into an open repository, including two additional files (Figure 1): (a) *project.yaml* defining the model parameters and any ranges over which they should be sampled; (b) *model.R (or .java, or .py)* refers to those named parameters and calls the model code.

model.R

```
# Initialize variables
sd          =  variables$sd

# Run model
logPopTraj[1] =  log(1000)

for ( t  in  1:1000 )  {
  logPopTraj[t + 1] =
         logPopTraj[t] -0.1
         + rnorm(1, mean=0, sd=sd)
}
```

project.yaml

```
project_name: ecotas
runner_class: Rrunner

base_params:
    variables:
        sd: 0.1
repetitions:
       generators:
        - key: sd
          gen_type: linear_step
          start: 0.05
          step_size: 0.01
          count: 10
```

Figure 1 – example configuration for a simple stochastic population model. The variable 'sd' is parameterized in a yaml file, and then used by the model code.

To execute a set of runs locally (the simplest case) Tzar is then simply run from the command line, specifying the version of the model to use (Figure 2). More information about Tzar, and some more complex examples, may be seen on the project wiki [11].

URL address for code version control repository

```
java -jar tzar.jar execlocalruns http://rdv-framework.googlecode.com/svn/trunk/projects/ec3
--revision 394
```

Can specify a **required version** of the model

Figure 2 – running Tzar locally.

2.2.2 Uncertainty propagation example: global sensitivity analysis

Once a model has been configured for Tzar by adding these two files in the repository, it can be downloaded by any user and easily re-parameterised locally. For example, to vary across a range of parameter values requires only the addition of a 'repetitions' section to the yaml file which defines the parameters (see Figure 1).

[11] https://tzar-framework.atlassian.net/wiki/display

2.3 Distributing jobs with Tzar

2.3.1 Scheduling runs on multiple machines

To distribute work across multiple machines, runs must be scheduled into a central database. Tzar is then started on a set of 'nodes' (real or virtual) and each node will poll the database for jobs (Figure 3). Results may be written to a shared directory if this is feasible. This approach has been tested on physical machines in a lab network, and on the Australian Nectar research cloud.

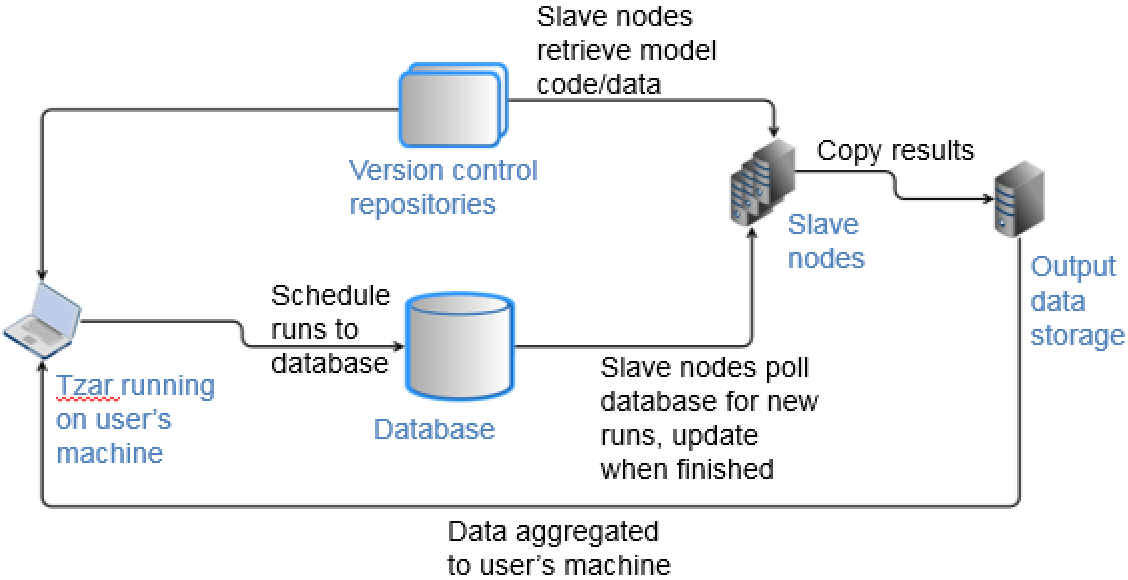


Figure 3 - distributing jobs with Tzar.

The necessity for a central database is another feature which historically required some user expertise: however, the use of Docker (section 2.4) permits the database to be encapsulated within the distribution and shared with very little effort from the user.

2.3.2 Managing jobs with a GUI

Scheduled jobs in Tzar may be monitored and managed with a Django-based GUI which permits jobs to be cancelled, re-run and edited. The source code for this GUI is also openly available on our repository [12].

2.4 Tzar plus Docker – an easier way to share computations

Docker[13] is an open platform to build, ship, and run applications. Docker enables applications and models to be quickly assembled from components and eliminates the friction between development and production environments. A sample Docker container might comprise just the model and its dependencies, making it completely portable. As a result, model developers can build any application in any language and through Docker Hub[14], they can also share Docker images with end-users through public or private repositories. Thus, any computer running Tzar which has Docker installed can access not just the versioned model and code, but the entire environment necessary to run that code. In this way, reproducibility in terms of algorithms, libraries and package versions is guaranteed without the user needing to install or update libraries on the machines which are picking up assigned jobs.

[12] http://tzar-framework.googlecode.com/svn/trunk/
[13] https://www.docker.com
[14] https://registry.hub.docker.com/

The database needed by Tzar to distribute jobs among different computers can also be ‘Dockerized’ to ease the process of setup and configuration. A Docker image composed of a Postgres 9.3 server with scripts to create a Tzar database is available[15] online.

3 Case study – eHabitat+

In order to demonstrate the value of the Tzar/Docker combination, we chose to share and distribute ‘eHabitat+’ a complex model which (a) uses a variety of very specific Python and GRASS libraries, complicated for a user to replicate, and which (b) is embarrassingly parallel, in that it must be run many times on a large number of independent datasets (namely, separate ‘ecoregions’ and sets of geographic areas which are protected for conservation purposes). eHabitat+ is an enhanced habitat mapping tool for protected areas, used in the Digital Observatory for Protected Areas (DOPA) (Dubois et al. 2013). Based on several environmental variables derived from remote sensing, eHabitat+ first performs a landscape segmentation step inside each protected area to identify the main ecological features and habitats of each reserve to be identified. Secondly, each segment (i.e. functional habitat type) is individually processed in order to generate similarity maps at a larger spatial extent. In other words, eHabitat+ computes, for each habitat type inside a protected area, a map of the probability of finding similar biophysical conditions and ecological characteristics elsewhere within the ecoregion (Olson et al. 2001). Finally, based on these similarity maps, several landscape metrics are computed for each functional habitat type by identifying and measuring all contiguous areas that present similar characteristics.

The eHabitat+ software tool is open source (Martínez-López, 2014) and it is based on several numerical and scientific Python libraries for multivariate analysis, such as NumPy, SciPy, Multiprocessing, Scikit-learn, as well as on GRASS GIS 7 for the automatic image segmentation step (Neteler et al., 2012). The segmentation algorithm requires two main parameters, minimum patch size and similarity threshold, that were estimated empirically.

In order to allow eHabitat+ to be executed using the Tzar distributed computing framework, an Ubuntu 14.04 Docker image[16] was set up with all the necessary software and library dependencies. The two extra configuration files required by Tzar (see section 2.2.1), were also added to the eHabitat+ repository on GitHub[17], together with a brief description of the steps needed to fetch the Docker image and start a container, as well as the processing[18].

---

15 https://registry.hub.docker.com/u/javimarlop/postgres93-tzar-docker/
16 https://registry.hub.docker.com/u/javimarlop/ubuntugis-docker/
17 https://github.com/javimarlop/eHabpy
18 https://github.com/javimarlop/eHabpy/blob/master/ehab_tzar_docker.sh

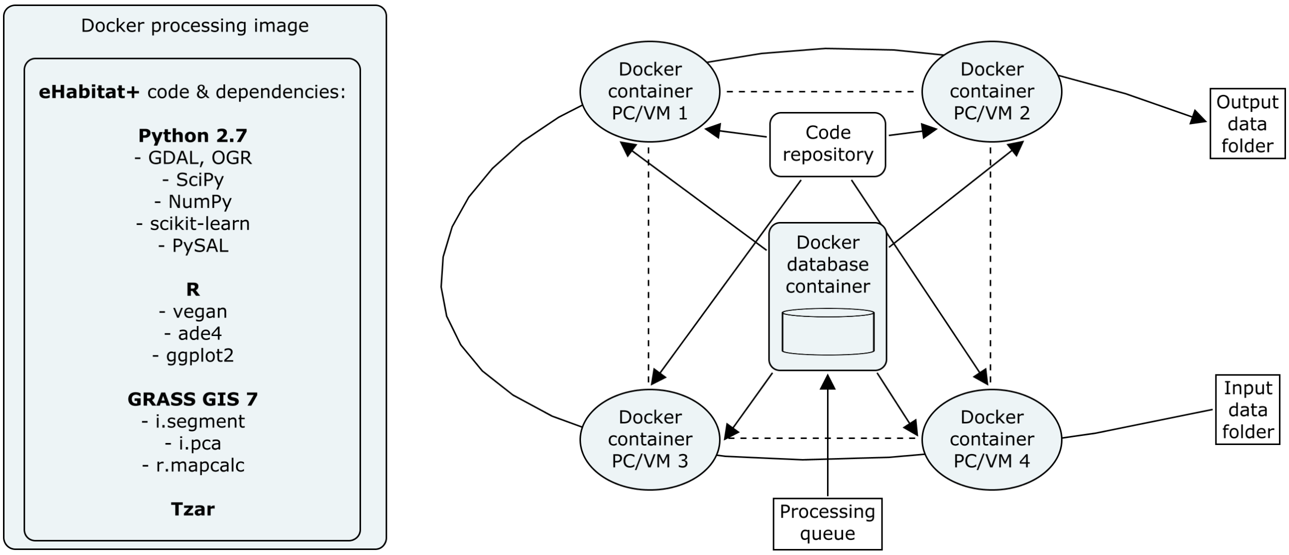


Figure 4 – running Tzar and Docker for distributed computing.

This simple procedure allowed us to run eHabitat+ over a large number of protected areas by distributing the processing among several computers, without having to set up all the necessary environment and dependencies for each node. The conceptual diagram of the procedure can be seen in Figure 4. The user uploads a processing queue into a Dockerized database (shown as a cylinder on the diagram) and the jobs are picked up and executed by Docker running containers installed on each node (PC or virtual machine; circular boxes). Each node contains the environment required to run eHabitat+ (components listed in the box on the left). Before executing jobs, the node fetches the eHabitat+ code from the repository. When a job is executed, the node processes several features using the same input data and stores the output in a common location.

4 Discussion and Conclusions

In the work described we have successfully managed to ease the re-use and sharing of a potentially complex model - in this case, the eHabitat+ model for segmentation of functional habitat types and mapping of similar areas in the landscape. The model uses a variety of languages and libraries, and requires repeated executions across a range of ecoregions and protected areas. Both these characteristics make it a potentially difficult analysis to repeat without numerous installation, tuning, configuration and aggregation steps. However, using Docker, the code and data have been successfully bundled into an open repository, which effectively contains the execution environment as well as the database required to orchestrate repeated runs of the model, and using Tzar, those runs can be smoothly distributed among a set of slave nodes. Most importantly, the provenance of the results is recorded in terms of the input data version, the exact nature of the model code and the specifics of the execution environment. We believe this is a useful step in making reproducible computing available to a wide variety of scientists.

To further develop the usefulness of the Tzar project, we will next explore whether provenance can be documented in more detail using models such as PROV-O, and will investigate the possibility of integration with standards such as the OGC Web Processing Service, through a mapping between the Tzar configuration files and the WPS process description.

References

- ✔ Dubois, G., Schulz, M., Skøien, J., Cottam, A., Temperley, W., Clerici, M., Drakou, E., et al., (2013). *An introduction to the Digital Observatory for Protected Areas (DOPA) and the DOPA Explorer (Beta).* EUR 26207 EN, EC. Luxembourg: Publications Office of the European Union.
- ✔ Gordon, A., Langford, W.T., Todd, J.A., White, M.D., Mullerworth, D.W. & Bekessy, S.A. (2011). Assessing the impacts of biodiversity offset policies. *Environmental Modelling and Software*, *144*, 558-566.
- ✔ Gordon, A**.**, Bastin, L., Langford W.T., Lechner, A.M. & Bekessy, S.A. (2013). Simulating the value of collaboration in multi-actor conservation planning. *Ecological Modelling*. *249*, 19-25. doi: 10.1016/j.ecolmodel.2012.07.009.
- ✔ Langford W.T., Gordon A., Bastin L., Bekessy S.A., White M. & Newell G. (2011) Raising the bar for systematic conservation planning. Trends in Ecology and Evolution, 26, 634-640.
- ✔ Martínez-López J. (2014). eHabitat Python and GRASS GIS library. doi:10.5281/zenodo.10612.
- ✔ Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, *31*, 124-130.
- ✔ Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C. & Kassem, K.R. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51(11),* 933 – 938.

GIS of Traditional Folk Culture

Jiří Kozel[1], Pavel Bohumel[2], Petr Kovács[3], Andrea Kýnová[4]

[1,2,3,4] Masaryk University, Institute of Computer Science

Abstract

This paper describes the process of the construction of the GIS of traditional Czech folk culture between 1750 – 1900. Our project is the first that collects records about all the aspects of folk culture in the Czech Republic systematically. We started with the design of a data model well-suited for heterogeneous folk-culture metadata. At present, we have about 35,000 records in the database, an editing application, and a web map presentation with spatiotemporal filtering written in Google Closure and OpenLayers 3.

Keywords

GIS, folk, culture, data model, web map, filtering

1 Introduction

Traditional Czech folk culture has a long and rich history, especially in the region of Moravia, which covers one third of the Czech Republic. Folk culture was an important vehicle of patriotism during the Austro-Hungarian domination, which took place for four centuries. Many of those customs, songs, crafts, and arts are still alive among ordinary people. Four items from the UNESCO Intangible Cultural Heritage List (314 items in total) are located in the Czech Republic, including the recruit dance ”verbuňk“ and the ”Ride of the Kings“ located in Moravia (UNESCO, 2015).

1.1 State of the Art, and Objectives

Before our project started, there had been no systematic collection of all the aspects of the traditional folk culture in the Czech Republic. There exist some web map presentations related only to a single topic of folk culture, e.g. “Hrady a zámky” (http://www.hrady.cz/), related to architectural objects, or the Folk Costumes Map currently being developed at the Czech Institute of Folk Culture.

Our goal is to create a systematic collection of metadata records of all the aspects of traditional folk culture in the region of Moravia between 1750 – 1900, and to build a web GIS that would enable editing, presentation, and advanced filtering of the collected data. This unique collection and developed applications enable subsequent research of a given topic, and help to preserve our cultural heritage.

Nowadays the database contains approximately 35,000 records collected by ethnological experts. Each record has attributes that describe place, time, category, and source. Furthermore, we have an editing application and a web map presentation with spatiotemporal filtering written in Google Closure and OL3. This paper focuses especially on the geoinformational aspects of the building of this collection and GIS.

1.2 Who Stands Behind

This paper forms a part of the project “Geographical information system for traditional folk culture (1750–1900)”, in short ”GISTraLiK“. The project is supported by a grant from the Ministry of Culture of the Czech Republic as project No. DF12P01OVV015. Two cooperating partners are involved: the Department of European Ethnology and the Institute of Computer Science, both belonging to Masaryk University in Brno.

2 Data model

The first and the most crucial step was to create a data model well-suited for heterogeneous metadata concerning folk culture. It took hours of discussion among ethnological and GIS experts. We had to deal with a certain level of uncertainty (especially in the case of temporal designation), different levels of geographic localities ranging from municipalities to ethnological regions, and the extremely wide scope of folk culture including customs, products, songs, artworks, etc. Basic concepts of the data model are illustrated in Figure 1.

2.1 Record

The record is the main class of the model. It represents either a mention of a custom, craft, song, social relations, event, etc., or an artwork (e.g. painting, ceramics, or literature).

Each record has several important mandatory properties including keywords, spatial and temporal validity, and the source of the represented mention or artwork. The multiplicity of keywords is 1–n, the multiplicity of other attributes is 1–1. These attributes allow us advanced filtering and record searching in the presentational part of the GIS.

The textual description is an attribute with a free domain, whereas keywords are selected from controlled vocabulary (see the next subsection).

Temporal validity is represented by year range (i.e. integer range) with an optional text. This allows us to store and search for even approximate temporal information, such as “beginning of the 19th century”, “mid 1830s”, or “2nd half of the 19th century”.

Spatial validity is represented by one one locality (see below for details).

A source is either a document that contains a mention about folk culture, or an object that contains an artwork (see below for details).

2.2 Locality

A locality is an abstract class representing a spatial unit; there are five subclasses: a municipality, parish, manor, legal district, and an ethnological region. The first four subclasses are administrative or territorial units that were legal for at least a part of the period 1750–1900. Ethnological regions are used by ethnologists to demarcate areas with similar folk culture.

The spatial extent of each locality is represented by one polygon or multipolygon. It means that we are not able to track the changes of boundaries in time, we just use the widest known delimitation of a given locality in the given period.

The overlapping of localities of the same type is allowed. There are three reasons for that: the changes of boundaries in time, it is not always possible to delimit a historical boundary exactly, and ethnological regions simply overlap even in reality.

2.3 Keywords

Keywords represent the ethnological classification of records. Keywords form a hierarchical acyclic structure. The highest level contains the most general keywords, whereas the lower levels contain more specific keywords, e.g. Music -> Musical instrument -> Bagpipes. One keyword can appear as a subkeyword of more keywords, e.g. Food -> Egg and Custom -> Easter -> Egg. The connection between a record and a keyword carries information on whether the keyword (e.g. Egg) is used in a general context or whether it relates just to a specific branch of the structure (i.e. to a specific meaning).

2.4 Source

A source is either a document (e.g. a chronicle, book, article, or a map) or an object that contains an artwork (e.g. a museum exhibit). Every source carries complete citation information in order to be identified easily (e.g. authors, year of publication, publisher, magazine, museum, exhibit No., etc.).

An attachment (usually a photo or a scan) and a URL that leads to the web page with the electronic version of a given source are very important optional attributes of a source. These attributes enable users to study a record directly without a physical visit to an archive, library, or a museum.

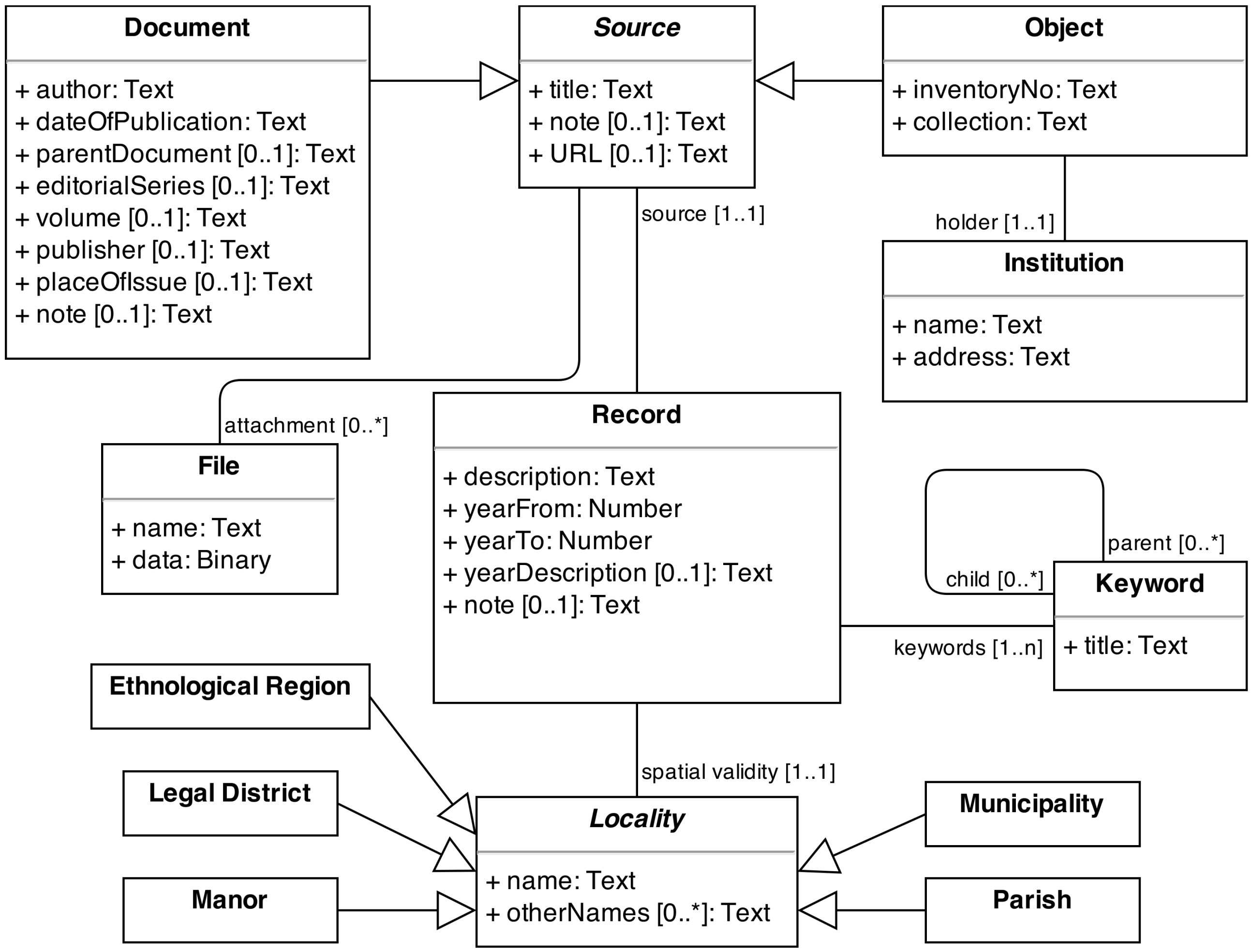


Figure 1 - Basic Concepts of the Data Model

3 Data acquisition

Before the records collection started, a complete index of keywords and complete register of localities had been created.

3.1 Keyword Index

The compilation of the keyword index was done specially by ethnologists because of their expert knowledge of the traditional folk culture phenomena. Attention was paid especially to covering all the phenomena equally to avoid a later expansion of the index. This complex task took about a year to complete. Finally, the index contained 1260 keywords separated into 3 hierarchical levels; the first level contains 28 keywords (Drápala, Doušek, Křížová, Pavlicová, and Válka, 2013). The whole index was also translated from Czech into English.

3.2 Localities Register

The register of localities, with the exception of ethnological regions, was made according to historical documents, such as old cadasters (http://archivnimapy.cuzk.cz/) or maps of Moravian regions by Schenkl, 1841–1845. The numbers of localities by type are available in Table 1. A list of 15 ethnological regions was prepared by ethnologists.
The register initially contained only names of localities and their identifiers.

Subsequently, we also filled the geometry attribute specifying the locality area using a polygon or multipolygon. The geometry was derived especially from the vector database of old cadasters made by CUZK.

3.3 Records and Sources

First, ethnologists made a list of about 200 museums, archives, and collections which cover not only all aspects of traditional Moravian folk culture, but also the whole geographic area of Moravia and the whole temporal period (Horáková & Šipöczová, 2013). Similarly, a list of scientific publications and magazines related to the field of interest was made.
Then ethnologists started to collect sources and records. The latter are related to the source, existing keywords, and an existing locality. We had a three-month testing phase to evaluate the data model, and after the evaluation we published the appropriate methodology (Drápala & Malecká, 2013). Today there are about 35,000 records in the database processed by ethnologists according to the methodology.

4 Database and Web Applications

4.1 Database

Prior to data acquisition, a data model had been implemented in the database, and an input application of acquired data had been built. Because of contractual reasons, we used Microsoft SQL Server with "geometry" spatial type as the database engine. Two databases are used: the "source database", which contains the source data that are managed by the editing application, and the "publication database" with a flattened model used exclusively for the web presentation. The flattened model enables providing and querying the data faster. The publication database is automatically updated once per 24 hours from the source database.

4.2 Web Applications

Since the input application is not public and it is a text-form application without maps, it was omitted from this description.
Parallel to the records and sources acquisition, we started to build a web presentation application with an interactive map and records filtering. It is located at http://gistralik.muni.cz/. We used OpenLayers 3 (http://openlayers.org/) and Google Closure libraries (https://developers.google.com/closure/library/) for client-side interactivity. The server side that serves data from the database to the client is managed only by ArcGIS REST API (http://resources.arcgis.com/en/help/arcgis-rest-api/), used because of contractual reasons. The technical aspects of the client-side solution are described in section 4.3.
The client side has two separate pages: search in the database (vyhledavani-db.html) and search in the map (vyhledavani-mapa.html). Both pages enable to search records by advanced filtering; the former enables to browse found records whereas the latter visualizes the number of found records per locality in a simple choropleth map.
The filtering of records is a very important part of the two pages. Filtering according to locality, time range, keywords, source, and a full-text search is

enabled. The current state of the filter control immediately updates the list of found results or the numbers of found results in the map. Furthermore, the state of the filter is stored in the URL so that everyone can copy the link and go back to the current state anytime later.
The web presentation exists in Czech and will be translated into English by June 2015.

4.3 OpenLayers 3 in Web Presentation Application

The ArcGIS REST API is used only for providing vector and attribute data in the JSON format that is similar to GeoJSON. We created OpenLayers 3 class ol.format.EsriJson as a subtype of ol.format.Feature. This class is able to parse ESRI JSON and create ol.Feature instances including geometries (e.g. point, polygon, or multipolygon). We use ol.source.ServerVector class for storing features on client side.
The locality layer is represented by ol.layer.Vector with custom styles (ol.style.Style) and the whole map is rendered to HTML canvas. Even with massive caching of styles and preprocessing (geometric generalization), we reached the limits of the browser reaction time, because too many polygons and vertices are visualized (see following table). Whereas the zoom of the legal districts layer is usually smooth in current browsers, the zoom of the municipalities layer is not.

Locality Type	**Number of Features**	**Number of Vertices**
Legal districts	77	14,040
Manors	415	44,960
Parishes	873	54,487
Municipalities	3,163	176,557

Table 1 - Numbers of Features and Vertices per Feature Type.

Two further principal extensions of the OL3 library were made. First we created ol.object.Type and ol.feature.Type, which are a simplification of FeatureType of ISO 191xx series (ISO, 2001). This helped us to load and store features and their properties on the client side automatically. Then we extended ol.Object so that every property may have a validator. The validator is a function called anytime an attempt of setting a property value is made. If the validator fails, the new value is not set. This helps us to build robust model classes with business logic separated from the view components.

5 Conclusions

This paper presented the building of the GIS of traditional Czech folk culture in the region of Moravia between 1750 – 1900. The project is realized by GIS and ethnological sections. Our field of study – folk culture – is not a field of study where GIS would be used commonly, so attention was paid especially to the data model and to the ability of web presentation to find and visualize the intensity of folk-culture records according to the filter used by the user easily. The created GIS enables subsequent research of Moravian folk culture and helps to preserve our cultural heritage. The project is going to continue at least until the end of 2015.

References

- ✔ Drápala, D., Doušek, R., Křížová, A., Pavlicová, M., & Válka M. (2013). Glossary of Geographic Information System of Traditional Folk Culture (1750-1900). Brno: muniPRESS. In Czech.
- ✔ Drápala, D., & Malecká, J. (2013). Methodology of Passportisation Resources and Processing of Records of Geographic Information System of Traditional Folk Culture (1750-1900). Brno: muniPRESS. In Czech.
- ✔ Horáková, A. & Šipöczová, E. (2013). Geographic information dream. *Prace Etnograficzne, Tom 41 (2013), Number 4,* 287–295.
- ✔ ISO. (2001). Draft Technical Specification 19103, Geographic information - Conceptual schema language, ISO/TC 211 N 1082.
- ✔ Schenkl, K. (1841-1845). Maps of Moravia and Silesia regions.
- ✔ UNESCO. (2015). Intanglible Cultural Heritage. Retrieved March 6, 2015, from http://www.unesco.org/culture/ich/en/home.

Towards a decision support system for environmental emergencies management in poor settlements in the Kathmandu Valley (Nepal)

Marta Giambelli, Alfonso Vitti, Marco Bezzi

DICAM, University of Trento.

Abstract

This study is the first output of the active collaboration among ASIA Onlus* and the University of Trento (Civil, Environmental and Mechanical Engineering Department) in the analysis of environmental risk occurrences in the Kathmandu Valley. Aim of the project is the implementation of a practical decision support system tool for environmental emergencies management for the poor settlements in Kathmandu. An intense field work and analysis has been done for the collection of input data for the model. The population of informal urbanized areas has been involved with a participatory approach to discover perception and behaviors during environmental risk occurrences and to share knowledge and procedures in risk management. The study focuses on one poor-area in the Kathmandu Valley considering flooding, earthquake and fire as environmental risks and aims to be extended to all the city with the direct involvement of local authorities.

* ASIA Onlus: Via San Martino della Battaglia, 31, 00185 Roma.

Keywords

Decision support system, environmental emergencies management, Kathmandu Valley, poor settlements, network analysis, FOSS4G.

1 Introduction

In many parts of the world and mostly in developing countries, processes such as badly planned and managed urban development, environmental degradation, poverty and inequality along with weak governance, are driving levels of disaster risk to new heights (United Nations International Strategy for Disaster Reduction [UNISDR], 2010). Since 2005, countries have been addressing this challenge through the Hyogo Framework for Action (HFA), which aims to achieve a substantial reduction of disaster losses, in lives and in the social, economic and environmental assets of countries and communities by 2015 (UNISDR, 2014). Moreover, the Bangkok Declaration on Disaster Risk Reduction (DRR) in Asia and the Pacific (2014) included the objectives of enhancing resilience at local levels, promoting further use of science and technology and innovation in DRR, and building coherence between HFA and current processes on the sustainable development goals and climate change arrangements. Nepal is in the early stages of disaster management planning and in order to address the HFA, the project "Mainstreaming Disaster Risk Reduction in Megacities" (Earthquakes and Megacities Initiatives [EMI], 2010) has been started. It defined a Structure and Implementation of a Competent Disaster Risk Management Function for Kathmandu Metropolitan City (KMC). A new unit called the Disaster Management Section was created, structured for preparing and responding to all types of emergencies, from major disasters to the “every-day” risks (EMI, 2010).

This study arises from an agreement between the University of Trento and the NGO ASIA Onlus. A Migrant Mapping Initiative was launched in collaboration with the NGO INCLUDED, testing the methodology in Balkhu settlement, in Kathmandu. The fieldwork allowed the collection of data and the possibility for understanding and being familiar with the slum's and city's contexts. The work described in this paper has the objective of creating the bases of a decision support system for environmental emergencies management in poor settlements in the Kathmandu Valley. The Valley is in fact prone to multiple hazards, such as fire and flooding and, all along the country, there is the earthquake's risk. Embracing the objective of enhancing resilience at local level, special focus is given to community's development building, in particular in migrant poor settlements, which are not recognized by the government and in which vulnerable groups live. The aim of the decision support system is to make rescuers prepared and ready to act in the case of environmental emergencies, giving them support in the paths' choice to reach the emergency place, through the network analysis. On the other hand, guide lines of behavior are provided to slum's dwellers to evacuate the settlement and reach an appropriate assembly area. For this purpose, the tool is implemented in GIS (Geographic Information System), addressing the suggested action of using technology and innovation in DRR, although it should be accessible, available and affordable to national governments and local communities. Therefore, the use of an open source software is inescapably required. A case study of Balkhu then provides an application of the analysis and emphasizes the efforts that should be done, in order to understand social and demographic features of the settlement, community perceptions on environmental vulnerabilities and to spatially connect these information.

2 Study area: Nepal and the Kathmandu Valley

Nepal is a small country landlocked between China and India. Its population numbered approximately 26.5 million, with almost 80% of people living in rural areas. The Kathmandu Valley is the political, commercial and cultural hub of Nepal, and it is located in the Central Development Region. Spread across an area of 665 sq km and at an altitude of 1336 mslmm, the Valley encloses parts of three districts, Kathmandu, Lalitpur and Bhaktapur (Fig.1), which had the highest population densities in 2011, reaching the maximum in KMC with a value of 4408 people/sq km.

Nepal's development challenges need to be appreciated against the political backdrop of the last two decades. In 1995 the Maoist rebels started an armed conflict, which lasted for ten years. In 2008 the monarchy was abolished and the country became a Federal Democratic Republic. Nevertheless, currently major political parties are still focused on finding common ground for the creation of an inclusive Constitution. Moreover, Nepal is classified as a least developed country with a large fraction of population under poverty level. The extent of poverty in Nepal has dropped to 25% in 2010/2011, but this is an overall figure of progress and not all groups have shared equally in poverty reduction (UN Country team in Nepal, 2013). Social development indicators in fact reveal unequal progress in gender, ethnic, religious and caste groups, and geographical regions. Agriculture is the mainstay of the economy, providing a livelihood for more than 70% of the population. However, a heavy reliance on tourism and farming makes Nepal's economy very sensitive to climate variability.

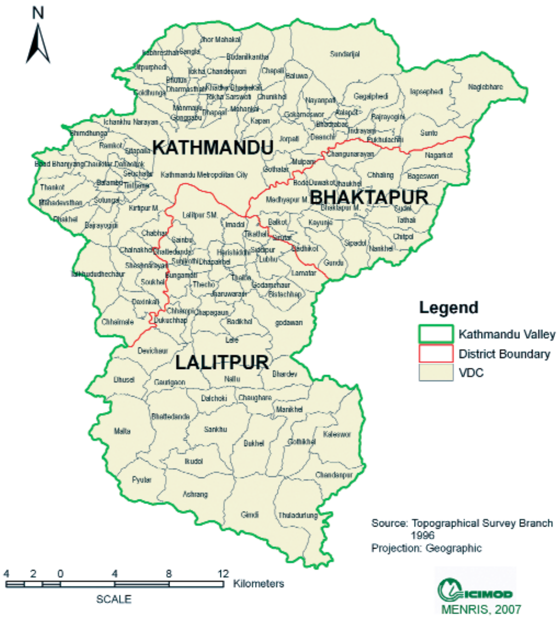


Figure 1 - Map of Kathmandu Valley, consisting of three districts: Kathmandu, Lalitpur and Bhaktapur (ICIMOD, 2007).

2.1 Urban vulnerabilities

Haphazard and unplanned urban development often lead to an increased variety of vulnerabilities. In the Kathmandu Valley urban areas are characterized by the lack of appropriate infrastructure such as water supply and sewerage systems (Toffin, 2010). This steady urbanization lead to negative impacts, such as traffic congestion, atmospheric pollution and a total collapse of the former fragile ecological equilibrium between man and his environment (Toffin, 2010). The rivers have also undergone tremendous pressure from the increase in demographic growth and in economic activities. This situation converges to numerous urban vulnerabilities, such as (1) a lack of emergency vehicular access and vulnerable infrastructures; (2) a lack of open spaces; (3) uncontrolled solid waste management, like open dumping practice and (4) unhygienic conditions (Bhattarai & Dennis, 2010). Moreover, sub-standards houses, which are often characterized by exorbitant price, has increased property and land prices, making most housing unaffordable to the common Nepalese and leaving them with no other alternative but to squat on open lands.

2.1.1 Slums and squatter settlements

The local NGO Lumanti (2008) defines slum communities as areas of "poverty, low income, inadequate living conditions and sub-standard facilities". Another important feature is their location on marginalized land: the majority of these settlements in the Valley is established on the floodplain of rivers. Most of the remaining settlements are in areas prone to landslides, while another place for occupancy is ghats, where traditional public buildings are on the verge to collapse (UN Country Team in Nepal, 2013). According to a study of UNSD (United Nation Statistical Division) of 2007, it is estimated that 92% of Nepal urban population live in slums. Although this figure is implausibly high, slum settlements are expanding in size and number along with the expansion of urban areas, reaching the number of 40 settlements and 2735 households (Lumanti NGO, 2008). As the political crisis that Nepal has undergoing, any attempt at addressing to these ongoing problems is passed on to NGOs and even housing rights activists are fighting to inscribe housing rights in the future constitution (Toffin, 2010).

2.1.2 Flooding, fire and earthquake

The major rivers flowing through the Kathmandu Valley are Bagmati, Bishnumati and Manohara. The occasional torrential rains, mostly during monsoon period, have caused flooding problems in the core areas of the city, inducing loss of life and damage to private properties, especially in the floodplain.

Fire is another common hazard in the Valley and the most vulnerable area is the central core, where houses are old and made with timbers and closely attached to each other and often characterized by a mixed use, which can worsen the risk. In the recent past, urban house fire has threatened people's lives due to the poor infrastructure and virtually non-existent fire and rescue services, making the vulnerability of the people of Kathmandu exacerbated by disasters, their socio-economic and ecosystem status (Raj Aryal, 2011).

Earthquakes are potentially the greatest threat, as most of Nepal lies in a high-risk seismic zone. According to seismic records, the Country faces one earthquake of magnitude 7 or greater every 75 years, on average (EMI, 2010). In this case, wealthy people are not necessarily better positioned than the poor. A risk assessment study illustrated the implications of a Mid-Nepal Earthquake scenario on Kathmandu Valley, predicting 21% of heavily damaged buildings, 1.3% of the total population of the Valley dead and finally 3.8% of the total population in the Valley seriously injured (Japan International Cooperation Agency [JICA], Ministry of Home Affairs of Nepal [MOHA], 2002). Moreover, according to Bhattarai et al. (2010), schools, hospitals and public offices are not earthquake resistant. Currently the Government is working on draft Disaster Risk Management Act, however it is not paying sufficient attention to preparedness (Dixit, 2015).

3 The project “Migrant Mapping Initiative”

The general objective of the project “Migrant Mapping Initiative" was to develop a mapping tool to understand the context of migrant urban settlements across different levels of poverty. The methodology implemented followed four steps:

1. preliminary phase, to make stakeholders aware of the potentiality of GIS and to define objectives, database creation and existing data needed;
2. identification of the community;
3. mapping the settlement and collecting information regarding houses' features;
4. data collection.

Regarding the identification of the community, it was a challenging phase, as not all the communities visited were willing to participate in a research that would not benefit them directly and/or immediately. The mapping fieldwork has been managed printing a zoomed satellite image and drawing polygons of households and line of paths directly on it, as the GPS was not effective because of the overcrowded area. These elements were then digitalized through the QGIS software, which provides an OpenLayers plugin, allowing the addition of a satellite image, used as base map. Furthermore, data collection was addressed with particular attention to the community participation within the project. The process included a Focus Group Discussion and questionnaires, held by both community youths and other volunteers. In this phase a training to orient each enumerator with the survey was organized. The database creation should identify spatial data to treat the problem, thus even data validation and coherence are unavoidable.

Some typical challenges experienced during an evaluation of the use of space-based information for disaster management in several countries are (1) non availability of baseline and thematic spatial data in the framework of GIS; (2) baseline and thematic spatial data available, but scattered at different locations with non-uniform data standards that limit data integration and finally (3) baseline and thematic spatial data well organized centrally, but not integrated with risk related data (Ravan, 2010). In particular, most of the Kathmandu Valley's data collected are from KVTDC (Kathmandu Valley Town Development Committee) and OSM (OpenStreetMap). The layers used in the analysis are reported in Tab.1 along with the related source. However, disaster specific data are needed, especially in the preparedness phase, such as hazards map related to the three environmental risks considered.

Maps	**Source**
Administrative boundaries	KVTDC
Roads	OSM
Hospitals	KVTDC
Fire brigades	OSM
Army barracks	KVTDC
Balkhu maps	Fieldwork

Table 1 - GIS maps used in the analysis and sources.

4 Results from “Migrant Mapping Initiative”

4.1 Map and location of Balkhu settlement

Balkhu settlement is located in the 14th ward of Kathmandu District, along the Bagmati river, in the south-western part of the city. The community came into existence about 9 years ago and there are roughly 350 households living in the area and approximately 1400 people. In the municipality map, the area appears as “Shahid Smarak park”, thus it is not recognized by Government and the community is not entitled to basic services. The settlement is divided in 4 sectors and it is characterized by poor housing conditions and materials vulnerable to fire. Furthermore, the internal paths are narrow and labyrinthine, as it is possible to observe in Fig.2, making it difficult to rescue people and save their belongings, during a disaster situation. The area is exposed to different vulnerabilities due to its location, as depicted in Fig.2: on the west side of the community, there is a very busy road, crossed by big trucks, as the proximity of a fruit market in the north; the soil where houses are built on is packed plastic waste and they have problems with sinking soil. With the fruit market on one side, and waste dump in the north side of the community, the area has many waste hazards due to the garbage dump and problems with foul smell, with potential airborne diseases related. There is also industrial waste and toxic fumes from a factory that produces utensils, located in front of the settlement.

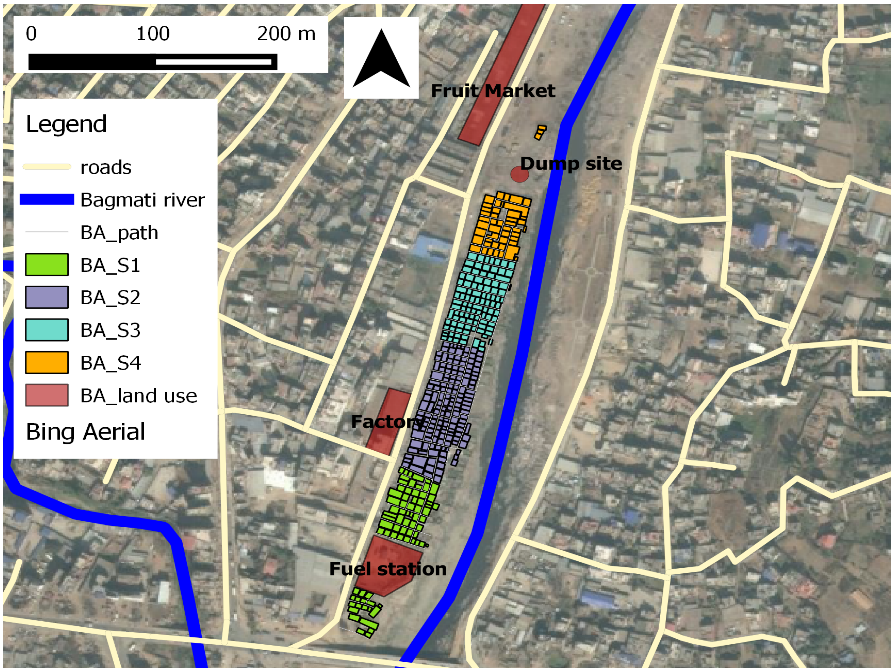


Figure 2 - Map of Balkhu settlement and land use points of interest.

4.2 Focus Group Discussion and questionnaires

One section of the Focus Group Discussion was dedicated to environmental hazards, to understand the vulnerability of Balkhu community and its perception around the issue. In particular, between fire, flooding, earthquake and bank erosion, the community stressed flooding and fire as major concerns. They pointed that homes located in the river proximity have been under direct threat of flooding from the river, as themselves drawn in a map created for the occasion. However, every year during monsoon and heavy rainfall, community experienced material loss and damage, but no human casualties have occurred thus far. One leader estimated about 50 houses get affected every year by flood water entering in houses and, at worse, the water comes up to about 0.6-0.9 m. Furthermore, they stressed that there are no organizations supporting them for design of emergency plans, hence they did not know where to go, or where to meet. The community’s way to respond to these risks is for the youth group to be alert and warn the community when water level rises, but there are no formal precaution measures. From questionnaires, demographic and sociological data were collected, useful in order to consider an eventual relocation, as solution to the vulnerable area and status of illegality of the settlement.

5 Network Analysis

The Network Analysis is a key part of this decision support system for the emergency preparedness and response in the Kathmandu Valley. The analysis is here applied into two environments, involving different stakeholders: the first one is related to the arrival of rescuers in the emergency place, while the second one is relevant to the evacuation of people from the settlement to reach safe areas. The Network Analysis has been applied in the software open source GRASS GIS, which is currently used in academic settings. It has powerful capabilities and can be integrated into QGIS. The latter is one of the most used software in International Cooperation' applications, as it is highly flexible and easy to use. In this manner the academic world would be linked to a more user friendly world, making results more affordable to the Municipality of Kathmandu or to other final users. Anyway, in both the cases, a training on these two GIS software would be necessary.
The procedure followed for the Network Analysis is the one suggested by the GRASS Tutorial developed by the University of Trento (2010). After setting up the data and creating the network (*v.net*) the analysis' process followed for each rescue service is composed by three steps:

1. allocation across the network to a competence area based on distance, in order to evaluate the coverage of services in the Valley (*v.alloc*);
2. allocation based on transit time, in order to identify the rescue service responsible for Balkhu area (*v.alloc*);
3. search for the shortest and fastest paths to reach the settlement (*v.net.path*).

The time considered in the analysis is referred only to the transit phase, while even call and preparedness times will be added. The road network used for the study covers the Kathmandu Metropolitan City and the Lalitpur Sub-Metropolitan City (LSMC), connected to the points which represent the rescue services, while the destination is the studied slum of Balkhu. In the second application, the network consists of the settlement's internal paths, connected to the urban road network and then to specific points, represented the safe areas.

5.1 Input data

<u>5.1.1 Open spaces as assembly areas</u>

One important concern of strategies to improve preparedness for response is the identification of suitable assembly areas. In general, for the safety of an individual, a 2 sqm of open space is needed (Bhattarai et al., 2010). This value could be effective: a standing man occupies 0.27 sqm, but in an emergency situation it should be enlarged. Supposing the presence of tired or injured

people, we may actually hypothesize an occupation space of 2 sqm.
The evacuation problem is a complex issue, as there are heterogeneous aspects to deal with, like the psychological component of people and their behavior in a stress situation. Open spaces should be chosen taking into account the proximity to the evacuated settlement and they should be easily and safely reachable by both dwellers and rescuers, without hampering the access to the settlement. Furthermore, the choice of the open place depends on the kind of environmental emergency, thus for each of them different reasoning should be done. In order to make the process of evacuation easier for the slum's dwellers, it has been structured by sectors. These areas were identified using a satellite image and through the knowledge acquired thanks to the fieldwork (Fig.3). However, the current physical status of these areas has to be improved, in order to respond efficiently to an emergency situation.

5.1.2 Nepal road status and classification

The main road network inside the Valley consists of corridors, one from north to south and one from east to west, along with a Ring Road surrounding the cities of Kathmandu and Patan (Lalitpur). Moreover, there are several radial roads, some originating from the city core area and others from the Ring Road. Apart from these, there are urban roads, most of which are narrow and heavy built-up on both sides, fostering a condition of almost every day jammed roads, which affects the travel time of emergency vehicles and increases response time of emergency services.
The OSM layer has been chosen for the analysis, but the maximum speed of roads is missing and the information regarding the access way is useless. Therefore, in this first phase, all the roads are considered accessible from both the ways and from the type, maximum speed was assumed. Because of no traffic data are available, the urban speed of 40 km/h is supposed to be the maximum one in primary links, even if road rules are not binding for rescuers' vehicles, in an emergency situation.

5.1.3 Rescue services

The rescue services included in the analysis are hospitals, fire brigade stations and armed police barracks. According to JICA and MOHA (2002), in the Valley there are 47 hospitals and only three fire stations, located in Kathmandu, Lalitpur and Bhaktapur Municipality. The third rescue service considered is the Nepal armed police, which has been actively participating in disaster management and leading some emergency operations. It is involved in the collection of information, resource identification, emergency response task assignment, communication with the public, coordination with other security forces, first aid and medical assistance and damage assessment (United Nations Development Programme [UNDP], 2012).

5.2 Application

5.2.1 Evacuation plan for Balkhu settlement

The evacuation plan for Balkhu settlement, depicted in Fig.3, has been organized per sectors: this choice makes the escape easier and faster, as people know the sectors' division of the settlement and where friends and relatives live. In this manner, the psychological aspect would be mitigated.

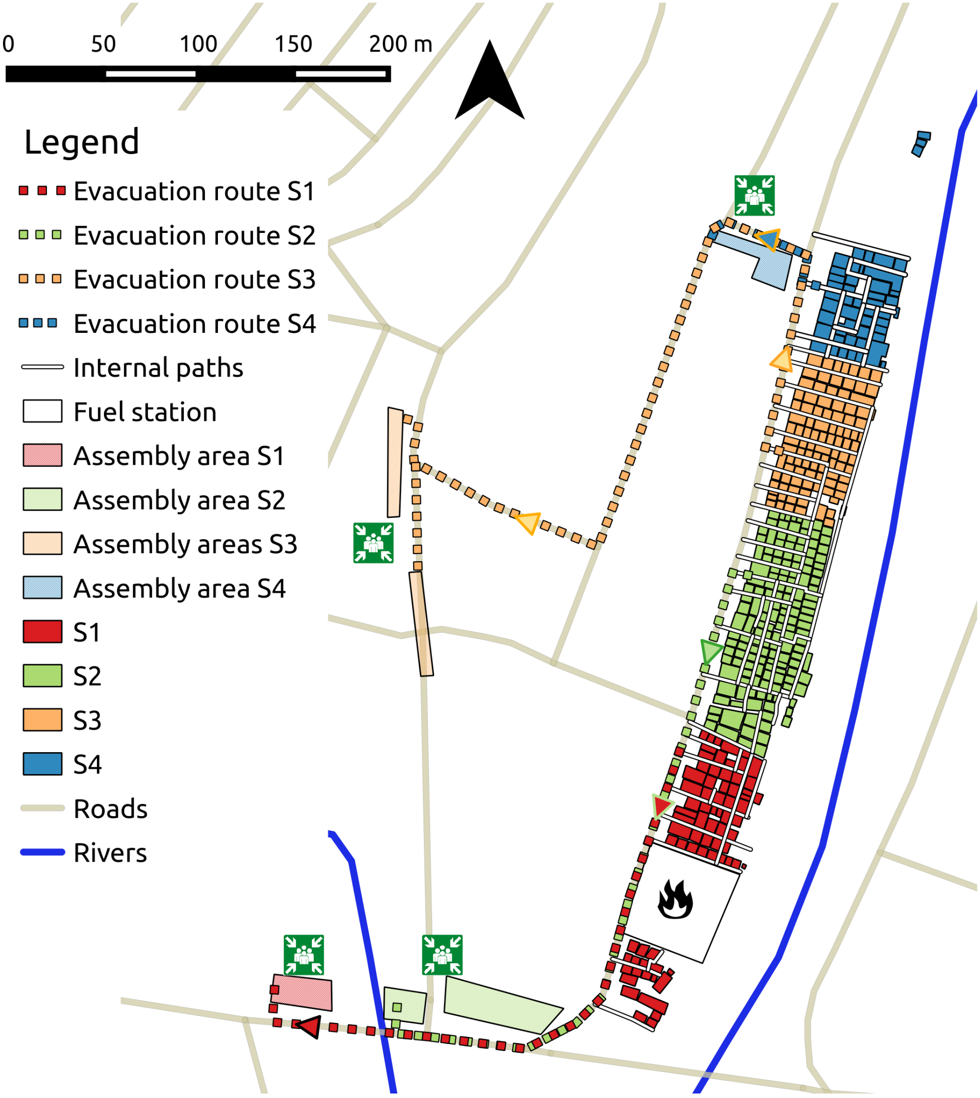


Figure 3 - Evacuation plan for the entire settlement: evacuation routes and assembly areas for each sector are depicted.

The evacuation plan and the emergency routes for the slum's dwellers have been evaluated in favor of security from the farthest point of each sector to the open space related to it. In this regard, results are reported in Tab.2. Furthermore, in the network analysis, the main road southward was closed for Sector 3 in order to not overcrowd the way. From the analysis on rescue services, it results in fact the entrance road for fire engines: therefore, in this case, a longer path is considered preferable in order to facilitate rescuers action.

5.2.2 Network analysis for rescue services

From the allocation of subnetwork to a competence area results that hospitals are homogeneously dislocated, such as army barracks, while fire brigades suffer from a lack of resources. In this paper, only the fire brigades' results are reported, as an explicative example of the analysis. There are in fact only two fire stations, which have to cover an extended area, as shown in Fig.4.

Emergency routes for rescue services responsible for Balkhu		
Rescue service	**Time (min)**	**Length (km)**
Hospital	4	1,5
Fire brigade KMC	6	3,5
Fire brigade LMSC	5	4
Army barrack	5	4

Evacuation routes for Balkhu's dwellers		
No Sector	**Time (min)**	**Length (m)**
1	5	385
2	6	455
3	7	550
4	2	150

Table 2 - Information about emergency routes for rescue services and evacuation routes for each sector of Balkhu, calculated as fastest paths.

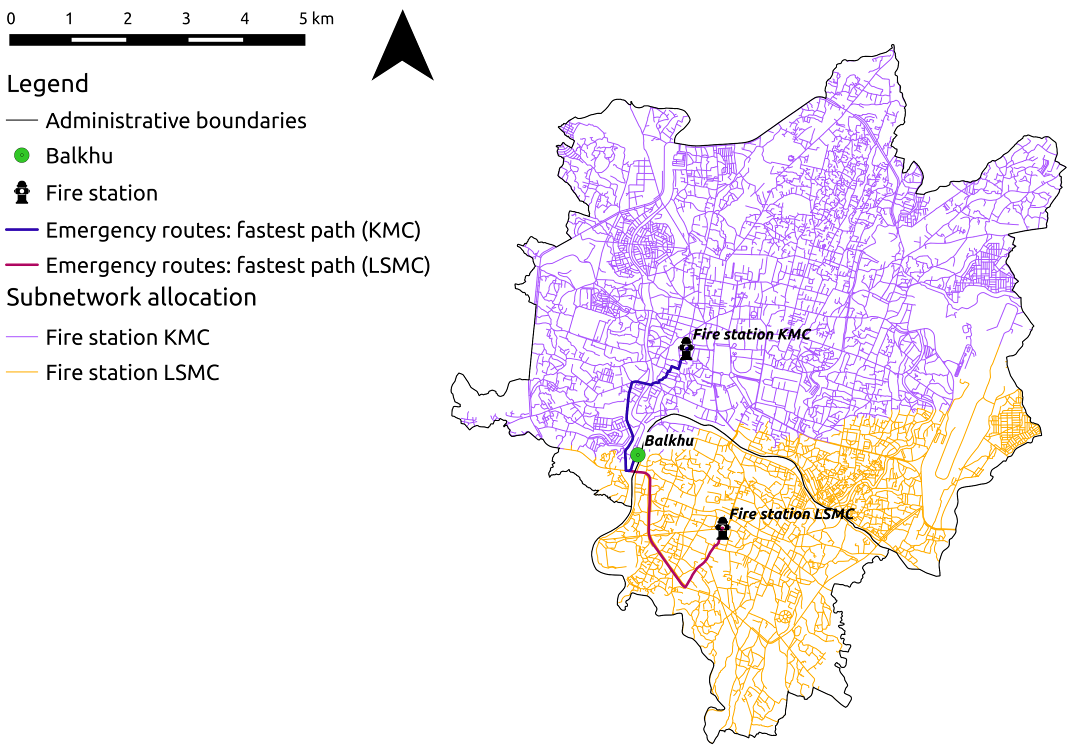


Figure 4 - Fire brigades' subnetwork allocation based on time, and emergency routes calculated as fastest paths.

As Balkhu falls in the middle of the two subnetworks, both of them were considered in the analysis. It results that in maximum 5 min the LSMC fire fighters could arrive in the area, although located farther from the settlement. In Tab.2, information about the fastest paths for each rescue service is reported, noticed that in maximum 6 min of transit time, all the rescue services would reach the settlement.

6 Discussion and perspectives

In order to reduce disaster losses, more efforts should be done on Disaster Risk Management: while initially more emphasis was given to disaster relief, recovery and reconstruction, thus eventually prevention and preparedness are recognized as central phases in order to address DRR. The improvement of response capacity of territorial systems to adverse shocks and the risk reduction are in fact widely accepted as drivers for the sustainable urban development. Nepal situation is critical and without improving housing spatial distribution and the transportation network, urban vulnerabilities are likely to increase (Bhattarai et al., 2010). As stated by Van Westen (2013), the use of GIS has become “an integrated, well developed and successful tool in disaster risk management”. Hazard assessment, elements-at-risk mapping, vulnerability and risk assessments all have important spatial components (Van Westen, 2013).

GIS has been chosen as frame for the decision support system and an emergency application in Balkhu settlement, as pilot area, is reported, applying a Network Analysis with the maps collected or created after an intensive fieldwork. The potentiality of the Analysis as emergency support tool is shown, involving rescue services and slums' dwellers as stakeholders and operating into two different spatial scales. Fastest paths were calculated, supporting rescuers in reaching the settlement and slums' community in the evacuation of the area. This work represents the first approach for an integrated emergency management in the Kathmandu Valley and it defines a methodology which can be extended potentially to the whole poor settlements in the Valley, or in other similar areas. However, it is recognized that more comprehensive knowledge of the Valley is needed. In fact, although the transit time of rescue services results comparable with the evacuation of people in Balkhu settlement, time for the call and preparedness of rescuers should be evaluated, along with other emergency related issues. In case of fire, for example, position of hydrants should be assessed, as well as water sources available. Additionally, a better roads network map is necessary, in order to fill missing information, like speed and accessible roads in terms of width and ways. As traffic jam is a common problem in the Kathmandu Valley, connecting this information with the analysis would make it more efficient and realistic. Furthermore, especially regarding emergency issue, the data collected should be verified in the field, in order to understand the effective resources available and their spatial distribution within the Valley. In this regard, the database should be open and continuously updated, in order to record changes in terms of infrastructures and viability. A future prospective could also include a climatological data update, to improve rescuers reaction and people mobilization. Moreover, environmental hazards maps are needed in order to create risks maps, crossing the information with vulnerability. Local authorities, with the support of NGOs, should initiated a focus on poor settlements, as these communities are still unseen, unheard and abandoned by the State, even if located in visible places. The objective was to begin the development of a decision support system, trying to define a procedural path to identify key elements of emergency plans, with a human face, as suggested by the UN Country team in Nepal (2013). People in fact have to be at the centre of development planning and awareness generation activities should be carried out. The final objective should include the improvement of the community's resilience, giving them the necessary resources and capacity of organization both prior to and during times of need.

References

- AMCDRR, Asian Ministerial Conference on Disaster Risk Reduction. (2014).
- *Bangkok declaration on disaster risk reduction in Asia and the Pacific 2014*, Kingdom of Thailand 22-26 June 2014.
- Bhattarai, K. & Dennis, C. (2010). Urban Vulnerabilities in the Kathmandu Valley, Nepal: Visualizations of Human/Hazard Interactions, J*ournal of Geographic Information System, 2010, 2, 63-84.* doi:10.4236/jgis.2010.22012.
- Dixit, K. (9-15 January 2015). Preparing to be prepared, *Nepali Times*.
- EMI & NSET-N &German Federal Foreign Affairs Office. (2010). *Structuring and Implementing a Competent Disaster Risk Management Function*. Kathmandu Metropolitan City, Nepal, February 2010.
- JICA & MOHA. (2002). *The study on Earthquake Disaster Mitigation in the Kathmandu Valley*. Kingdom of Nepal.
- Lumanti NGO. (2008). *Status of Squatter Communities in the Kathmandu Valley*, unpublished.
- Raj Arjal, K. (2011). Revisiting disasters responses in Nepal, Repùblica Opinion. From http://archives.myrepublica.com/portal/index.phpaction=news_details&news_id=27943
- Ravan, S. (2010). Spatial data to complement the use of space-based information for disaster management. *Geoinformation for Disaster and Risk Management. Examples and Best Practices*. Joint Board of Geospatial Information Societies, United Nations Office for Outer Space Affairs.
- Toffin, G. (2010). Urban fringes: squatter and slum settlements in the Kathmandu Valley, *CNAS Journal, Vol.37, No.2*.
- UN Country Team in Nepal. (2013). *A country Analysis With a Human Face 2011, updated in February 2013*, Kathmandu, Nepal.
- UNDP. (2012). *Capacity Assessment for Nepal. Urban Search and Rescue, Sustainable Training and Resilient Disaster Management*.
- UNISDR. (2014). *Progress and Challenges in Disaster Risk Reduction: A contribution towards the development of policy indicators for the Post-2015 Framework on Disaster Risk Reduction.* Geneva, Switzerland.
- UNISDR. (2010). *Disaster through a different lens. Behind every effect there is a cause*.
- University of Trento. (2010). *GRASS Tutorial – ver 6.4.0.* From http://www.ing.unitn.it/~grass/docs/tutorial_64_en/index.html
- UNSD. (2007). *Slum Population in Urban Areas*, New York: UN Statistics Division.
- Van Westen, C. J. (2013). Remote sensing and GIS for natural hazards assessment and disaster risk management. T*reatise on Geomorphology, Vol.3, Remote Sensing and GIScience in Geomorphology, 259-298*, 2013.

Alpinescapes: making the Alpine fringes visible trough OpenData and Crowdsourced Cartography

Michele Ferretti[1],Daniele Villa[1],Paolo Tagliolato[2,3]

[1] DAStU, Politecnico di Milano, Milan, Italy
[2] ISMAR, CNR, Venice, Italy
[3] IREA, CNR, Milan, Italy

Abstract

A recent redefinition of the Landscape concept promoted by EU institutions (European Commission, 2000), (JPI Cultural Heritage, 2010) stresses the importance of its social dimension, and of societal efforts in its construction and production (Castiglioni, 2009). Similarly, Societal Challenges permeate the debate around Cultural Heritage (European Commission, 2014). According to these documents such issues can be tackled by a combination of ICT technologies, applied to more transparent and participatory models of territorial governance (European Commission, 2014).
Building upon these premises, the Alpinescapes project poses itself as a novel tentative to link such multiple dimensions, by the means of FOSS4G (Free and Open Source Software for Geospatial), OpenData and Open licensing to collect, provide access and ultimately contribute to, the existing open content on the alpine landscape of Valsassina valley in Northern Italy.
Apart from building an easy-to-use platform for promoting tourism in the area, this experience sets an example in bridging the gap between the Users/Citizens and the Public Administration, in the collective representation of a territory, by the means of FOSS4G tools and OpenData.

Keywords

OpenStreetMap, OpenData, Cultural Heritage, Crowdsourced GIS, OpenGovernment, Web-mapping

1 Introduction

Seminal documents such (European Commission, 2000) and (JPI Cultural Heritage, 2010), posed a crucial milestone in underlining the Social dimension of Landscape, considered at the same time construct and product of contemporary societies, in a circular process in which citizens perceive, (re)construct and ultimately influence with their narratives, the Landscape. Such duality in the role of society becomes pivotal in the preservation and reproduction not only of the physical features of an area, but also of its culture and heritage (Castiglioni, 2009). The tight interconnections among these three components (Society, Landscape and Heritage) is further stressed by H2020 3rd pillar *Societal Challenges*, which advocates for a combination of new ICT methodologies for harvesting modern fluid forms of data, in conjunction with innovative, open, paradigms of Governance (European Commission, 2014). From this premises follows that, for any action pertaining to territorial transformations, societal involvement in the form of civic participation becomes paramount.

Against this background the *Alpinescapes* project has been established with the aim of valorising and preserving the Landscape and Cultural Heritage of *Valsassina*, an alpine valley within the *Grigne* mountain range in the province of Lecco (Northern Italy). Despite the beauty of its landscape, Valsassina, in particular its innermost municipalities, suffers from a relative uneasiness of access (due to the geo-morphological connotations of its surroundings). This condition partially threatens the diffusion of a tourism based economy, which is on the contrary a flourishing sector in nearby territories, such as the world renowned Lake Como Riviera.

Henceforth, to conjugate EU’s directives with the needs of a local community, this case study has chosen Open Geospatial Technologies and Open (geo)Data not only to act as a mere background or simply as a medium for collecting and representing already present information about the landscape, but as its entire foundation. Open community efforts such as OpenStreetMap and Wikipedia are in fact projects that, by virtue of their characteristics, can be considered the most suitable to allow a participatory process in which the end-user become at the same time sensor/producer, as well as consumer, of the information thus gathered (Goodchild, 2007).

2 Materials and Methods

Given the two main aims of the project (allow the exploration of the alpine landscape and contribute to open databases) a crucial part in devising the application has been the choice of its architecture, which had to exhibit an agile and open structure, allowing for a constant insertion and update of data (hopefully also produced by its users).

To reach such goals the project has been built with a lean format: At its core it's a single-page web-app, built on top of the open source project *icì* (Summers, 2014), with the relatively straightforward anatomy of a client-side application. It's built with FOSS geospatial javascript libraries, such as LeafletJS (http://leafletjs.com), able to retrieve JSON data via *HTTP GET* requests, and to serve them using the latest *HTML5*-compliant technologies. In this sense, the application becomes almost a “container” for displaying external data. This configuration, without the direct inclusion of any server or back-end applet not only avoids the complexity of maintaining and updating such infrastructure in face live ecosystems, such as the OpenStreetMap Database (OSM), but also accounts for the aforementioned flexibility in connecting multiple data sources.

With such characteristics in mind, the application offers mashups of geo-referenced information from heterogeneous data sources: 1) *Wikipedia* offers geo-tagged crowd sourced articles 2) *OpenStreetMap (OSM)*, from which ad hoc geographic features are fetched, related to the mountainous environment, i.e. hiking and MTB trails, alpine refuges etc. 3) geo-referenced information about the local area history and heritage from *Comunità Montana della Valsassina*, a territorial public body. A fourth database, which at the moment of the writing is undergoing an evaluation finalized to it's release in an open format (the licensing should be compatible with the other above mentioned data sources), is constituted by layers of environmental interest from the Geographic Database of Lecco Province. These data, which will eventually be imported into OSM, will serve to enrich the map with more detailed landscape elements, not already traced.

To access the open data archives the application uses standard *HTTP GET/POST* protocol requests, connecting to the APIs of different query engines, such as Overpass API (http://overpass-api.de) in the case of OSM, or MediaWiki (https://www.mediawiki.org) for Wikipedia. The requests responses, obtained in plain *(Geo)JSON* are injected in real-time in the web-page, according to the features chosen by the user, and listening for his interaction with the map (panning/zooming, geolocation..). In order to achieve cross-browser and cross-platform compatibility standards, the web-application is written within the framework of Twitter Bootstrap (http://getbootstrap.com) and makes use of libraries such as JQuery (http://jquery.com) and Modernizr (http://modernizr.com). This combination ensures a responsive, mobile-first application, making *de facto* useless the implementation of different viewers to experience the content on multiple screen sizes.

The user interface is further devised according to common UX/UI practices: The app presents itself with a full screen map, which is a version of Mapbox Terrain layer customized with Studio (https://www.mapbox.com/mapbox-studio). A simple responsive navigation toolbar is present on top of it. The actual map-controls menu, featuring functions such as the geolocation and a geocoder, a LeafletJS plugin built on OSM Nominatim (http://wiki.openstreetmap.org/wiki/Nominatim), is present on the left-hand side of the screen, while a togglable content menu is present on the right-hand side. From this menu the user can display the paths and points of interest, as well as further refine his search, by selecting the features contained within a fixed-radius (500m) dynamic geographic filter, provided by MapboxJS APIs (http://mapbox.com). This option conveys an impression of the closest elements without the need to actually perform server-side spatial queries. Lastly, to foster contribution towards the open databases onto which is built, the app offers an “Edit in OSM” button, which directly links to OSM, as well as is able, with a functionality inherited by Summers (2014), to discriminate the eventual lack of references/images within a Wikipedia article, warning the user and asking for his help.

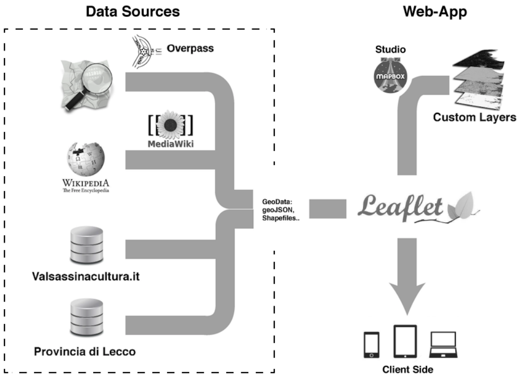


Figure 1 - App Conceptual Architecture

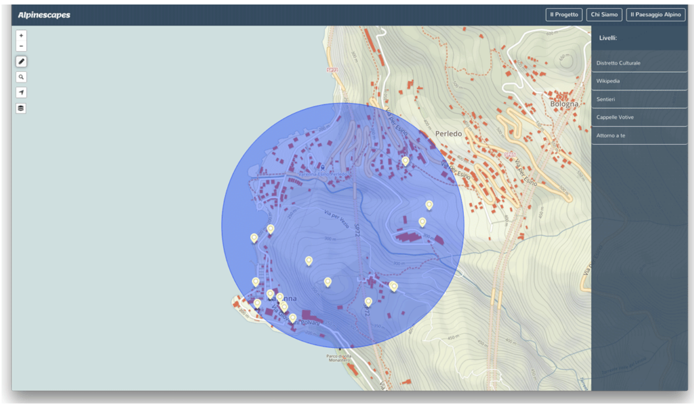


Figure 2 - User Interface (detail: geo-search radius)

3 Partial Conclusions

Alpinescapes thus poses itself in the wider disciplinary tradition of approaches exploiting user-generated content (Ratti, Pulselli, Williams, & Frenchman, 2006) and VGI (Goodchild, 2007) mixed with mobile computing, to render and translate complex ecologies of Information into more easy to access and interactive geo-visualizations. This let the users, to paraphrase Amoroso, Hudson-smith, Phillips, & Speed (2013), "*become instrumental in how the changing landscape is interpreted, rather than being passive viewers of the represented landscape*". Despite the straightforwardness of its architecture, the web-app further presents a rather positive initiative combining different datasets into an agile and open architecture, mainly suited for interacting with open licence databases. This operation is strategic in illustrating to the public sector, with a proven, factual example, the feasibility and the convenience of operating with such kind of data, and in particular with FOSS tools. An operation that should be especially relevant for the public sector in times of increasing expectations yet decreasing resources (European Commission, 2014).

4 Acknowledgements

The Alpinescapes project has been made possible with funding from an Italian-Swiss Interreg agreement. Among the partnership subscribers can be found the *Comunità Montana della Valsassina* (Italian side) as well as *SUPSI Lugano* (Swiss side).

References

- ✔ Amoroso, N., Hudson-smith, A., Phillips, M., & Speed, C. (2013). Patch-Scape: Pads, Pods, Phones and Spatial data. In Digital Landscape Architecture DLA Conference 2013. Bernberg, GER.
- ✔ Castiglioni, B. (2009). Di che è il Paesaggio? La partecipazione degli attori nella individuazione , valutazione e pianificazione. (B. Castiglioni & M. De Marchi, Eds.) (I., p. 172). Padua, IT: C.L.E.U.P. “Coop. Libraria Editrice Università di Padova.”
- ✔ DPSCE. (2012). Le aree interne: di quale territori parliamo? Nota esplicativa

sul metodo di classificazione delle aree (pp. 1–14). Rome, IT. Retrieved March 16, 2015 from http://www.dps.tesoro.it/aree_interne/doc/Nota Territorializzazione AI_03 marzo_2013.pdf

- ✔ European Commission. (2000). European Landscape Convention - Convention européenne du paysage. Strasbourg, FR.
- ✔ European Commission. (2014). HORIZON 2020 WORK PROGRAMME 2014 – 2015 Europe in a changing world: inclusive , innovative and reflective Societies (Vol. 2015, p. 112).
- ✔ Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. GeoJournal, 69(4), 211–221.
- ✔ JPI Cultural Heritage. (2010). The Joint Programming Initiative on Cultural Heritage and Global Change : a new challenge for Europe - Vision Document (pp. 1–46). Rome, IT.
- ✔ Ratti, C., Pulselli, R. M., Williams, S., & Frenchman, D. (2006). Mobile Landscapes: using location data from cell phones for urban analysis. Environment and Planning B: Planning and Design, 33(5), 727–748. doi:10.1068/b32047
- ✔ Summers, E. (2014). icì. E. Summers. Retrieved July 5, 2014 from https://github.com/edsu/ici

Operation-based revision control for geospatial data sets

Máté Cserép, Roberto Giachetta

Eötvös Loránd University

Abstract

Revision control is useful and sometimes essential tool in the maintenance of changing datasets, and the need for version control emerges in several data centric environments, such as Geographic Information Systems (GIS). However, current revision control software focus primarily on textual information as no method is available for the efficient storage of modifications on any binary content. Beside this issue of space efficiency, retrieving semantic information regarding the alterations applied between selected revisions is also an unsolved problem for most cases. In this paper we study the application of operation-based revision control on geospatial data, specifically vector features and remotely sensed raster imagery. We present an approach relying on a unified data model and operation metadata as modification deltas. The method is independent of storage solution, thus it is applicable for both files and databases, while also persisting the required semantic information about the changes. Our evaluation shows that operation-based revision control handles geospatial data more efficiently than current state of the art revision control software, including systems created especially for maintaining geospatial data.

Keywords

revision control systems, geographic information systems, remote sensing, distributed systems

1 Introduction

The evolution and spreading of data capturing methods ranging from simple GPS devices (Wang, et al., 2010) to large scale imaging equipment (including very high resolution and hyperspectral cameras, LIDAR, etc.) (Zhang, 2010) resulted in an exponential growth in the amount of spatial data maintained by companies and organizations for which cloud computing technologies and distributed systems must be utilized (Yang, et al., 2010). Also, usually multiple versions of the same data exist due to additional modification. Hence, to manage data efficiently a spatial revision control system is required which fits to the requirements raised by Geospatial Information Systems (GIS).

Revision control systems are widely used tools in software development primarily aimed on managing versions of software source code during implementation (Ruparelia, 2010). Due to source code being written in plain text format, these systems are generally usable for any text document. There are also solutions available for text based storage of complex objects (Firmenich, et al., 2005). However these systems are not optimized for storing versions of binary data, as its format is too general for tracking changes within the binary code. Some revision control systems – for instance, Git[1] – offer limited support for binary content (Dobŏs & Steed, 2012).

In GIS several file formats exist for storing geospatial information, most formats being binary. Only a handful of data formats have textual representation, including Well-known Text (WKT), GML and GeoJSON. Although some new solutions have emerged in the revision control of geospatial data including GitSpatial[2] and GeoGig[3] (formerly GeoGit), these systems are rather concentrating on specific areas. No universal solution has been suggested with respect to all kinds of spatial data and possible operations performed on it.

Treating spatial information as binary data produces several major drawbacks compared to textual revision control. Firstly, it erases all semantic information about the applied changes, like the executed spatial operations. Secondly, even in cases when only a smaller fragment of the given geospatial dataset was modified, a single editing operation could result in the change of a larger fraction in the binary data, needlessly increasing the required storage.

The demand for revision control systems in GIS software capable to exploit the peculiarities of spatial operations arose as early as the beginning of the '90s (Easterfield, et al., 1990). Alongside the several attempts made to utilize the concept of version control in standalone GIS solutions (Wachowicz & Healey, 1994), the idea of DBMS[4]-integrated systems (Newell, et al., 1992) (Peuquet & Duan, 1995) were also researched in cooperation with the database community. However, no sufficient solution was implemented in contemporary GIS software to satisfy the previously defined requisitions of the storage space compactness and the persistence of semantic information.

In this article, we present a solution for the efficient management of geospatial data using operation-based revision control. The model is implemented as part of the *AEGIS geospatial framework* (Giachetta, 2014), a generic library for geographic and remote sensing data processing.

2 Revision control models and methods

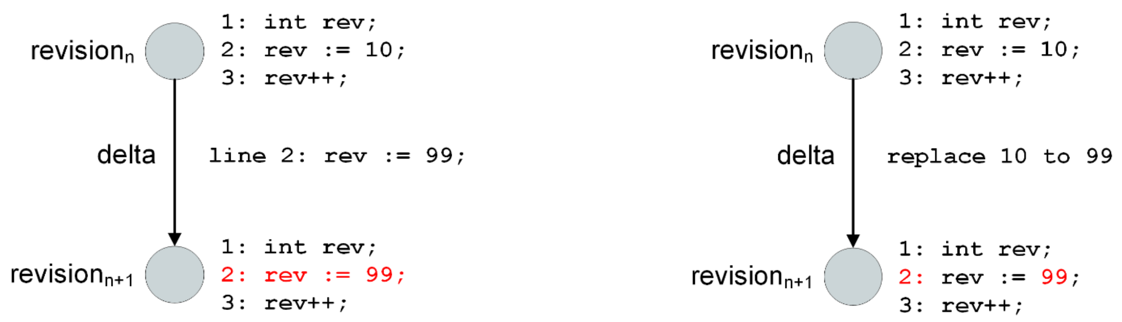
Preserving a full copy of every version of a document can easily take up a significant amount of storage space even with a relatively short version history. As a commonly used solution, most modern revision control tools spare storage space by computing and persisting only the difference (*delta*) between succeeding revisions, and only preserving the full state of a few special versions – like initial or final (*head*) revisions.

The two main categories for producing the changeset between versions are *state-based* and *operation-based* deltas. An example for both models in case of

1 http://git-scm.com
2 http://gitspatial.com
3 http://geogig.org
4 Database management system

text documents is presented in Figure 1.

In the state-based case the general method is to decompose the document into smaller, more easily manageable parts. This is due to comparing two versions of a document as a whole would induce several difficulties when creating, applying or merging deltas. For textual documents an often applied practice is to dissolve a file into lines, ignoring the special structure (syntax and semantics) of the content. Compared to this unstructured data processing, another approach is to – partially – statically analyze the document and decompose it in a structured or semi-structured way. This latter method reduces the occurrence of possible unresolvable merge conflicts (Apel, et al., 2011), however also tailors the manageable content of the revision control system to a pre-defined format. While these techniques provide a poor storage efficiency for unstructured binary data, they can be used for tracking changes in textual geospatial information (like GitSpatial handles GeoJSON), and structured models are also feasible for specific binary file types (like GeoGig parses shapefiles).



(a) State-based *(b) Operation-based*

Figure 1 - Example of revision control models

In contrast, operation-based deltas store the actually performed operations between two succeeding versions in the revision control system, instead of the alterations in the state. While this approach also has benefits in managing textual data (Shen & Sun, 2001), it is notable for this concept to not necessarily requiring any special knowledge about the managed data format from the revision control system to create or merge changesets. As a utilization in the graphical field Chen, et al. (2011) presented a revision control system using the operation-based model for the image editing application GIMP. The core idea is the storage of graphical operations using a directed acyclic graph (*DAG*), where each vertex contains a performed editing operation. This representation is suitable for nonlinear revision control and does not require storing the state of the modified image.

Different applications of revision control may require different models. Operation-based revision control is favorable over the state-based when the following conditions are met.

- The amount of data affected by a single modification is generally large. In this case state-based deltas require large storage space, whilst the size of operation-based deltas is independent from the size of the affected data. For example, an image operation such as histogram equalization can modify all pixels of the image, requiring the entire image to be stored again in case of the state-based model.

- The data is heterogeneous, and cannot be divided into small parts, where the modification of the individual parts can be monitored. The state-based model is mostly effective when small fragments of data are identified as location of the modification.
- The query of former revisions is performed infrequently. Frequent queries would cause performance issues, as operations need to be re-executed to present the individual versions.
- Both data and operations can be managed through a single, high level model. This is required to enable the uniform handling and storage of the operation-deltas, and the proper identification of information stored within binary content.

Due to these reasons, revision control of geospatial data is more favorable for the operation-based model. The data is heterogeneous, stored as binary content, and operations usually modify large parts of the dataset. Previous revisions are not required often (as in case of software source code), but still, revision history needs to be maintained and looked up frequently.

An example for geospatial workflow consisting of four operations can be seen in Figure 2. First, the remotely sensed image is registered for the specified reference system. Second, the image is enhanced by using histogram equalization. Third, selection of the area of interest is defined by intersection using a specified input polygon. Finally, for classification, thresholding is applied to the area.

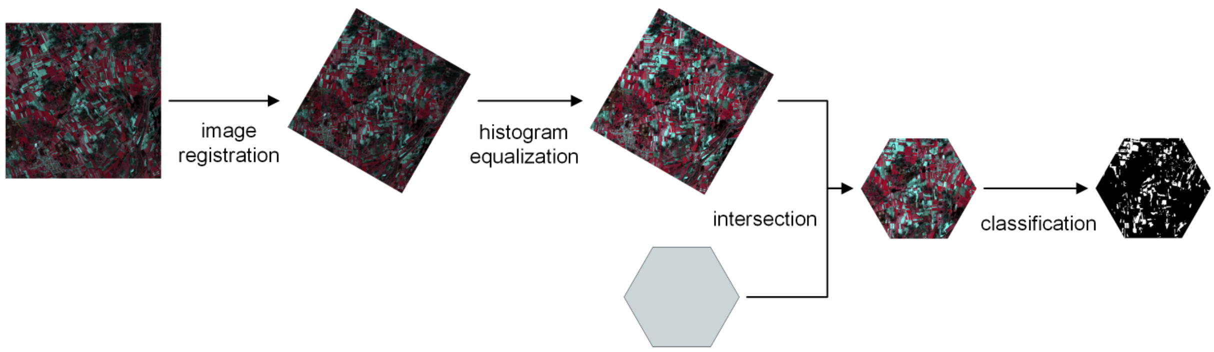


Figure 2 - Example workflow with 4 operations

3 Operation-based revision control model on geospatial data

The core concept of our solution was inspired by the idea described in (Chen et al., 2011). To create an expansively applicable model, the representation of geospatial data and operations were designed by raising the least possible requirements against them.

3.1 The baseline model

The data model purely assumes that all geospatial data types in the system directly or indirectly inherit or implement a common ancestor class or interface. A well-known compliant example for the before mentioned minimalist specification is the *Simple Feature Access* (SFA) standard[5] (Herring, 2010) by the Open Geospatial Consortium (OGC)[6]. The SFA is an object-relational mapping for geospatial data, providing an abstract *geometry* as a common base for spatial objects, including collections. The standard deals only with the handling of vector data, but can be easily extended to support other features. For example, support for raster imagery can be introduced by specializing polygons to support raster data (denoted *raster polygons*, see Figure 3). By using this abstraction the source of data is irrelevant in terms of the revision control system.

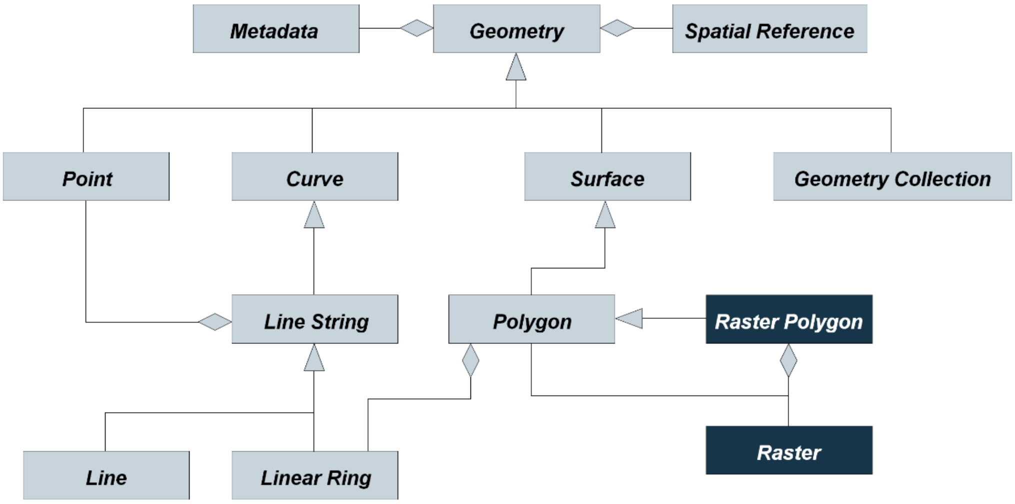


Figure 3 - The extended Simple Feature Access data model (UML notation)

Geospatial operations can be defined as a mapping between geometry objects, and can be represented as transformation objects. The objects can be described by the applied method and the values of arguments (if any). This concept is introduced in the OGC *Spatial Referencing by Coordinates* (SRC) standard (Cooper, 2010) for coordinate transformations, but can be easily generalized for all kind of geospatial operations, including raster image processing. Using this approach, only the descriptors of the operation are required to be stored as operation-deltas, which is compact data, and independent of the size of the actual transformation object.

For example, in case of the workflow presented in Section 2, histogram equalization and intersection only need the source geometries ($geom_i$) and no additional arguments. Thresholding requires the predefined threshold (th), whilst registration requires the spatial reference system identifier ($SRID$) and the geographic coordinates of the image corners ($coord_i$) in addition to the source geometries (image and intersection polygon). An operation-based revision control system does not require any specific knowledge about the previously executed spatial transformations (as described in Section 2), therefore no further requisite is necessary towards the model of operations.

5 Also an ISO standard: ISO/IEC 19125:2004
6 http://www.opengeospatial.org

op_1 = (registration, $geom_1$, SRID, $coord_1$, …, $coord_4$)
op_2 = (hist. equalization, $geom_1$)
op_3 = (intersection, $geom_1$, $geom_2$)
op_4 = (thresholding, $geom_1$, th)

As the core structure for storing the version history the *directed acyclic revision graph* is used. In this representation each vertex symbolizes the corresponding version and contains the ordered sequence of operations performed between the current and the predecessor revision, while edges denote the semantic relationships in the version system. A possible graph for the example workflow is presented in Figure 4 with rectangles denoting geometry storage and circles denoting operation-delta storage. The initial revision contains the source geometries, whilst further revisions are computed using the specified operation.

$$revision_1 = (image, polygon)$$
$$\forall i \in 2...5: revision_i = (revision_{i-1}, op_{i-1})$$

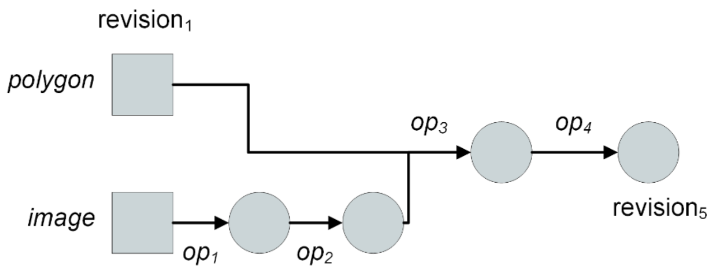


Figure 4 - Revision graph of the example workflow

3.2 Data storage

The high-level data and operation model enables the usage of all SFA compliant storage methods as source. This includes vector file formats, such as Shapefiles, GML or GeoJSON, as they can be simply converted to geometry format. The extension for handling of raster data also enables image file formats, such as GeoTIFF. Furthermore, other storage mechanisms can also be used, such as spatial databases. Thus the model is not only applicable for file based storage of revisions, but can be used on any spatial data source.

As stated in Section 3.1, operation-deltas are stored using descriptors, which can be serialized to a compact text or binary representation. Even more, the descriptors can be aggregated to the geometries as attributes. As most spatial formats enable the storage of attributes beside the geometry, these formats allow straightforward usage for storing the revision history of a geometry. For example, Shapefile stores descriptive information within a dBase file. Although specialized software is required for deserialization of the revision history, the stored geometries are readable by most geospatial software.

3.3 Centralized revision control

Our revision control system applies a natural, object-oriented approach for representation, where versions are embodied by *revision descriptor* objects. These descriptors contain meta information about the versions (e.g. the identifier and the predecessor revision), while the changesets encompass the executed editing operations are stored and can be obtained independently for each version. This representation architecturally separates the attributes and the deltas, accelerating several revision management algorithms where the latter information is irrelevant.

The states of the spatial data managed by the revision control system may be interpreted as sets of geometrical objects. Deltas between any two versions can contain the addition of new or the deletion of existing geometry objects from the state set beside the transformation operations needed to be applied.

Using the above described revision model in linear version control systems, it becomes a straightforward task to construct the cumulative changeset between any two revisions. However, to update the working copy of a given user to a former version, the system has to revert the performed editing operations in a reversed order on the actual state. While the inverse of state-based deltas are easily constructible, operations can be irreversible. In such cases it is the task of the revision management system to support alternative approaches handling the problem. Multiple possible fallback mechanisms exist to overcome this issue, like recreating the selected revision from the other end of the version history or exceptionally storing the pre-transformed state of the affected geometry object in the original revision changeset in case of irreversible operations.

In an operation-based revision control system like the designed model, it was already shown that only the initial states of the geometries are needed to be stored, further modifications are represented with transformation objects. For efficiency reasons, fittingly selected revisions can also be snapshoted, which store the entire geometry set, so it can be retrieved without the necessity of reapplying operations. This feature makes available some balancing options in the storage space against time efficiency issues. A typical application of this method is to create a snapshot of the head revision of the main – or every – branch in a version management system, and store reverse directed inverse deltas in the changesets. In such models the fresh – significantly more often checked out – revisions can be obtained with an outstanding performance, therefore also many existing revision control tools (e.g. Subversion[7]) already use a similar methodology.

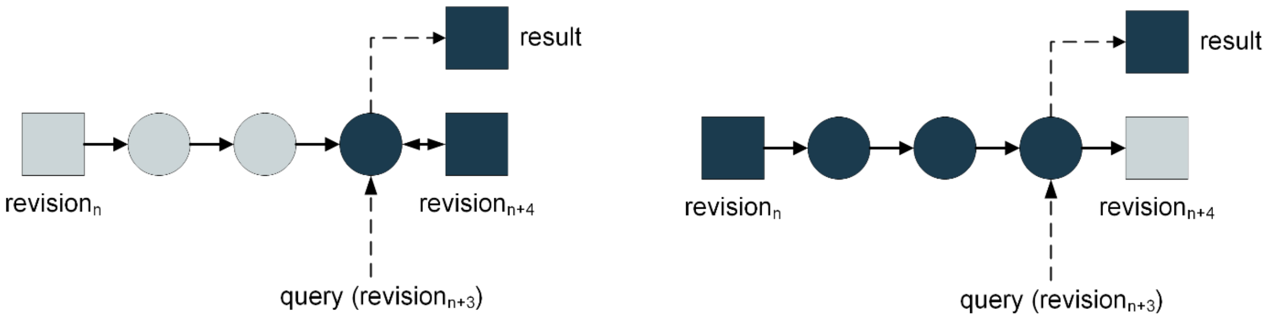
An example query is presented in Figure 5. Assuming that $revision_n$ is stored as snapshot, $revision_{n+4}$ is the head revision, and additional revisions using operation deltas, the query method for $revision_{n+3}$ depends on the reversibility of the final operation. If the operation is reversible, the revision can be reconstructed using a reverse delta from the head revision (Figure 5a). Otherwise, the revision is reconstructed from $revision_n$ by execution of multiple operations (Figure 5b).

This model also complies with the nonlinear revision control paradigm (branching and merging) by enabling multiple revision descriptors to reference a common predecessor. With this extension, branching itself becomes a simple task. However, there are multiple issues left with merging. For example, the

7 http://subversion.tigris.org

aggregated changesets might contain not only a single, but multiple editing operations per branches for a geometry object – the atomic unit in the model. Therefore the corresponding order of the alterations must be defined to resolve the conflict, for which task the involvement of the user is generally unavoidable.

The support of merging also requires some slight modifications of the revision descriptor objects. As a descriptor declares a single other version as its predecessor (and the changeset defines the modifications between these two revisions), the information about which other version(s) were merged into the current revision is lost. Therefore the identifier of all the merged versions shall be stored in every revision descriptor as this meta information might be important or helpful to the human users.



(a) Query using reverse delta *(b) Query using forward deltas*

Figure 5 - Revision query in the operation-based model

3.4 Distributed revision control

As cloud computing and distributed data storage became a generally widespread and popular research topic in recent times, most third generation version control tools (O'Sullivan, 2009) – like Git or Mercurial[8] – also abandoned the centralized paradigm to implement distributed data management and revision control instead. This concept ensures a higher availability of service and better distribution of processing and load (Reuter, et al., 1996).

The presented revision control model also supports this paradigm. A higher level software layer can be created above the model, which is responsible for the network communication and cooperation between repositories. The synchronization of the individual repositories can be performed by one of the following methods.

- *Data based synchronization* of deltas, snapshots and head revision. As in the state-based model, this method requires transferring large loads of data, and replacement of remote content to the new content. Update of the remote repository can be performed quickly.
- *Operation based synchronization*. This method requires only the transfer of operation-deltas, thus is more applicable in low bandwidth environments. Based on the operation deltas, the remote repository can rebuild any snapshot and the head revision. Therefore this method requires additional execution of operations.

8 http://mercurial.selenic.com

> The benefit of this approach is that the rebuilding of revisions is not needed to be performed immediately. The construction can be delayed until the next query of the repository.

The method of synchronization is not required to be determined beforehand. Both methods can be supported by the system, and based on current environmental conditions, the appropriate method or even a hybrid approach (e.g. transfer of the head revision, but reconstruction of the snapshots) may be chosen automatically by the system with each synchronization.

Special forms of repositories can also be formed, which are only used as remote sources, such as the “bare” repository in case of Git. When using operation based synchronization, no rebuilding is required, as the revisions are reconstructed anyway by the client repository. This approach saves both storage space and execution time.

3.5 Implementation considerations

To prove the usability of the presented revision control model the implementation was carried out as part of the AEGIS geospatial framework as the system satisfies all requirements described in Section 3.1, including a general, abstract data model based on SFA and operation support which provides operation metadata (Giachetta, 2015). The implementation of both the AEGIS and the revision control system was carried out using the *.NET/Mono Frameworks* to utilize the possibilities and the simple usage of this object-oriented development platform.

As it was stated, the physical storage of the version history is independent from the revision control model. The various implementations of the storage component do not contain any specific knowledge about the inner architecture of the revision system, but they are solely responsible for storing and returning the requested data regarding a particular physical storage (e.g. a file system or a spatial database). Despite the specific inner data representation of the system, the sample implementation of the revision control model presented in the current section is easily connectible via the AEGIS framework, allowing the usage of widespread geospatial formats and services.

4 Experimental results

The practical applicability of the designed and implemented version control framework can be measured through comparing the developed software with existing revision control tools in the research field. In the interest of quantifying the result of the comparison, storage efficiency, the most significant attribute of revision management systems was chosen and examined in Section 4.1.

As the version history for a project increases in size, the recreation of an arbitrary revision with a satisfactory performance in speed could meet obstacles, as the cause of this issue and possible solutions were discussed in Section 2 and Section 3.3. In Section 4.2 we measure and analyze this other substantial aspect of our version control tool for different scenarios.

4.1 Storage efficiency

Three existing version management tools were selected as the basis of the benchmark: Subversion, Git and GeoGig. The former two software are probably the most used revision control tools in the world (O'Sullivan, 2009) for general purposes, while GeoGig is the only other known and formerly[9] actively developed software specifically for tracking changes in (binary) spatial data.

The tested data format was the Shapefile, as it is a widely used binary storage format for vector-graphic information supported by all four compared software. This type of binary data is usually handled by general delta producing algorithms in most classical revision control tools.

Most version control tools used in production – like Subversion and Git – compute not only the deltas instead of storing the full snapshots of the data, they also apply lossless compression to further reduce the required storage space. In order to avoid any remarkable distortion in the results, a similar method[10] had been implemented in the prototype tool created with AEGIS. Since Git uses a l*oose object format* (Chacon, 2009), the storage space allocated by Git was determined after the repacking of the objects was enforced manually into *packfiles*.

Two test were carried out to measure storage efficiency. In the first one a repository of all four revision control tools were populated with 1.000 revisions of randomly generated subsequent states of a single Shapefile, while the editing operations evaluated between the versions were also provided for our prototype operation-based system. The experimental results of this test are shown in Figure 6a, where for reference purposes the storage space requirement without any revision control tool applied was also indicated, exceeding 33 gigabytes. The outcome of the analysis clearly shows that while Subversion and Git required about 3-5 percent of the unversioned storage space, the AEGIS prototype demanded a much more limited amount of space, around 0.2 percent. The results of the GeoGig testing might be a surprise, hence its repository occupied more than 10 percent of the storage space of the full snapshots of all versions altogether. The explanation for this result might be the loose, superfluous and redundant inner data representation of GeoGig. This unfortunate aspect of the software was significantly improved compared to previous versions, however the tool is still in a beta phase and the progress of development slowed down recently.

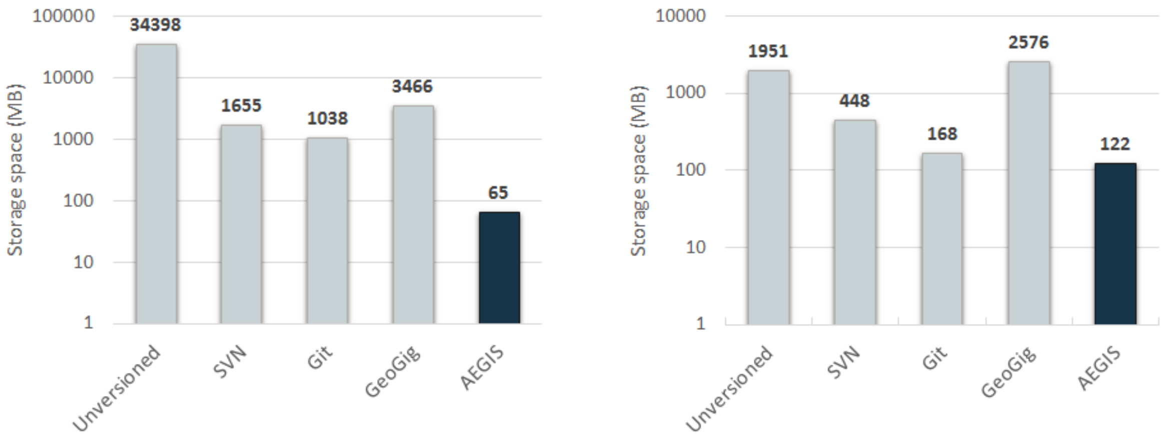
In order to evaluate the storage requirement efficiency in a real-world scenario, the second test case was based on the available history of the OpenStreetMap datasets about Hungary in Shapefile format[11]. Altogether 11 revisions were used covering over a year of timespan. The evaluation of the second test case is displayed on Figure 6b, revealing similar results to the previous test on generated data. While the gathered and stored semantic information regarding

9 The development of GeoGig was unfortunately suspended a few months ago.

10 As both SVN and Git uses the *DEFLATE* procedure for compression purposes, we used the *LZMA*, since both are based based on the algorithm described by Lempel and Ziv (1977).

11 The shapefiles were downloaded from the publicly accessible history available on the www.geofabrik.de website.

the editions is a secure benefit of the operation-based model, the vantage of AEGIS in space requirement is evidently relaxed in this case. The reason is that only a few revisions were analyzed in quantity and a significant part of the modifications were additions of new data as the OpenStreetMap dataset about Hungary increased by over 60% in size over the last year.



(a) Generated dataset *(b) OpenStreetMap dataset*

Figure 6 - Storage efficiency comparison of revision control tools

This section proved that an operation-based approach can not only preserve crucial semantic information about the editing operations, but even a prototype level implementation of a geospatial revision control system can comply a better performance in storage space requirement than the general binary delta algorithms applied by the version management systems in production.

4.2 Computation performance

Alongside the advantages emphasized at the end of the previous section, the potential drawbacks regarding the computational efficiency of the operation-based architecture must also be examined, because in this model the changesets contain sequences of transformations instead of state modifications. Since the evaluation of certain spatial operations can be quite time consuming and these transformations have to be executed each time when recreating a revision from the state of another one, it might cause a remarkable effect on the performance of such a system. The problem can be diminished with proper implementation decisions, for instance storing the head state of the branches and persisting reverse deltas in the changesets ensures that the fresh revisions are always easily producible by the framework. Moreover, automatically saving snapshots by a declared heuristic can also guarantee a maximum limit of revision recreation time at a cost on storage space. However, as stated in Section 3.3, not all operations are reversible.

As a practical demonstration of our before mentioned theory we created three repositories with our prototype tool, containing the same spatial version history of a few hundreds revisions. The changesets attached to each version primarily consisted of time expensive transformation operations. The three repositories applied different models related to the changesets: the first one used forward deltas, the second one applied reverse deltas and the last one also worked with reverse deltas, but persisted a snapshot of the actual state at every 20th revision on each branch. Figure 7 shows the recorded average checkout times for the initial, the head and randomized revisions in the repositories, presenting the performance advantage of reverse deltas with snapshots applied[12].

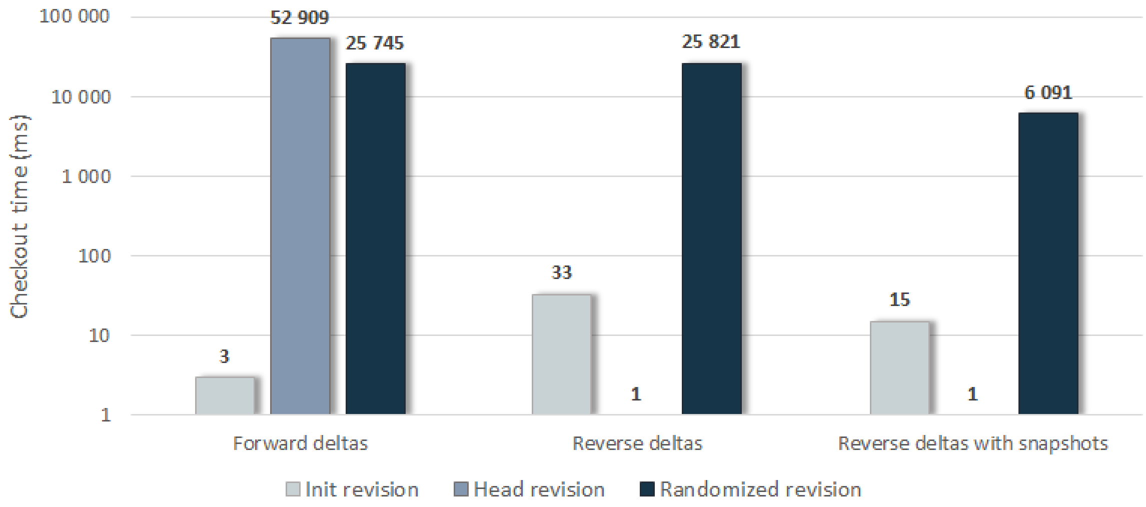


Figure 7 - Speed performance comparison of different methods

5 Conclusion and future work

In this article we presented an approach to geospatial data management using operation-based geospatial revision control model. As demonstrated, using a high level data model the method may support both vector and raster data and a variety of storage solutions. We argued that despite the inevitable computational overhead of this architecture, considering the peculiarities of the geographical information systems, this model is capable to deliver formerly unobtainable benefits on the field of revision control of spatial data, like persisting the semantic information of modifier operations. Our article presented a detailed design on that concept enabling not only linear control, but also branching and tagging, and usage of distributed repositories. Evaluation has shown, that our approach is capable of matching and even excelling efficiency of currently available solutions both in terms of storage space and computation performance.

Future work includes the development of a production ready revision control tool by further performance improvement of the prototype implementation. The thorough examination of complex snapshot creation heuristics is also required to find convenient and balanced solutions for snapshot management. The application of revision control in distributed data management is another future research topic, as AEGIS is capable of distributed data processing using the MapReduce paradigm (Giachetta, 2015). Revision control may be utilized for different purposes in distributed environment, such as data backup, recovery and transfer (Vrable, Savage, and Voelker, 2009).

12 In a production environment naturally a better heuristic should be chosen, for instance regarding not only the quantity, but also the complexity of the changesets.

References

Apel, S., Liebig, J., Brandl, B., Lengauer, C., & Kästner, C. (2011). Semistructured merge: Rethinking merge in revision control systems. *Proceedings of the 19th ACM SIGSOFT Symposium and the 13th European Conference on Foundations of Software Engineering (ESEC/FSE '11)* (pp. 190-200). New York: ACM.

Chacon, S. (2009). *Pro Git* (1st ed.). Berkley, CA, USA: Apress.

Chen, H.-T., Wei, L.-Y., & Chang, C.-F. (2011). Nonlinear revision control for images. *ACM Trans. Graph., 30*(4), 105-115.

Cooper, P. (Ed.). (2010). *The OpenGIS Abstract Specification - Topic 2: Spatial referencing by coordinates* (version 4.0 ed.). Wayland, MA, USA: Open Geospatial Consortium.

Dobŏs, J., & Steed, A. (2012). 3D Revision Control Framework. *Proceedings of the 17th International Conference on 3D Web Technology (Web3D '12)* (pp. 121-129). New York: ACM.

Easterfield, M. E., Newell, R. G., & Theriault, D. G. (1990). Version Management in GIS - Applications and Techniques. *Proceedings of the European Conference on Geographical Information Systems (EGIS 1990)* (pp. 1-8). Amsterdam: Elsevier.

Firmenich, B., Koch, C., Richter, T., & Beer, D. G. (2005). Versioning structured object sets using text based Version Control Systems. *Proceedings of the CIB W78's 22nd International Conference on Information Technology in Construction* (pp. 105-112). Dresden: Technische Universität Dresden.

Giachetta, R. (2014). AEGIS - A state-of-the-art spatio-temporal framework for education and research. *OSGeo Journal, 13*, 68-77.

Giachetta, R. (2015). A framework for processing large scale geospatial and remote sensing data in MapReduce environment. *Computers & Graphics*.

Herring, J. R. (Ed.). (2010). *OpenGIS Implementation Standard for Geographic Information: Simple Feature Access – Common Architecture* (version 1.2.1 ed.). Wayland, MA, USA: Open Geospatial Consortium.

Newell, R. G., Theriault, D., & Easterfield, M. (1992). Temporal GIS - Modeling the evolution of spatial data in time. *Computers & Geosciences*, 427-433.

O'Sullivan, B. (2009). Making sense of revision-control systems. *Communications of the ACM, 52*(9), 56-62.

Peuquet, D. J., & Duan, N. (1995). Peuquet, D. J.; Duan, N. *International journal of geographical information systems, 9*(1), 7-24.

Reuter, J., Hänßgen, S. U., Hunt, J. J., & Tichy, W. F. (1996). Distributed revision control via the World Wide Web. *Lecture Notes in Computer Science, 1167*, 166-174.

Ruparelia, N. B. (2010). The History of Version Control. *SIGSOFT Softw. Eng. Notes, 35*(1), 5-9.

Shen, H., & Sun, C. (2001). Operation-based revision control systems. *Proceedings of the 3rd Annual International Workshop on Collaborative Editing Systems.* New York: ACM.

Vrable, M., Savage, S., & Voelker, G. M. (2009). Cumulus: Filesystem backup to the cloud. *Trans. Storage, 5*(4), 14:1-14:28.

Wachowicz, M., & Healey, R. G. (1994). Towards temporality in GIS. *Innovations* *in GIS, 1*, 105-115.

Wang, J., Ghosh, R., & Das, S. (2010). A survey on sensor localization. *Journal of Control Theory and Applications, 8*(1), 2-11.

Winter, S., & Frank, A. (2000). Topology in Raster and Vector Representation. *GeoInformatica, 4*(1), 35-65.

Yang, C., Raskin, R., Goodchild, M., & Gahegan, M. (2010). Geospatial Cyberinfrastructure: Past, present and future. *Computers, Environment and Urban Systems, 34*(4), 264-277.

Zhang, J. (2010). Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion, 1*(1), 5-24.

Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory, 23*(3), 337-343.

Land cover validation game

Maria Antonia Brovelli[1], Irene Celino[2], Monia Elisa Molinari[1], Vijaycharan Venkatachalam[1]

[1] Politecnico di Milano, Como Campus, Via Valleggio 11, 22100, Como, Italy

[2] CEFRIEL – ICT Institute Politecnico di Milano, Via Fucini 2, 20133, Milano, Italy

Abstract

Land cover data constitutes highly useful information to monitor the extension and status of land resources, hence it has been realized how important it is to have accurate land cover data. Here, an interactive WebGIS is built in order to validate GlobeLand30 global land cover data. The Game with a Purpose (GWAP) human-based computation technique is adopted. The system is based on crowdsourcing, i.e. multiple users play the game to validate land cover classifications, thus increasing the confidence level of the validation.

Keywords

Land cover validation, WebGIS, Game With A Purpose, Crowdsourcing.

1 Introduction

Land cover data represents very valuable resource for many different studies related to the environment and sustainable development. However, the usage of this data in many applications cannot disregard its validation and the knowledge of its classification accuracy. Among the land cover validation techniques, Web-based applications are growing increasingly. The approaches proposed by Fritz et al. (2009) and Bastin, Buchanan, Beresford, Pekel, and Duboi (2013) represent notable examples. Geo-Wiki (Fritz et al., 2009) is certainly one of the most popular applications used for validating global land cover using crowdsourcing. The disagreement maps between different global land cover products are provided, where users validate these products using Google Earth. In the application of Bastin et al. (2013), authorized users have to visually assess land cover and may provide uncertainties information at various levels: from a general rating of their confidence to a quantification of the proportions of land-cover types within a reference area.

The objective of this work is to develop an interactive WebGIS for the crowsourcing-based validation of GlobeLand30, a new global land cover dataset at 30 meters resolution derived from the classification of Landsat (TM and ETM+) and HJ-1 satellites images according to the pixel-object-knowledge-based (POK-based) approach (Chen et al., 2014). The dataset has been produced by the Chinese government and released as open data in September 2014. It is available for the two baseline years of 2000 and 2010.

Within a research study aimed to evaluate the classification quality of GlobeLand30 on the Italian area (Brovelli, Molinari, Hussein, Chen & Li, 2015), this dataset has been compared with more accurate Italian land cover maps. Results show a degree of disagreement that ranges between 10% and 20%. For the non-coherent data another level of validation is needed.
The implemented WebGIS is intended to involve citizens to classify the non-coherent pixels. In comparison to Geo-Wiki, which makes use of Google Earth in its Web application for the validation purpose, we implemented a Web application which makes use of high resolution aerial photos in the form of a gaming environment which may attract many more citizens for the validation process. Currently, the area of Como municipality (Lombardy, Italy) has been considered for the WebGIS implementation.
In the following sections the classification quality of GlobeLand30 in Como region, the developed WebGIS characteristics and the implemented gaming approach are presented.

2 GlobeLand30 accuracy on Como areas

The assessment of classification accuracy of GlobeLand30 on the Como municipality area has been calculated by means of a comparison with DUSAF (“Destinazione d’Uso dei Suoli Agricoli e Forestali”), the Italian acronym for “Use Categories of Agricultural and Forest Soil” (Credali et al., 2011), the land cover data from Lombardy Region at scale 1:10'000. DUSAF is periodically updated through aerial photo interpretation integrated with regional databases information and it is currently available in five releases. Among these, the datasets referred to years 2000 and 2012 were selected for comparison with GlobeLand30 2000 and GlobeLand30 2010, respectively.
To enable the comparison, both datasets have been reclassified according to the first level of Corine Land Cover nomenclature, i.e. the hierarchical thematic legend adopted by DUSAF. As shown in Table 1, the degrees of disagreement between the datasets are coherent with those reported by Brovelli et al. (2015). The number of non-coherent pixels for the years 2000 and 2010 on which the validation process will be applied is equal to 17'773 and 18'510, respectively. Each pixel size is 30×30 m.

	Coherent pixels	**Non-coherent pixels**	**Overall Accuracy [%]**	**Disagreement [%]**
2000	66'701	17'773	79	21
2010	65'955	18'510	78	22

Table 1 - Comparison between GlobeLand30 and DUSAF on Como area: number of coherent/non-coherent pixels, overall accuracy and degree of disagreement.

The spatial distribution of non-coherent pixels (Figure 1) has been obtained by computing the difference between the compared datasets. These pixels will be validated by means of aerial photos at 0.5 m resolution provided by Blom CGR S.p.a. (http://www.blomasa.com/home.html)

Figure 1 - Distribution of non-coherent pixels (red areas) in the Como municipality area for year 2010.

3 WebGIS

A dynamic Web application was built which targets on crowdsourcing to validate the land cover dataset. Here PostGIS database has been integrated with Geoserver in order to store the aerial photos and other data. These data are made available via Web Map Service (WMS) in a browser-based interface by means of Angular JavaScript, Leaflet JavaScript library and server-side PHP code. The WebGIS consists of three panels, namely map panel, title panel and side panel. The pixel to be validated with aerial photo and five classification options are displayed in the map panel. The title panel is provided with a *Leaderboard* tab which displays the top three player positions and the player position of the current user, a *Best last Players* tab which shows the best players list according to the cumulative score achieved during the last ten game rounds, a *Badge list* tab which shows the list of badges that the player could win during the game play according to his performance, a *How to play* tab which shows a popup with instructions on how to play the game. The side panel consists of time counter, score board and the badge list.

The validation of land cover accounts for two different years (2000 and 2010). Two separate WebGIS were built for the years 2000 and 2010, respectively. In each of them non-coherent pixels (one at a time) are displayed to the user who has to choose a classification for them (Figure 2). Each non-coherent pixel is extracted randomly without following particular order. After submission of each answer, the user is displayed with another non-coherent pixel. The user is freely allowed to pan and zoom the map in order to correctly identify the pixel classification.

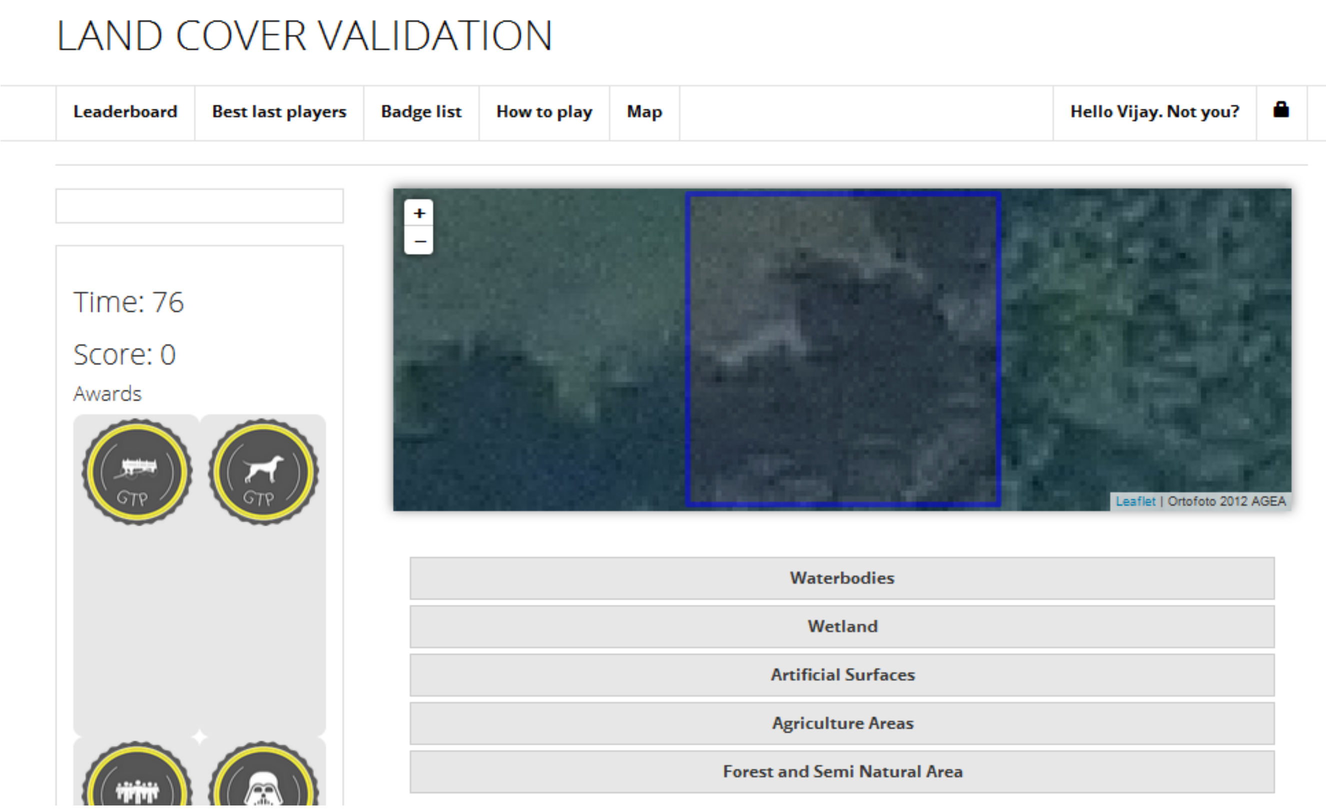


Figure 2 - Interactive Web application which displays a pixel (blue square box) for which user chooses a classification.

4 Gaming

The WebGIS is incorporated within a Game with a Purpose (GWAP) application, in line with the original proposal by Von Ahn (2006). A GWAP is a Human Computation application that puts a task in a crowdsourcing environment with a gaming flavor; the GWAP player has fun playing the game and, as a collateral effect, solves the Human Computation task. Here the purpose incorporated in our GWAP is land cover validation. This interactive game will engage the users to classify aerial images for their land use; as a consequence, the game will increase the quality of the land cover classification. For each pixel, the system knows the classified land use according to DUSAF and to GlobeLand30, which can be considered the “expert’s answers”; if the GWAP player’s answer matches an expert answer, the player will score a point in the game and will have fun. As incentives for the users to continue playing, we provide two mechanisms: on the one hand each game round gives some points, so a leader board is built to let the players compare their performance with other players; on the other hand, specific answer combinations assign players with badges that are aimed to make the game more attractive. Meanwhile, the system will collect the answers from multiple players; by cross-checking those players’ answers, the global land cover map will be validated by increasing the confidence about the correctness of the most “popular” answer.

5 Conclusions

The land cover validation game is a novel interactive system that allows citizen scientists to validate land cover maps while having fun. This game gives the right classification with a high level of confidence. This work is targeted at people who have a basic level of knowledge in image interpretation. Our purpose is to demonstrate that validation process could be increased by crowdsourcing. This game is a work in progress and we are currently studying how to assign proper scores to the players according to choices made on different land cover classifications. User information collected from this game will be stored in the database for further statistical evaluation to obtain the confidence level of the GlobeLand30 dataset.

Acknowledgement

The authors would like to thank Blom CGR S.p.a for providing the high resolution photos used for GlobeLand30 validation process.

References

- ✔ Bastin, L., Buchanan, G., Beresford, A., Pekel, J.F., & Dubois, G. (2013). Open-source mapping and services for Web-based land-cover validation. *Ecological Informatics*, 14, 9-16.
- ✔ Brovelli, M.A., Molinari, M.E., Hussein, E., Chen, J., & Li, R. (2015). The first comprehensive accuracy assessment of GlobeLand30 at a national level: methodology and results. *Remote Sensing*, 7, 4191-4212. doi:10.3390/rs70404191.
- ✔ Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., & Mills, J. (2014). Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS J. Photogram. Remote Sensing.* http://dx.doi.org/10.1016/j.isprsjprs.2014.09.002.
- ✔ Credali, M., Fasolini, D., Minnella, L., Pedrazzini, L., Peggion, M., & Pezzoli, S. (2011). Tools for territorial knowledge and government. In D. Fasolini, S. Pezzoli, V. M. Sale, M. Cesca, S. Coffani, & S. Brenna (Eds.), *Land cover changes in Lombardy over the last 50 years* (pp.17-19). Milano, ERSAF.
- ✔ Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., Obersteiner, M. (2009). Geo-Wiki.Org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1, 345-354. doi:10.3390/rs1030345.
- ✔ Von Ahn, L. (2006). Game with a Purpose. *IEEE Computer,* 39(6), 92-94.

FOSS4G-based energy management system for planning virtual power plants at the municipal scale

Luis Ramirez Camargo[1,2], Roland Zink[1], Wolfgang Dorner[1]

[1] Applied Energy Research group, Technologie Campus Freyung, Deggendorf Institute of Technology, Deggendorf, Germany

[2] Institute of Spatial Planning and Rural Development, University of Natural Resources and Life Sciences, Vienna, Austria

Abstract

The exploitation of renewable energy sources (RES) implies a transformation of the current energy supply infrastructure towards a spatially decentralized structure. This process has a significant impact on the use of space because energy is no longer generated far away from the demand, but it can be provided directly on-site. Hence, geographic information systems (GIS) have become a well established tool for the determination of sites for RES plants for heat and electricity generation. However, planning energy supply systems with a high share of RES, such as virtual power plants (VPPs), requires to consider also the temporal variation of the generation profiles of RES and the local energy demand. This paper presents the work in progress on a FOSS4G-based tool for the spatio-temporal modeling of municipality wide VPPs. The detailed simulation of potential RES-based supply alternatives and energy demand allows to identify suitable locations and to define the optimal size and technical characteristics of the individual RES energy generation plants in a VPP. This information can serve to assess technical renewable energy potentials at the municipal scale and to design a RES deployment roadmap customized to better fulfil the local energy demand.

Keywords

Integrated spatial and energy planning, renewable energy, virtual power plants

1 Introduction

A high share of renewable energy sources (RES) in the energy matrix will contribute to put an end on the dependency on fossil fuels and to reducing the amount of CO_2. In the case of the European Union the goal is set at 80% decrease in greenhouse gas emissions from the level of 1990, which would imply a share of 75% RES in the gross final energy consumption and 97% in the electricity consumption by 2050 (EC 2011). These ambitious objectives require a change in the energy generation paradigm away from centralized generation towards distributed generation.

With this transformation of the current energy supply infrastructure to a spatially decentralized structure two urgent problems arise: (1) wind and photovoltaics (PV) have a strong temporal variation and (2) total annual yield as well as the temporal variation strongly depend on the location of the plant. The temporal variation in power generation is not only a disadvantage of the meteorological influenced wind and PV technologies, but also applies for controllable combined heat and power (CHP) biomass plants, if the plants are driven by the heat production. Virtual power plants (VPP) were established as a concept to interconnect individual plants to compensate these spatio-temporal variations of one technology through variation of plants at different locations. Furthermore, VPPs compensate the deficits of one technology through the use of a portfolio of different technologies.

Research based on GIS technologies has already led to a number of concepts to identify optimal locations considering a maximisation of annual yield (Zink, 2010; Angelis-Dimakis et al., 2011; Calvert, Pearce & Mabee, 2013). However, there is still an open gap between the annual yield modelling of locations, using GIS, and the planning of VPPs, considering the strong regional and temporal variations of different technologies at a minute-by-minute or hourly basis.

2 Energy management system for planning virtual power plants at the municipal scale

The proposed tool consists of three modules (see Figure 1): (1) a series of GIS-based components, used to determine the technically usable potential for renewables within a municipality at a high spatio-temporal resolution (individiual installations of up to one square meter spatial resolution and hourly or intra-hourly time steps). (2) A GIS-based bottom-up model for estimating the energy demand for heating, water heating and electricity of buildings at a high temporal resolution (up to a quarter hours). (3) A newly developed VPP design module to select the plants configuration with the best match to the local energy demand. In addition, the required installed capacity of manageable RES, such as biomass, and the energy and power storage capacities, which are necessary to fulfil the demand of the local energy system is calculated. All modules are implemented in Python and rely on GRASS GIS, gdal, ogr, and Pktools.

The GIS-based modules to determine technical RES generation potential include the estimations of photovoltaic, solar thermal and wind energy potential. The potential estimation of the first two technologies relies on high resolution digital elevation models, vector data of the building footprints, global radiation, temperature data and technical parameters of common building-integrated photovoltaics and solar thermal plants. Calculation tools are the modules r.horizon and r.sun of GRASS GIS. The wind energy generation potential is assessed using a combination of reanalysis data and a GIS-based wind park location selection methodology.

The bottom-up model for estimating energy demand for heating, water heating and electricity demand of buildings requires georeferenced vector data in form of building’s footprints or, if available, 3D-buildings with level of detail 1 (LOD1) combined with historical data to match the individual buildings to building typologies. The typologies provide data for building`s envelope components quality that is necessary to run an adapted version of the resistance capacitance model defined in the EN ISO 13790:2008. This methodology is used for the calculation of the energy requirements for space conditioning. Further input parameters such as the dimensions of the building´s envelope components and the heating use areas are also estimated from the building´s footprints or retrieved from the LOD1 building models. The electricity demand of every building is defined with standard load profiles and the total electricity consumption per building is determined based on population data.

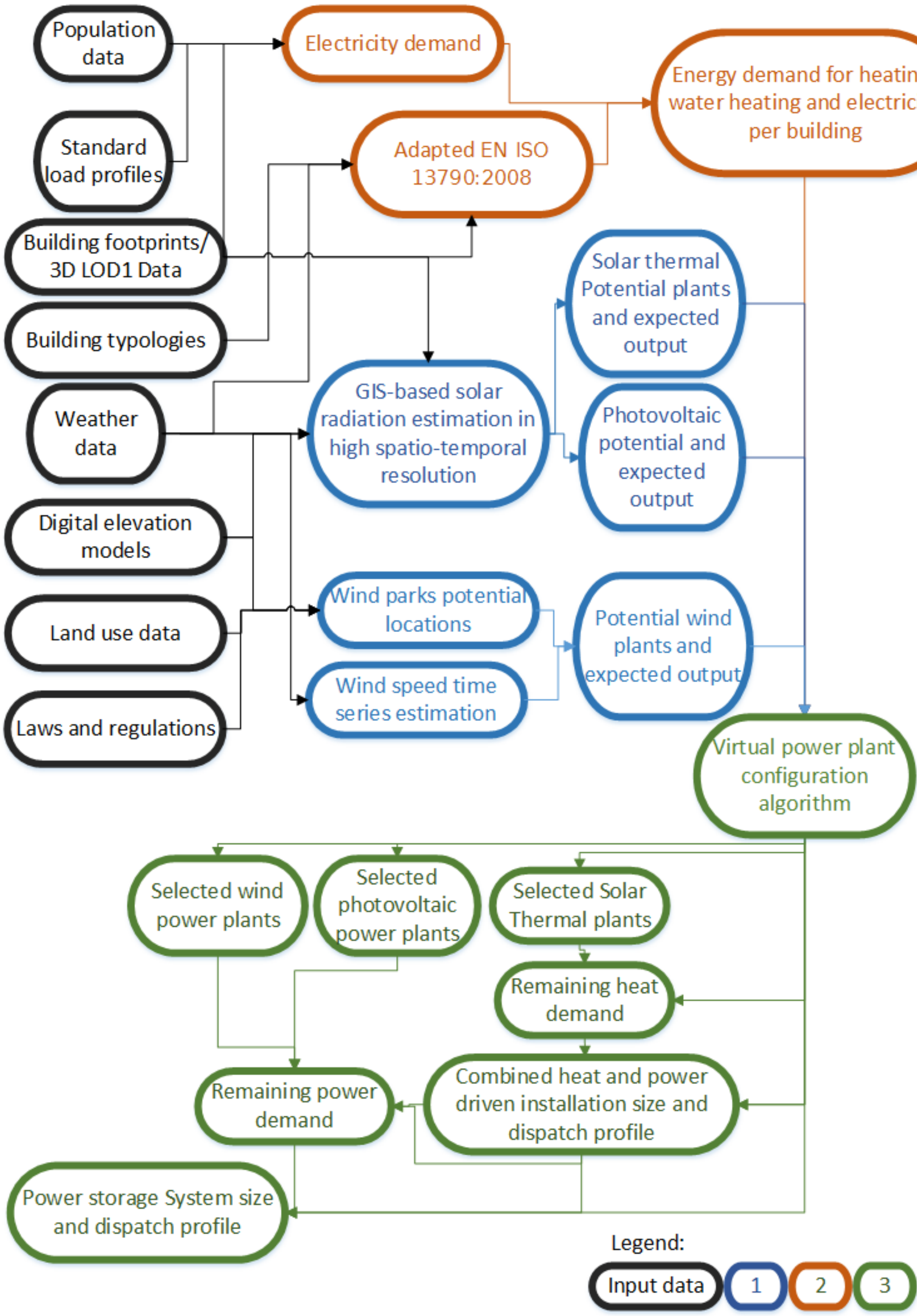


Figure 1 - Overall workflow of the tool.

The module to determine the VPP configuration searches for the combination of energy generation plants from fluctuating RES that has the best match to the local energy demand. The match to the demand is evaluated considering the amount of properly supply energy as well as the excess energy that every plant would contribute to the municipal energy system. The result is a data set representing the selected plants. Furthermore, dispatch profiles for biomass heat-driven CHP plants and storage systems are calculated in order to achieve a loss of power supply probability of zero. This means that the demand is fulfilled in all simulated time steps. The computed dispatch-profiles are the basis to calculate the CHPs and storage systems sizes. The algorithm based on a decision tree solves the problem firstly for every individual building for the heat and water heating demand and secondly municipality wide for the electricity demand.

3 Interim results

First results using the algorithm based on a decision tree approach to match supply and demand and the algorithm for sizing the storage system with data from a German rural municipality (Waldthurn, Bavaria) were presented in detail in Ramirez Camargo et al. (2015). A first VPP design calculated by the system for a part of Waldthurn in order to achieve a PV penetration rate of 40% of the yearly total demand is presented in Figure 2. This scenario was calculated for the electricity demand of 438 domestic buildings. Also the state of charge (SOC) of a storage system that does not allow any energy dumping and the state of charge of an optimally sized storage system were calculated (red and blue lines respectively). The SOC is based on the corresponding time series of the PV energy production minus the demand. The black and the dashed red colored lines in Figure 2 represent the power and energy capacities of the storage system that do not allow any energy dumping. The dashed yellow and dashed blue lines are the energy and power capacities of the optimally sized storage system.

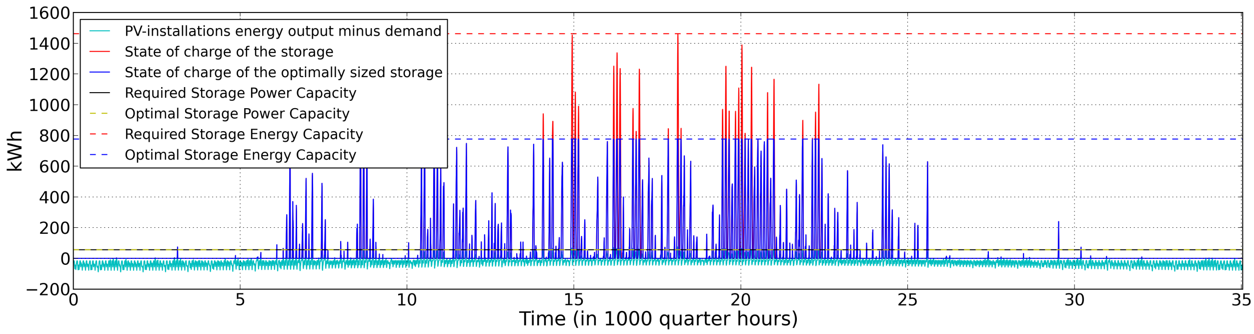


Figure 2 - Time series for selected PV-configuration for a penetration rate of 40% of the yearly demand.

A preliminary version of the GIS based bottom-up model for predicting the energy demand for heating is available in Ramirez Camargo (2012). This was implemented as a combination of Visual Basic for Applications (VBA) and the proprietary software ESRI ArcGIS. The current version that is presented here is programmed in Python and relies on GRASS GIS. Besides of the improvements in calculation time, the estimation of the solar radiation hitting the building`s envelope is improved by using the output of the r.sun module of GRASS GIS. Moreover, the model can work with LOD1 3D data and not only with German but also Austrian buildings typologies. Additionally, the calculation methodology is now compatible with German standard load profiles of electricity demand.

4 Future research

The proposed tool in its current status integrates relevant algorithms to determine renewable energy potentials as well as energy demand. The solar thermal potential and wind module as well as the sizing algorithm for the CHP plants are work in progress. The merge of all modules and the test of the tool in at least two case study municipalities, one in Germany and one in Austria, are the next milestones. The final objective is to deploy the tool for municipalities and regional suppliers allowing them to concretize a technically appropriate RES deployment roadmap and conceive itself as VPP. The fact that the tool relies on well established FOSS4G components is not only an advantage in terms of flexibility and computational efficiency, but should also facilitate its use and distribution.

References

- ✔ Angelis-Dimakis, A., Biberacher, M., Dominguez, J., Fiorese, G., Gadocha, S., Gnansounou, E., Guariso, G., Kartalidis, A., Panichelli, L., Pinedo, I. and Robba, M. (2011). Methods and tools to evaluate the availability of renewable energy sources. Renewable and Sustainable Energy Reviews, 15(2), 1182–1200. http://doi.org/10.1016/j.rser.2010.09.049
- ✔ Calvert, K., Pearce, J. M., & Mabee, W. E. (2013). Toward renewable energy geo-information infrastructures: Applications of GIScience and remote sensing that build institutional capacity. Renewable and Sustainable Energy Reviews, 18, 416–429. http://doi.org/10.1016/j.rser.2012.10.024
- ✔ European Commission (2011). Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions: Energy Roadmap 2050.
- ✔ Ramirez Camargo, L. (2012). A GIS-Based Method for Predicting Hourly Domestic Energy Need for Space Conditioning and Water Heating of Districts and Municipalities. Institut für Nachhaltige Wirtschaftsentwicklung, BOKU-Universität für Bodenkultur, Wien.
- ✔ Ramirez Camargo, L., Zink, R., Dorner, W., & Stoeglehner, G. (2015). Spatio-temporal modeling of roof-top photovoltaic panels for improved technical potential assessment and electricity peak load offsetting at the municipal scale. *Computers, Environment and Urban Systems*, *52*, 58–69. http://doi.org/10.1016/j.compenvurbsys.2015.03.002
- ✔ Zink, R. (2010). Einbindung von GIS in einen nachhaltigen Planungsprozess von Photovoltaik-Freiflächenanlagen auf kommunaler Ebene, in: Strobl, J.; Blaschke, Th. & Griesebner, G.: Angewandte Geoinformatik 2010. Berlin, p. 954-963.

Sharing geodata through university libraries: the case of Politecnico di Milano

Ilaria Muratori[1], Marcella Samakovlija[1]

[1] Politecnico di Milano. TeDOC - Area servizi bibliotecari di Ateneo (University Library Services)

Abstract

University libraries often have huge collections of cartographic documentation, both in traditional paper format and in complex digital format.
It's not only a matter of images or vectors, but also a problem of dealing with complete databases with thematic attributes and details.
In Italy, the libraries' systems of classification and research of these resources are still far to be well developed.
TeDOC, that has a large collection of historical and actual maps, both on paper and on digital format, and is responsible of the distribution of these resources to professors and students within teaching and research institutional activities of Politecnico di Milano, is developing a new approach to make these resources more accessible and easier available through the implementation of a geoportal.
The goal is to build a geolibrary, a library that puts at users disposal geoinformation and documents for which the primary search mechanism is place.
The issues faced within this project are related to metadata creation for different geospatial information, technical policy to assure protected access to resources with particular use constraints (e.g. resources available only for specific categories of users such as students), the opportunity to share resources within federated catalog both Italian and foreigners, and lastly to develop monitoring techniques to get statistical data about users.

Keywords

Metadata standard, sharing geodata, university geolibraries, geoportal

1 Which was the problem?

At TeDOC we have a large collection of digital and paper maps. These documents were searchable through the library catalogue and our web site[1].
In the recent years we became more and more aware that we needed to change the approach to our cartography conservation, organization, storage and accessibility.
The tools available to manage library documents at Politecnico, such as the discovery tool PRIMO and the catalog software Sebina, are not suitable to deal with cartographic documents. This problem is not due to a particular lack of those tools, but it rather depends on the specific characteristics of cartography, for which the best search method is place.
The search through the library catalogue has always been very difficult mostly because each cartographic series consists of numerous map sheets and the

[1] www.tedoc.polimi.it

user, without an appropriate geographic reference, is not able to identify which is the sheet that represents the portion of land of his interest.
On the other hand our old web site was organised into sections provided with map tiles of each cartographic series:

- For paper maps it worked with interactive tiles that brought to the catalogue record of the map containing the information with the borrowing details;
- For digital maps it provided static tiles containing the map coordinates and unique codes to make the request of the files.

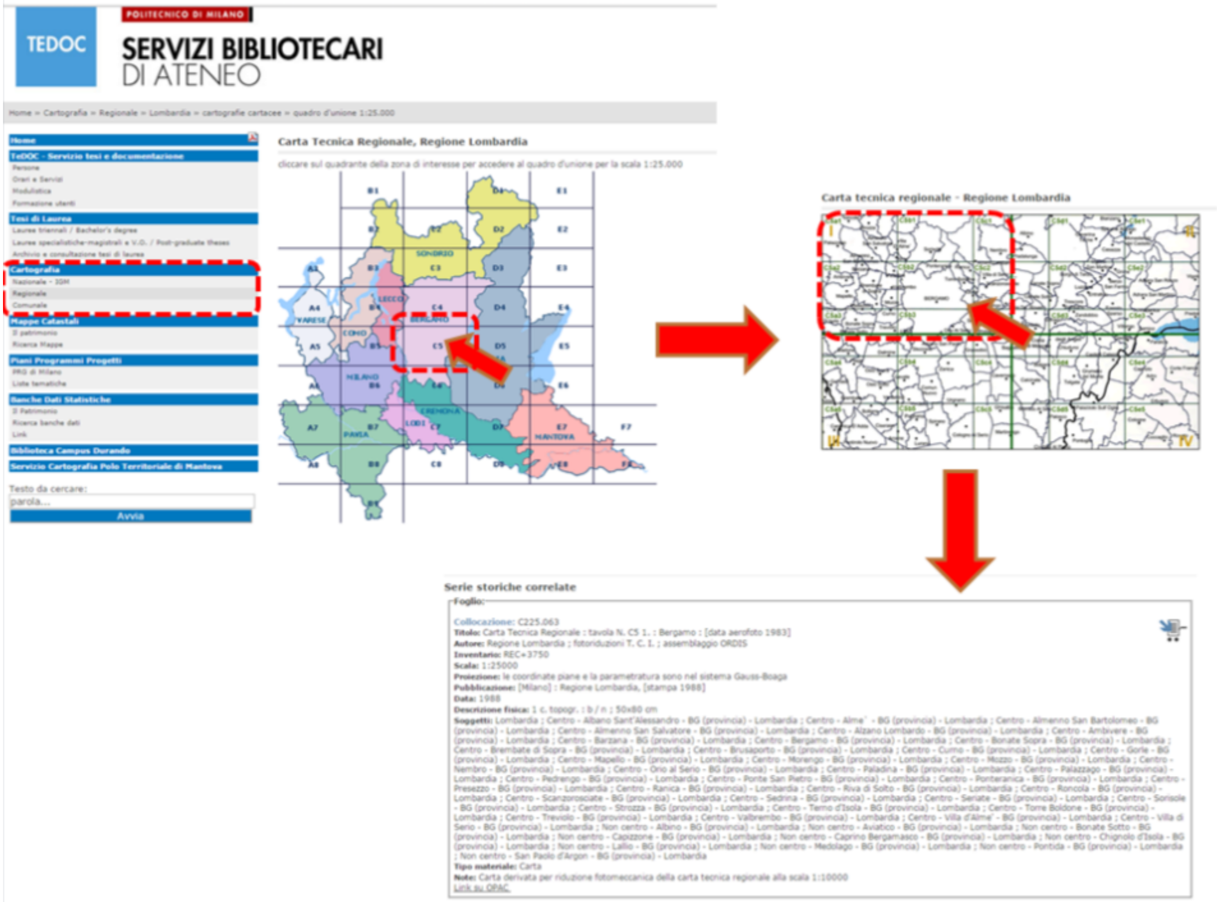


Figure 1 - TeDOC old web site: searching paper cartography through interactive tiles linked to library records.

Because map tiles are always different for each map series, the cartographic documents were divided into sections based on the owner/publisher, that in Italy usually corresponds to local administrations or regional agency, and in sub-sections based on the scale of the maps.
As you can imagine this system was quite difficult and put the user at the risk not to find what he was looking for, just because resources were too difficult to be found.

2 The technical choice - Why OGP?

After a deep analysis through several software and spatial search engines, our choice fell on the software open source OpenGeoportal[2]. The Open Geoportal (OGP) is a collaboratively developed, open source, federated web application to

[2] http://opengeoportal.org/

rapidly discover, preview, and retrieve geospatial data from multiple organizations. The project is lead by Tufts University along with Harvard and MIT. Several other partner organizations are assisting with the development[3].
Just because it has been developed by GIS librarians, this software is perfectly able to meet the need of university libraries in cartography management, catalogue issues, authentication and login needs and to dialog with the different tools in use in the libraries.

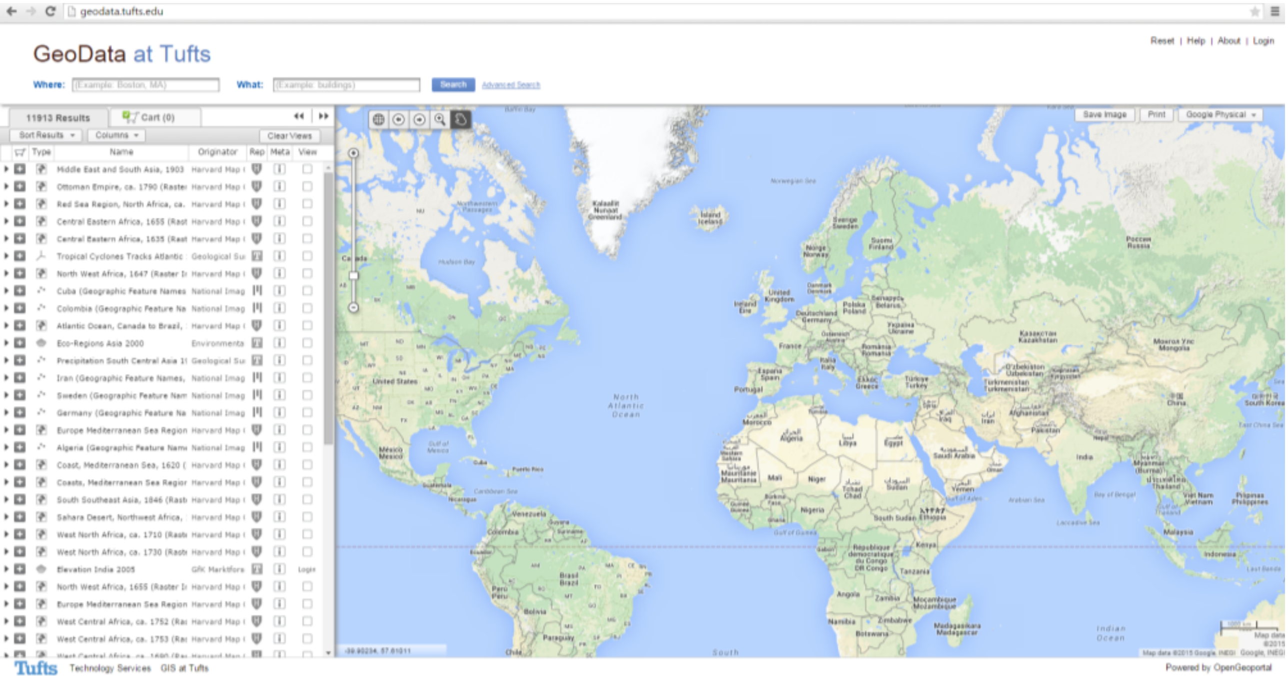


Figure 2 - GeoData@Tufts: Opengeoportal in production - http://geodata.tufts.edu/

3 The project

The project developed by TeDOC is supposed to solve several critical points, both of data search and of data delivery:

1. Difficulty by users in finding the cartographic resources needed;
2. Difficulty in delivery of digital geodata: an automatic download was not possible and the user had to come to our office to get the files;
3. Difficulty in producing statistical evaluations about resources utilization: any use survey had to be made by hand;
4. Need to find a unique method to catalogue and manage all kind of cartographic resources both digital and paper, raster and vector.

Thanks to the implementation of our geoportal based on the OGP web application the user can move and search into the different geographical areas in order to identify what are the maps available, regardless of the owner/publisher organization, or the data format, the scale of representation and the year of the survey. The geoportal provides results of the geographical search as a sort of map stratification or a core sampling of all the cartographic resources available for that land portion (e.g. for the territory of Milan there are municipal base maps at a very large scale, regional and national scale maps, cadastral maps, urban plans, historical maps, etc.). Furthermore it provides the

[3] Examples of OpenGeoportal in production: Harvard (http://calvert.hul.harvard.edu:8080/opengeoportal/), University of Wisconsin-Madison (http://maps.sco.wisc.edu/opengeoportal/), University of California, Berkeley (http://gis.lib.berkeley.edu:9081/)

download function both for selected metadata and files.
It is important to note that the implementation of this project involves, in addition to TeDOC, also the ICT Area, with informatics tasks such as the installation and customization of the software, and the Geodesy and Geomatics office of the Department of Civil and Environmental Engineering (DICA) for some technical-methodological support.

3.1 Analysis and re-organization of TeDOC cartographic resources

First of all, since the TeDOC cartographic collections are very different, we started with deep analysis of the digital collections to check the quality of metadata (if present) and, on the other hand, the methods of distribution to the users of each collections.
For example in the case of the Regional Base Map produced by Regione Lombardia, the geodata (in shp format) are organized into thematic groups and distributed to users in a compressed zip format, not in single layers. This is because each theme, such as for example the land use theme, is divided into many layers.

ente	DESCRIZIONE	codice carta	scala	formato	note
Comune di Milano	Carta Tecnica Comunale (2006)	0101	1:1.000	jpg	Vedi quadro d'unione
	Carta Tecnica Comunale (2006)	0102	1:1.000	dwg	Vedi quadro d'unione
	Carta Tecnica Comunale (2004) Altimetrie	0103	1:1.000	tiff	Vedi quadro d'unione
	Carta Tecnica Comunale (2004) Pertinenze	0104	1:1.000	tiff	Vedi quadro d'unione
	Carta Tecnica Comunale (1990)	0105	1:2.000	tiff	Vedi quadro d'unione
Comune di Cesano Maderno	Carta Tecnica Comunale (2005)	1701	1:2.000	dwg	Vedi quadro d'unione
	Carta Tecnica Comunale (2005)	1702	1:5.000	dwg	Intero territorio comunale
Comune di Cremona	Carta Tecnica Comunale (2003)	0201	1:1.000	shp	Intero territorio comunale
	Carta Tecnica Comunale (2003)	0202	1:1.000	dwg	Intero territorio comunale
Comune di Gessate	Carta Tecnica Comunale (2008)	0301	1:1.000	shp	Intero territorio comunale
Comune di Giussano	Carta Tecnica Comunale (1982)	1501	1:5.000	tiff	Intero territorio comunale
	Carta Tecnica Comunale (2012)	1502	1:1.000	dwg	Intero territorio comunale
	Carta Tecnica Comunale (2012)	1503	1:1.000	shp	Intero territorio comunale
Comune di Monza	Carta Tecnica Comunale (1993)	0401	1:2.000	dwg	Vedi quadro d'unione
	Carta Tecnica Comunale (2005)	0402	1:2.000	shp	Intero territorio comunale
	Carta Tecnica Comunale (2005)	0403	1:2.000	dwg	Vedi quadro d'unione
	Ortofoto (2005)	0404	1:2.000	tiff	Vedi quadro d'unione
Comune di Novate Milanese	Carta Tecnica Comunale (2010)	1601	1:1.000	dwg	Intero territorio comunale
	Carta Tecnica Comunale (2008)	1602	1:1.000	shp	Intero territorio comunale
Comune di Pavia	Carta Tecnica Comunale (2001)	0501	1:2.000	dwg	Vedi quadro d'unione
Comune di Rho	Carta Tecnica Comunale (1996)	0601	1:2.000	dwg	Vedi quadro d'unione
Comune di Sesto San Giovanni	Carta Tecnica Comunale (2001)	0701	1:2.000	dwg	Vedi quadro d'unione
	Aerofoto (2001)	0702	1:2.000	jpg	Vedi quadro d'unione
	Carta Tecnica Regionale, volo 1980-91	0801	1:10.000	tiff	Vedi quadro d'unione
	Carta Tecnica Regionale, volo 1994	0802	1:10.000	tiff	Vedi quadro d'unione
	Carta Tecnica Regionale, volo 1994	0803	1:25.000	tiff	Vedi quadro d'unione
	Carta Tecnica Regionale, volo 1994	0804	1:50.000	tiff	Vedi quadro d'unione
	Base Dati Geografica CT50 (a colori), volo 1994	0805	1:50.000	tiff	Vedi quadro d'unione
	Modello digitale del terreno DTM20, volo 1981-83	0806	1:10.000	grid-ascii	Vedi quadro d'unione
	Base Dati Geografica - CT10	0807	1:10.000	shp	Scegliere tematismi
	Basi Informative Ambientali della Pianura	0808	1:25.000	shp	Scegliere tematismi
	Base Dati Geografica di Sintesi	0809	1:250.000	shp	Scegliere tematismi

Figure 3 - Re-organization of TeDOC cartographic resources: extract of cartographic digital collections

The single layer is not significant in an analysis of the territory and, for the user, finding all the layers that belong to a theme would be quite difficult. That's why we decided to maintain the same kind of aggregation even within the geoportal: actually the various collections, or cartographic series, are

organized into general themes (e.g. land use, transportation infrastructure, administrative units, etc. ...) in order to make it easier to select and download the data.
Instead, the raster data, that are usually divided into several tiles, are ingested into the geoportal one by one so that the user can identify exactly which is the area represented by the map that he is selecting and that he may download.
Making our choices we have considered our experience gained through years and years of observation of our users who often have no expertise in GIS.
For what paper maps are concerned, they are registered in the University Library Catalogue and we adopt as metadata the library record of each map.

3.2 Metadata standards. A solution for cataloguing issues

The implementation of metadata is critical, as metadata provide valuable information about the data itself. In fact collecting, creating and organizing the metadata for our resources has been the most time-consuming and labor-intensive process we had to face.
We chose to follow international standard to guarantee time preservation and exchange opportunities.
The largest part of our digital cartographic resources had original metadata with very different levels of quality. For example "historical" datasets are given (by the owner/publisher institution) with incomplete or not precise metadata that need to be re-compiled almost completely. Instead more actual datasets have more comprehensive metadata that are filled according to European standards.
This fact can be explained by the delay, typical of Italy, in conforming to the international regulations, and mainly by the real difficulties encountered by the institutions to go back in time to create metadata for old geodata that were published when still the concern about metadata was not yet a critical issue.
And since our idea was to create a geoportal that could be shared with other universities, including foreigners, and at the same time to build a web application that can keep together (read, search) resources held even by other universities, it has been necessary to make a comparison of the different standards in use. We of course decided to provide bilingual metadata for all our resources.
It was therefore carried out an analysis of the metadata standards in use, of the information that they must include and the specific ways in which this information should be expressed. The starting point for this analysis was the INSPIRE Directive, not only for its prescriptive nature, but mainly because it introduces the idea of building a transnational structure for geospatial information and thus the need to catalog this information according to specific criteria that allow to go back over the path of the data history and that make the search process the easiest possible.
As known, the standard in force for Europe is ISO19139, while in the US it is in use the standard FGDC (but debates are ongoing on whether to move to ISO19139 which allows a more detailed description of the geographic data).
To follow the standard, we listed the following elements to be reviewed:

- keywords
- abstract
- lineage
- purposes
- data format

- use constraints

Our job was carried out following the OpenGeoPortal Metadata Working Group, a multi-institutional partnership of data and metadata experts, who are seeking to develop a set of common practices for creating and exchanging geospatial metadata.

Included in these best practices there are recommendations for the use of keyword and thesaurus, outlining encoding standards for the construction of free text fields, and methodologies about describing and relating metadata of collections to layers' metadata.

This is very important since our future target is to federate our geoportal with the others existing university geoportals to share geodata.

In a first step we decided to create the metadata template for some cartographic series. In particular, we intervened on the maps of Milano Municipality and Regione Lombardia, which are the institutions with whom we have the more frequent and consolidated relationships. For a better result in the template creation we consulted the cartographic departments of the two institutions, either to ensure the correctness of the information or to give a feedback information to the institutions themselves that may use our support in their processes of metadata review and creation.

Our target for this step was to define a complete procedure to be applied for the creation of the metadata for the wholeness of our cartographic resources.

The metadata template is filled with the basic info that are the same for all the elements of a specific cartographic collection.

The fields that must filled up creating the metadata template are:

1. contact information: Responsible Party (owner, publisher, resource provider, point of contact...)
2. series name
3. collective title
4. abstract
5. keywords and thesaurus name
6. lineage
7. use limitations
8. other constraints
9. distribution format

The field "title" has a part in common with all the maps of a specific cartographic series, but it usually contains a specification related to the single map tile or theme.

To create the metadata template we used ArcCatalog 10.3, then editing it with notepad++ if necessary. Afterwards, we passed through the metadata validation using the translator ARCGIS2ISO19139.xml

3.3 Procedures for Automatic Metadata creation

After having developed the metadata template we proceeded with metadata import for each element of each cartographic series.

First of all, it is necessary to project the element in its own coordinate system. A pyton script, designed at this purpose, run the projection command for all files of a particular series.

Afterwards, using another pyton script, we run the import metadata command for every file. At this point it is necessary to make some manual, or semi-automatic, changes to modify the fields "title" and "file identifier". We underline the importance of choosing a file identifier that can assure the

unique identification of the resource now and in the future.
At the end of this procedure we perform the validation (in batch) of all the metadata in the standard ISO 19139.
Finally the validated metadata can be ingested in OGP with the OGPingest tool that perform another check of the conformity of metadata to the standard.

3.4 Paper maps and geoportal: which solutions?

The TeDOC, as already mentioned, has a large collection of paper maps that have been cataloged according to the standard UNIMARC.
Our project wants to connect the catalog with the TeDOC geoportal to allow the search of paper maps together with the digital ones.
To do this, it is necessary to update all the library records by entering geographic coordinates in degrees and the Prime Meridian.
Our work is not yet complete. For the moment we started working on historical maps produced by IGM (Istituto Geografico Militare). We studied the procedures to derive automatically the coordinates of each map sheets from the tilesheet in shp format. Subsequently, we created a DB containing, for each library record, the coordinates of the bounding box and the permalink to the resource.
In the near future it should be possible, from that DB, to perform an automatic update of the catalog by entering the coordinates of the bounding box of each element in the library record. Afterward we will perform an extraction of library records in MARC format (compatible with our system) and, thanks to customization of the OGP made by MIT staff, it should be possible for us to ingest our MARC records into the SOLr OGP index.

4 Conclusion

The TeDOC geoportal, currently in test version, does not yet provide preview and download functions. We still have a lot of geodata to ingest and we are waiting to be able to ingest also all the paper maps. Anyway, since January 2015 the geoportal can be used by Politecnico students to make their search among the cartographic collections already ingested.

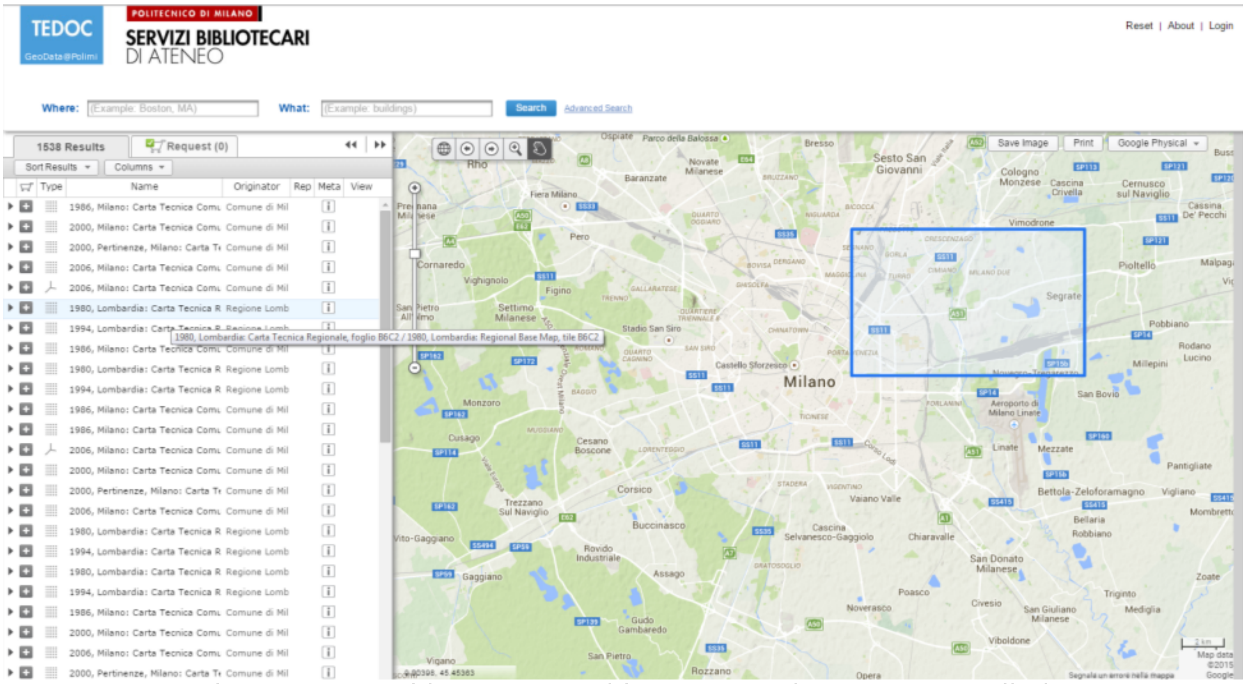


Figure 4 - Searching cartographic resources in GeoData@polimi

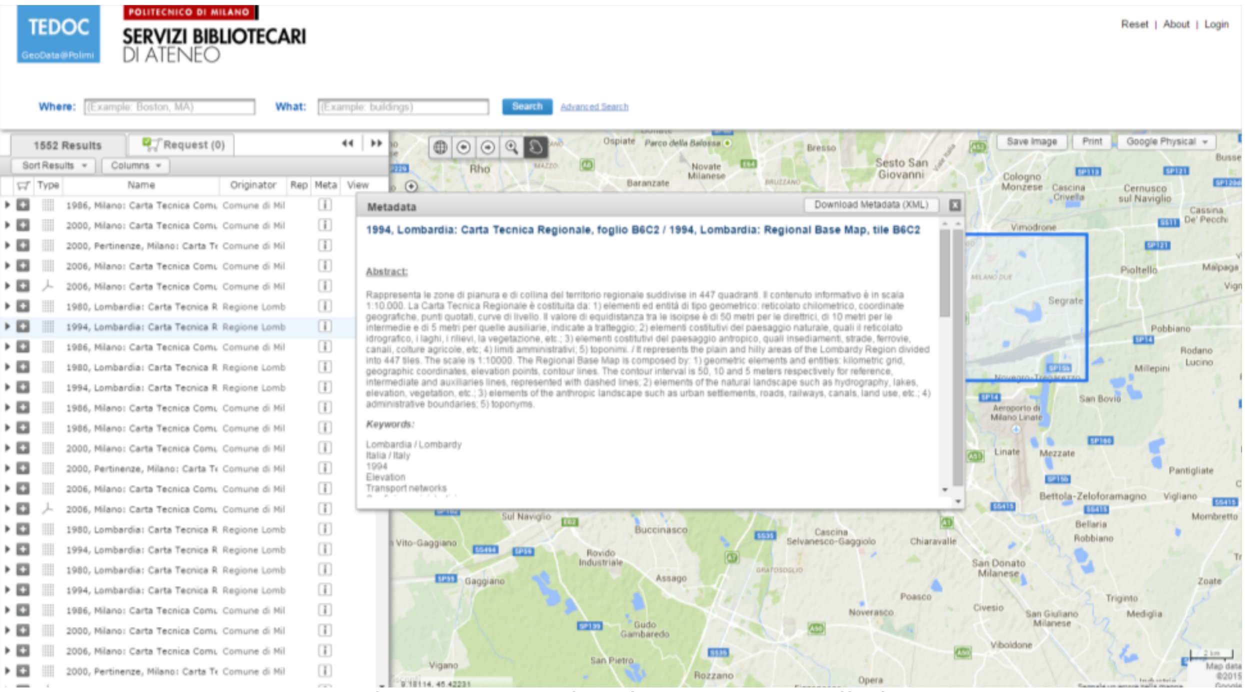


Figure 5 - Metadata in GeoData@polimi

In the next steps, the project will provide the opportunity to sharing the cartographic resources with the other university that are currently in the OGP group with different level of accessibility depending on the use constraints of the documents.

We finally want to note that during the last decade, while users demand of geodata was increasing, also the web availability of these information has greatly increased, that's why the use of a federated "geo-catalogue" will be a great opportunity for the diffusion and sharing information both directly owned by Politecnico and by other University in the world.

Politecnico users will be able to find everything about a geographic area just with a click of the mouse.
We hope that in the short term other Italian and European universities, in addition to those already operating overseas, will join this project to allow an ever wider sharing of knowledge and geospatial information.

References

✔Boxall, J. (2003). Geolibraries: geographers, librarians and spatial collaboration. The Canadian Geographer/Le Géographe canadien, 47(1), 18-27.
✔Donnelly, F. P. (2010). Evaluating open source GIS for libraries. Library Hi Tech, 28(1), 131-151.
✔Durante, K., & Wang, Z. (2012). Creating an actionable assessment framework for discovery services in academic libraries. College & Undergraduate Libraries, 19(2-4), 215-228.
✔Florance, P. (2006). GIS collection development within an academic library. Library trends, 55(2), 222-235.
✔Goodchild, M. F. (2000). Communicating geographic information in a digital age. Annals of the Association of American Geographers, 90(2), 344-355.
✔Hardy, D., & Durante, K. (2014). A Metadata Schema for Geospatial Resource Discovery Use Cases. Code4lib Journal, 25.

Translating urban history, research and sources, into interactive digital libraries

Fulvio Rinaudo[1], Rosa Tamborrino[1]

[1] Politecnico di Torino – Dept. Of Architecture & Design

Abstract

Many results of historical studies regarding Cultural Heritage assets are still recorded and stored into traditional books and reports. They quote essential documents, iconography and archival sources. The efforts done in the last years by digitizing libraries and archives contents represent just a first step towards a digital communication of historical research results. The challenge is in looking for a new digital access to outcomes linking also to the related sources.

The use of GIS technology and 3D modelling allow a new way to share historical information among the specialists who need to access those information to document, preserve and manage Cultural Heritage assets.

In particular basic historical information about urban centres can be archived and transmitted by using shape files, 3D models by using the most advanced *3Dpdf* format and complex historical research results connected and presented by using hyper texts which can be easily published as websites.

All these platforms produce critical readings linked to digital libraries where the users are able to access the data but also to integrate those data in digital platform.

Keywords

Digital History, GIS, 3D modeling, cities, urban spaces, digital libraries

1 Introduction

The definition of digital library is evolving from the old meaning of electronic or virtual library as a set of digitized documents and collections.

Today this simple concept can be extended by considering that digital libraries are not single entities, therefore they require high technological solutions to link the resources of many but preserving the transparency of these links to final users.

More advanced definitions are directed to guarantee a universal access to digital libraries as information services and to overpass the concept that digital libraries have to be realized as surrogates of paper documents: they have to be extended to digital artefacts that cannot be represented or distributed in printed formats (Stern, 2014).

By following the modern trends of digitization of all the information in order to open towards a better interoperability of the data among different specialists also the data concerning the documentation of Cultural Heritage assets have to fit this topic.

While the results of metric survey are today completely developed, stored and documented (by means of metadata information) in a digital way, up to now the results of historical research related to the comprehension of the importance, from a cultural point of view, of an asset (both tangible and intangible one), its documentation and the interpretation of the related sources, have been stored and transmitted by using the same media that has been used for the last centuries since the invention of the printing such as printed books, reports, etc.

In the last decades a lot of efforts have been devoted to a wide digitization of archives and libraries but this action cannot be interpreted as a ~~real~~ full digitization of written text and documents.

The realized digitization just prevent the losses and damages of the original documents but do not allow a real digital "translation" of the contents.

By considering the built environment as a space where a large part of the cultural assets are, the historical research on architecture and on cities is very important in order to share new developments in the documentation and advance in the understanding. This kind of historical approach is quite peculiar because the time is strictly linkable to the space. In the making history of architecture almost all the information used to understand and describe a specific asset can be easily referred to a geographic reference system and therefore the use of digital and interactive thematic maps can be one of the most intriguing way to share this kind of information: this is the case, for example, of huge cataloguing works performed in the past about architect's ideas and designs. Another important aspect of history researches is the study of the development of urban centres by considering the relationships between historical events and effective architectural and urban transformations.

A more complex kind of historical studies refer to historical reconstruction of a life cycle of a specific building or asset by using original documents and original analysis performed by a specialist. In this case hypertext (e.g. website) can represent an alternative way to allow the access both to the conceptual processing of the data and the original data themselves.

2 Effective digital conversion of historical data

Every book concerning a particular argument of the history of architecture contains many information that can be extracted and used for different aims or just to explore a specific aspect of the treated argument.

All those data can be grouped in the following main categories:

- Geodata: data which can be located in a specific cartographic reference system (both historical and modern ones);
- Historical sources: digitized documents and iconographies usually stored in public libraries and/or Archives and museums;
- 3D models: virtual 3D reconstruction of existing or not existing architectures.

For each of these categories specific solutions can be adopted by preserving the principle of interoperability and reuse of data strongly underlined and promoted by all the scientific communities and finally stated at European level by INSPIRE directive (also adopted by a lot of extra European Countries).

2.1 Historical geodata

By considering urban history, a huge amount of data used by historians (especially architectural historians) to support their studies can be located in a specific point of the Earth's surface or in a specific point of a local system (e.g.

when referring to a single building).
Some of them find a correct location on historical maps with different accuracy properties than modern digital map but in any case they can be located by using the traditional vector elements used in a traditional GIS layer (point, line or polygon).
Often historical maps can linked to modern reference and coordinate systems (e.g. UTM/WGS84) or by using rigorous transformations (when all the geodetic data concerning the historical map are known like in case of maps dating from XIX century) or by means of local transformations when some of the points of the historical maps can be easily identified in a modern map or on the ground.
In this last case GNSS techniques allow to define modern coordinates of the homologous points and then estimate the transformation parameters.
Each of these possible solutions has to be discolsed in such a way that all ~~the~~ users can judge in an indipendent objective way the quality of the adopted geo-referencing strategy.
In the aim of a rigorous translation the authors have to declare the basic information: use of official data released by official cartographic bodies, use of specific software or the values of the parameters used and the transformation (e.g. Molodensky model, etc.).
In case of local transformation a report about the points used and the coordinates of those points in both the coordinate systems and the residuals on the used points is required.

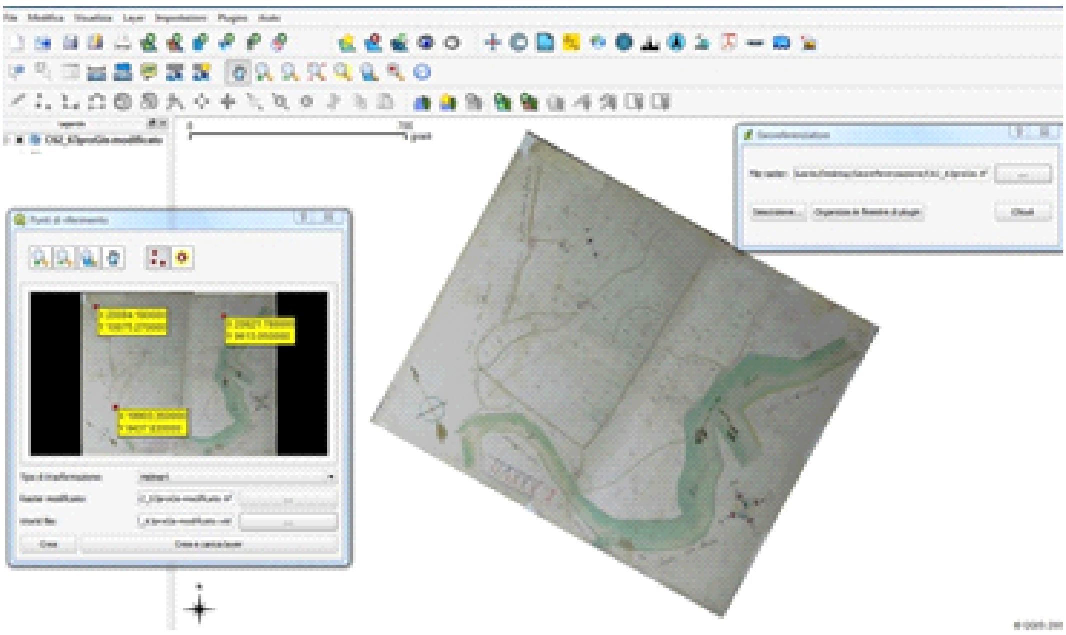


Figure 1 – Georeferencing of an historical map

Once the reference and coordinate systems have been defined for a specific map it is possible to insert the non-cartographic information by using one of the possible vector elements.
Obviously the vector element to be used depends on the scale of the map and on the detail of the data to be inserted. In case of traditional and historical maps those elements (points, lines or polygons) must be generated by means of a digitizing process which requires a correct interpretation of the historical

map.
The selected data are than stored as attributes and all the information saved in *shapefile* format in order to be accessible independently from the used software.
To be really accessible all the contents of the recorded data a complete and exhaustive information about the distributed data has to be annexed.
The organized historical information is the division among a traditional digitisation of documents and a digital library. The survey provides data to be processed, the historical approach provides the interpretation ranking data in layers.
The main information to be provided are:

- Metric contents:
  - datum and coordinate system (if known);
  - residuals on referencing points (in case of use of local transformations);
  - nominal scale.

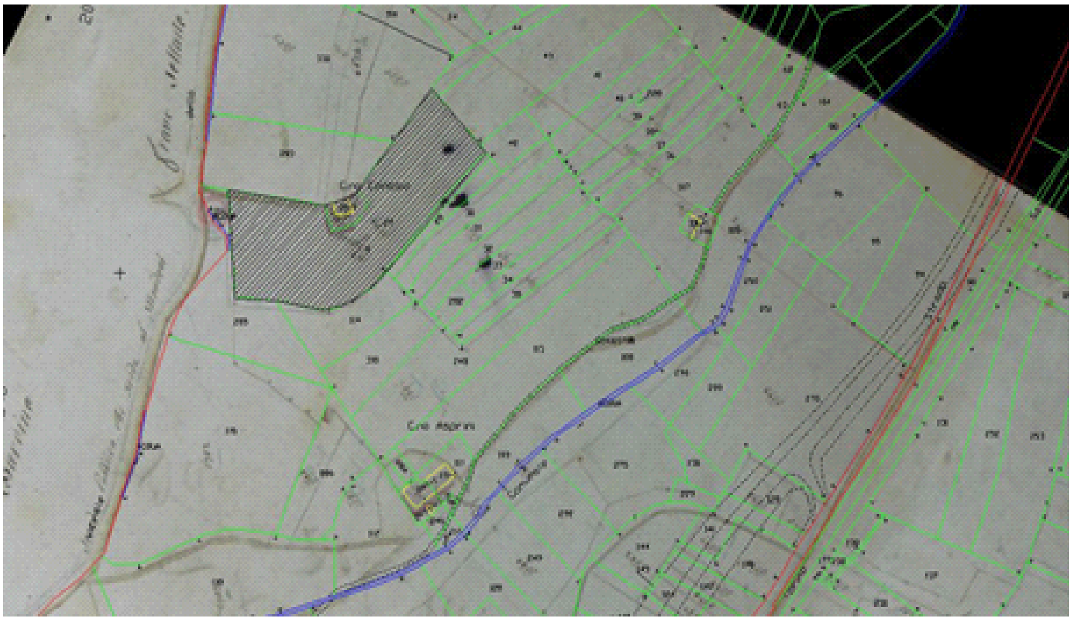


Figure 2 – Digitization of a raster map

- Attributes:
  - clear definition of the content
  - type (number, string, etc.)
  - glossary (in case of strings a clear definition of each term has to be given)

2.2 Historical sources

All historical researches are based on original documents usually stored inside Libraries and/or Archives or Museums.
In a traditional book they are usually cited in terms of footnotes and is up to the reader to check, in case of interest, the original documents which justify the historical reconstruction of an event.
In case of a digital “translation” of a book, the challenge is to produce a new access to the outcomes and to the sources. Translating research in a digital environment is quite different than produce a digitisation of a book or a report.

The full use of the digital implication boosts to change our mentality in communicating the research as an open data. By considering the outcomes as a step of the knowledge to share, the digital tool asked to be conceived as useful to check the evidence of the results by accessing sources also able to accept integration and new implementations. it is possible to conceive a direct link to the original documents by considering the two possible cases.
If the cited document is present in a digital archives directly accessible from the network a simple link allow the user to access the original information; if the document is not yet digitized or not accessible through the network a digital copy of it can be published by using mainly the *pdf* format easily visible from every kind of digital platform.
In this case the metadata have to be defined by the hosting digital library or Archive. Only in the case of new digitization some basic information have to be connected to the digitized document (e.g. author, resolution, copyright, etc.).

2.3 3D models

A lot of books on history of architecture use 3D representations (e.g. axonometric or perspective views) to show particular details of interest to demonstrate in a graphic way ideas, interpretations and designs.
A simple scan of these drawings cannot be considered a real digitization of them.
The ideal solution can be the realization of complete 3D models by using one of the useful packages (e.g. AUTOCAD, Revit, Rhinoceros, etc.) and to record them in *3Dpdf* format in order to ease the users to see the different logical layers of the 3D models, to rotate them and to extract the 3D metric information usually lost in every sort of 2D reproduction of specific views of the realized 3D model.
This is the most expensive digitization process of historical data and must be performed in such a way that the 3D model can really transmit the contents useful to understand what the author need to show to support his ideas.
The use of layer structure have to be studied carefully in order to allow the user to see the object in all the aspects required by the aims of the book to be digitized.
For those digital documents some information have to be delivered to easy the user to understand and properly manage the data.
Each layer have to be defined in terms of contents and in terms of origin (drawings, surveys, interpretation) and accuracy of the data by considering that each layer have to contain only single origin and single accuracy data.
These tools allow a visual translation of all data collected, by creating digital archives and especially by representing synthetic information.

3 A practical example: Alessandro Antonelli’s designs

In the following an example about a possible translation of a book of history of architecture is explained.
The considered book (Rosso, 1989) is a catalogue about the designs of Alessandro Antonelli, one of the well-known Italian architects of the XIX century: he designed the Mole Antonelliana the symbol of the city of Turin.

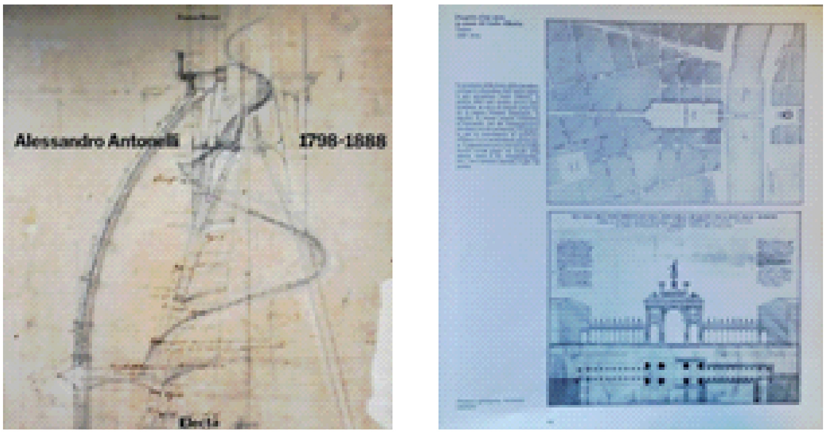


Figure 3 – The "Alessandro Antonelli 1798 – 1888" book cover and a sample of the catalog of the Antonelli's design

From this book the complete list of the Antonelli's designs was extracted and located on a digital maps by considering different scales and therefore different set of attributes for national list, regional list and municipality list of the analysed designs.
As an example the following figure shows the list of the Antonelli's designs at different scales. For each Italian region only the number of designs were recorded (due to the considered scale 1:1000000) while for the municipality of Turin (due to the considered scale 1:1000) more detailed information was extracted: e.g. year of the design, name of the object, property (religious, private, public), realization (yes or not), presence.
The *shp* files can be downloaded and opened by using a GIS platform (e.g. QGis). At municipality level hyperlinks allow to access, for each design, the original drawings, the 3D models realized by using Revit software and published as *3D pdf* format and in some cases also some video showing the Antonelli's designs in the existing urban context (both at the time of the design and at present times).
The books could contain several reproduction of drawings but the most of the data about drawings and documents were just noted in records and footnotes. Sources were preserved in different archives, such as the municipal archives, archives private or public related to the patronage, archive in the library of the school Academia Albertina collection of the Galleroa d'arte Moderna where the personal archive of the architect is kipped.

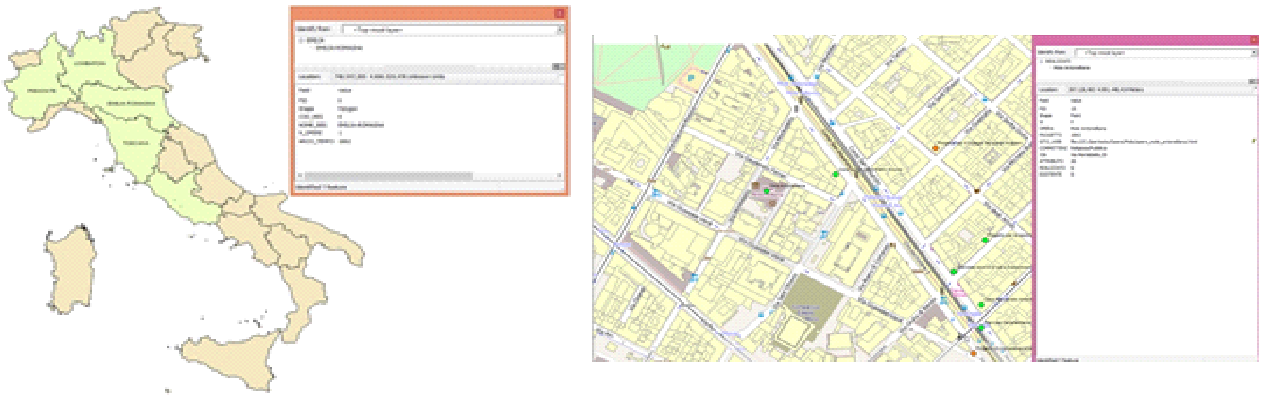


Figure 4 – Italian regions were Alessandro Antonelli conceive his designs (left), location of the Antonelli's designs at city scale (right). Beside the thematic maps generated by using a GIS platform the list of the attributes of the selected object

All the files could be stored as single elements inside a digital library by using indexes and all the metadata needed to help the search but a possible alternative solution is the use of an intermediate media such as an hypertext or a website able to drive the user through a correct use and comprehension of the produced documents.

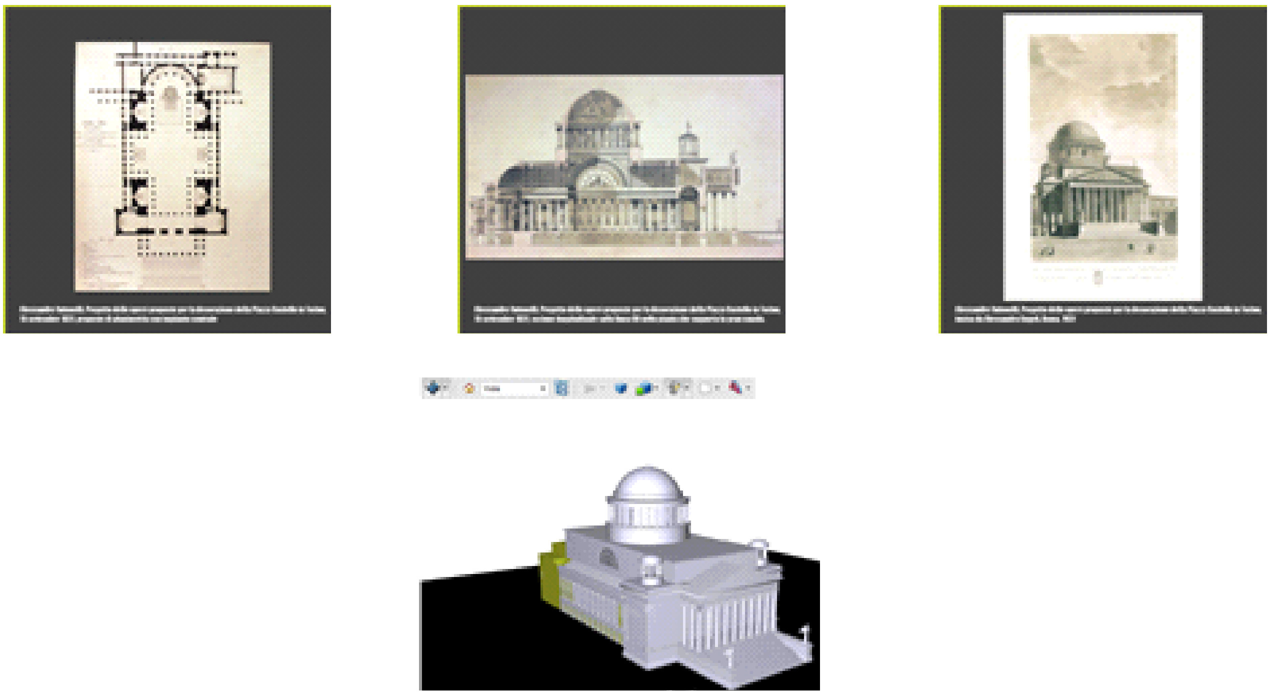


Figure 5 – Original design drawings and resulting 3D model on 3Dpdf file.

This hypertext can contains texts and comments, active buttons to allow the direct downloading of the *shapefiles*, *3Dpdf*, scans of the used original documents, video and related ancillary information.

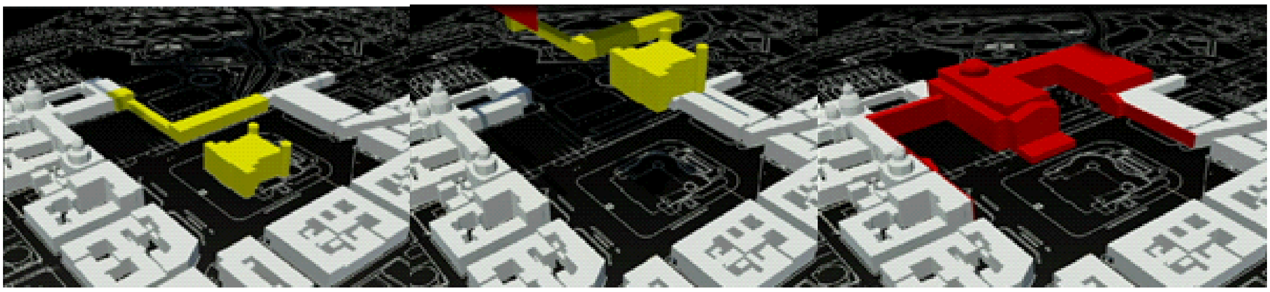


Figure 6 – Frames of a video showing the modification of the Piazza Castello in Turin due to a not realized design of Alessandro Antonelli

4 Conclusions

The data used for historical researches on the history of cities can be used in different situations and for different aims.
The development of the concept of digital libraries allows to share those data organised by the historical research and visualised in digital form by preserving the quality of the data themselves and by ensuring a real complete access to them.
The efforts to be done to reach this goal are not negligible, therefore historians have to start to collect and manage the basic data and their subsequent elaborations in digital form.

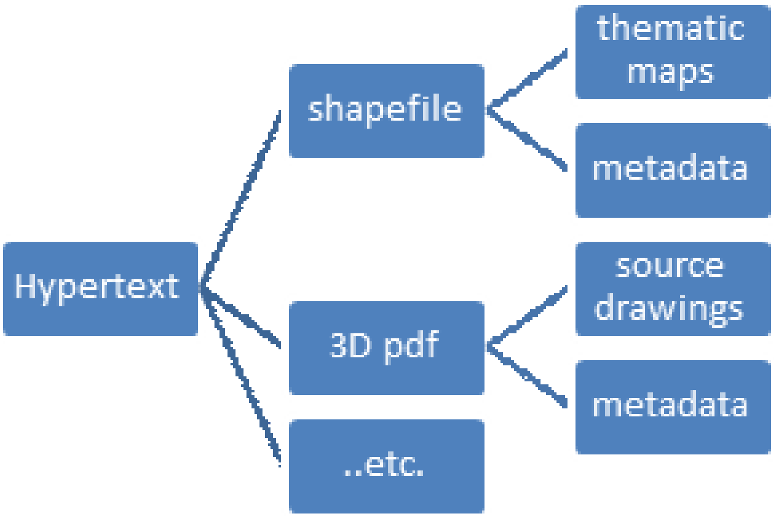


Figure 7 – Hierarchical structure of the documents which translate the content of a traditional book

The use of digital technologies is not only a way to communicate the results of an historical research but, especially during the research, the construction of databases organizing information affords a more complex management of a lot of data that can be enquired crossing many different sources. This digital approach to the research by historians allow a better understanding of the primary data and open new possibilities of interpretation.
This last goal is strongly supported by the development of the Digital History, a branch of the Digital Humanities defined from 1970 as the use of digital media and computational analytic for furthering historical practice, presentation,

analysis, or research.

References

Borgman, C. L. (2015). Big Data, Little Data, Open Data, and Libraries.

Fay, E., & Nyhan, J. (2015). Webbs on the web: libraries, digital humanities and collaboration. *Library Review*, *64*(1/2).

Meyer, H., Bruder, I., Finger, A., & Heuer, A. (2015, January). Building digital archives: Design decisions: A best practice example. In *Emerging Trends and Technologies in Libraries and Information Services (ETTLIS), 2015 4th International Symposium on* (pp. 59-64). IEEE.

Rosso F. (1989). *Alessandro Antonelli 1798-1888.* Electa

Singh, T., & Sharma, A. (2015, January). Research work and changing dimensions of digital library: A review. In *Emerging Trends and Technologies in Libraries and Information Services (ETTLIS), 2015 4th International Symposium on* (pp. 39-42). IEEE.

Stern, D. (2014). *Digital libraries: philosophies, technical design considerations, and example scenarios*. Routledge.

Acknowledgements

The figures of the present paper are based on the final works developed by the students of the course “GIS for Cultural heritage and Digital History” at the Politecnico di Torino in the academic year 2013-2014.

Acquisition of audiovisual Scientific Technical Information from OSGeo by TIB Hannover: A work in progress report

Peter Löwe[1], Margret Plank[1], Paloma Marín-Arraiza[1]

[1] German National Library of Science and Technology (TIB)

Abstract

This paper gives a work in progress report on the application of the TIB|AV Portal for audiovisual OSGeo content. The portal is a web-based platform for audiovisual media combining state-of-the art multimedia analysis with semantic based analysis, and retrieval. It meets the requirements by special libraries for reliable long term preservation, scientific citation via persistent identifiers, and applies metadata enhancement to enable innovative services for search and retrieval.

Keywords

Audiovisual content, Digital Object Identifiers, Citation, OSGeo, Preservation, Semantic Web

1 The Challenge of Scientific Technical Information

Communication within OSGeo involves organisation committees, project maintainers, developers, application of related topics, education and reach-out. From the library perspective, much of this consists of scientific technical information (STI) and is conducted on alternative channels beyond the traditional journal-based scientific discussion. The advent of ubiquitous digital recording equipment and screen-cast-software has led to a steady rise of audiovisual content, such as lecture recordings or data animations. This trend continues and is likely to accelerate further due to the widespread use of handheld end-user devices to access such content, a diversification of application scenarios (“software mash-ups”), and the growing number of OSGeo-approved software projects.

Currently the majority of this content is distributed via proprietary commercial Web 2.0 platforms (e.g. Youtube). Access to and reuse of such content is hampered due to the lack of appropriate metadata, citation, and long term preservation. These topics exceed the scope of the Web 2.0 audiovisual platforms that are currently used., However, theyare crucial for the discoverability, the long term availability and re-use of audiovisual STI in general. This is part of a larger challenge currently affecting academia regarding access to heterogeneous STI content as part of the general trend towards data-driven science.

2 The German National Library of Science and Technology

The German National Library of Science and Technology (TIB) ranks as one of the largest special libraries worldwide. Its task is to comprehensively acquire

and archive literature from around the world pertaining to all areas of engineering, architecture, chemistry, information technology, mathematics, and physics. TIB actively tracks relevant output from OSGeo projects, whichis significant for the field of applied computer science.
Moreover, the information portal of the TIB, GetInfo, provides access to more than 160 million data sets from specialised databases, publishers, and library catalogues.
The access to non-textual materials such as audiovisual media, 3D objects, and reseach data, as well as the use of these materials, also concerns the TIB . This is the main task of the Competence Centre for non-textual Materials. Tools and infrastructure that enable the easy publication, discoverability, and long-term availability of non-textual materials are developed to support users in scientific work processes.

3 The TIB|AV Portal Video Platform

Launched for practical use in April 2014, the TIB|AV-Portal is a bilingual (English/German) web-based portal for audiovisual media that optimises access to scientific videos in the field of science and technology. It was developed together with the Hasso-Plattner Institute and designed to overcome the limitations encountered in current commercial Web 2.0 video portals according to the requirements of a data-driven special library. By combining state-of-the-art multimedia retrieval techniques with semantic analysis, it provides content-based access to videos at the segment level, and link content to new knowledge. The processing workflow of the video analysis includes structural analysis based on video shot detection, optical character recognition, automated speech-to-text transcription, and visual concept detection.

3.1 Metadata of the TIB|AV Portal and the Linked Open Data Context

Automatically generated metadata are processed via linguistic and semantic analysis. That is, named entities are identified, disambiguated and mapped to an authoritative knowledge base. This knowledge base consists of subject specific parts of the Integrated German Authority Files (GND), available as Linked Data Services of the German National Library. The English labels were gained by mapping GND entities onto other authority files.
The portal uses it own schema for the authoritative metadata (formal, technical and content-describing) to describe and manage its content in a standardised manner (Lichtenstein et al., 2014).
The automatical and authoritative metadata of each video are stored in RDF form in a local RDF Store (also known as Triple-Store). The connections between elements stored in the RDFStore are annotated by using an internal ontology. It is planned to map and merge this ontology onto external vocabularies available as Linked Open Vocabularies in other to improve the interoperability of the data.
Furthermore, additional information concerning authors and institutions will be included to facilitate the searchability and discoverability of the content.

3.2 Features and Services

Each video is registered by a digital object identifier (DOI). A DOI name clearly identifies the video, akin to the use of ISBN in books. In addition, the TIB|AV

Portal offers a time-based citation link, enabling a citable DOI to be displayed for each video segment using the open standard media fragment identifier (MFID).
A visual table of contents provides a quick overview of the video facilitating access to individual video sequences. Content-based filter facets for search results enable the exploration of the increasing number of video assets. The term search is not only performed within authoritative metadata but also within metadata from video analysis, giving different term weight when searching. These techniques allow the users to search more efficiently and find content that otherwise would remain hidden. Producers of scientific films, such as the OSGeo communities, can upload their video to the TIB|AV-Portal free of charge. Once the quality of the video has been checked, it is published in a legally watertight manner, indexed according to international standards, transcribed, digitally preserved, and given a DOI name. This ensures an optimal discoverability of scientific films.

3.3 Collections

Films from the TIB´s fields are collected in both German and English. This covers recordings of conference presentations, panel discussions, and recordings of experiments (microscopic images, modelling, simulations and presentations of specific software), among others. Audiovisual content from the OSGeo communities has been steadily acquired since its launch in 2014. The content available by March 2015 ranges from conference recordings (FOSS4G-EU 2013, FOSSGIS 2014, FOSS4G 2014) to thematic-renderings of code evolution (e.g. http://dx.doi.org/10.5446/14652) and historic GIS footage dating back to 1987 (e.g. http://dx.doi.org/10.5446/12963).

4 The European Perspective on audiovisual Scientific Technical Information

The large amount of audiovisual STI generated by OSGeo exceeds both thematically (i.e. topics beyond TIB's acquisition foci) and qualitatively (i.e. multilingual content) the range of acquisition and customer services offered by TIB.
A comprehensive acquisition strategy addressing the multilingual and multitopical diversity of OSGeo's scientifically relevant output remains a challenge to be taken on by several special libraries. This will lead to added value for the OSGeo communities, libraries and their users, and also the general public. The experiences gained from the TIB|AV-Portal can serve as a starting point towards a modular best-practice based approach for this objective.

5 Conclusions

Digital audiovisual content has become an important communication channel for STI within and beyond OSGeo. This content is currently hosted on publicly available commercial web portals whose functionalities do not meet the needs for reliable long term scientific preservation, access, and citation. Since the production of audiovisual content diversifies and accelerates, best practices are needed to address and solve this challenge on a global scale. The TIB|AV-Portal for audiovisual STI meets the requirements to preserve such content and to provide innovative services for search and retrieval. Simultaneously, the research and development for improved services continues. Quality checked audiovisual content from the OSGeo communities is constantly being acquired for the portal as a part of TIB's mission to preserve relevant content in applied computer sciences for science, industry, and the general public.

6 References

[1] Berners-Lee, T. (2006, July 27). Linked Data-Design Issues. Available in : http://www.w3.org/DesignIssues/LinkedData.html

[2] Lichtenstein, A., Plank, M., & Neumann, J. (2014). TIB's Portal for Audiovisual Media: Combining Manual and Automatic Indexing. *Cataloging & Classification Quarterly*. 52(5), 562-577. doi: 10.1080/01639374.2014.917135.

[3] Strobel, S. (2014). Englischsprachige Erweiterung des TIB|AV-Portals: Ein GND-DBpedia-Mapping zur Gewinnung eines englischen Begriffssystems. *o-bib*. 1(1). 197-204. doi: http://dx.doi.org/10.5282/o-bib/2014H1S197-204.

An open-HGIS project for the city of Parma: database structure and map registration

Nazarena Bruno, Giorgia Bianchi, Andrea Zerbi, Riccardo Roncella

Dipartimento di Ingegneria Civile, dell'Ambiente, del Territorio e Architettura (DICATeA), Università degli Studi di Parma

Abstract

The observation of the structure of the city and its inhabitants, and how such system evolves in time, is essential in almost all land planning and urban development activities, as well as in any historical study. Historical GIS in the last decade proved to be one of the most promising frameworks, that can provide innovative tools and methodological approaches in this field, allowing new research perspective and concept in many fields. The paper describes the early stages of the development of a HGIS that considers different historical periods of Parma (Italy), with particular attention to cadastral documents and associated census data. These data are gathered from the main cadastres of the city: the Sardi Atlas of 1767, the Bourbon Cadastre of 1853, the Post-Unitarian Cadastre of 1901 and the Cadastre of 1940. Other social statistics, conserved in the local communal and state archives, such as census reports, building permits, deeds, transfer of property, wills, etc., were considered in the project, leading to a quite complex database structure. In the hope of making the HGIS as flexible and scalable as possible, and to make the system usable to the widest spectrum of users, an open-source database architecture based on Postgresql/Postgis has been developed. The paper provides information about the database structure and the map registration stages, with particular reference to the first historical period considered in the project, based on the Sardi Atlas (drawn in 1767), which can be considered as the first parcel-geometric cadastre of Parma.

Keywords

Historical cadastres, GIS, georeferencing, open-source, data-driven, image registration

1 Introduction

The analysis of the city from many points of view, has always been important for historians of different disciplines. The observation of the structure of the city, its transformations and inhabitants that live in it, can provide an idea of society as well as architecture or urban structure. Therefore, knowing the structure of the city and interrelating it with its inhabitants can reconstruct the image of the society and the social and economic dynamics.

To this aim, the analysis of cadastres and censuses is very important because these documents can be considered as main sources for knowing urban, architectonical and demographic structure of the city. Moreover, the investigation of these documents is particularly significant to identify the forms of land use, urban structure, pre-existences and distribution of the property.

Cadastres became privileged sources of territory knowledge in the Eighteenth century, when their realization became systematic in several Italian regions and the principles, which govern their structure still today, were identified. They included maps, based on topographic surveys, which showed the individual parcels. Each parcel was then included in a register containing information relative to its owner, type of crop and annuity. The paper will describe the first activities of a scientific project for the development of a Historical Geographic Information System (H-GIS or HGIS) of the city of Parma (Italy). In its very early stage the main purpose of the project was to provide a multi-user, low cost, information system for historians as a support for urban development studies. Now a further improvement of the project, addressing all the cartographic sources (up to modern age) and historical statistics (census reports, building permits, deeds, transfer of property, wills, etc.) preserved in city archives, that can provide thematic data to a wider range of user (e.g. architects, land planner, researcher, etc.), is in progress.

The city, in fact, thanks to the creation in 1767 of its first geometric-parcel cadastre, known as Sardi Atlas (from the name of the cartographer Gian Pietro Sardi who drew it), took part in this first phase of national cadastral survey and was among the first cities that adopted a modern cadastre.

In spite of the undoubted importance of documentary sources for historical city knowledge, in general their direct consultation is often difficult and very expensive in terms of time and effort in research. This because of problems related to conservation, lack of copies and reduced circulation, which limit the use of documents. To overcome these difficulties, computer technologies make new tools (above all GIS) available today for the analysis of historical data. Even if there is a difference between the original source and its digital reproduction, today digital scans of maps and documents and informative systems based on historical data are increasingly widespread.

On the other hand, relational databases allow, in a much simpler way, not only historical investigations related to the single historical period (synchronic analysis) but also to different historical periods (diachronic analysis) in order to know the urban evolution in time. Despite the irreplaceable value of the original source, the information system aims to reproduce the documents in digital format, without distorting their structure and contents, in order to broaden the number of users and improve the research techniques thanks to the computer.

Whilst in the late ’90 the growth of HGIS was slow and episodic (Gregory & Ell, 2007), from the very beginning of 2000 the scientific community developed a growing interest in such topics (Knowles, 2002). In many works one of the main features that make HGIS attractable for a wider range of scholars and researchers is an easier consultation of archival documents, thanks to the possibility also to visualize on a map information that are not spatial. For example in (Gauthiez et al., 2009, 2014) a HGIS for the city of Lyon (France) is described. The project, started in the late ’90, in based on the spatial localization of several buildings described in archival documents, in particular the building permits given from 1617 to 1763. The maps used as base for the spatial reference belong to the 1745 to 1830 period, with particular regards to cadastral maps of 1830.

Another project (Gregory et al., 2002), involving a more extended territory and a bigger amount of documents, is the GBHGIS realized by the University of Portsmouth on the whole Great Britain territory, since 1994. The original work was entirely focused on historical statistics: census reports, data on births, marriages and deaths, and on unemployment and poor law statistics. The data have been related to digital maps. Because the data had an high variability in time, a sequence of static maps in this project were not created, but each datum is date-stamped with a start-date and an end-date.

An interesting HGIS (Carrion et al., 2013) concerning medieval data related to Southern Italy, also with the implementation of a WebGIS prototype, was developed by Politecnico of Milan research group in the frame of two national projects (PRIN 2006 and PRIN 2009).

Usually GIS based on cadastral sources are a little different in structure: in this context not only national, but also regional differences make the comparison of choices and GIS architecture quite awkward. Similar to the Parma condition there are two projects, both about the city of Rome. The first, elaborated by the CROMA group of Roma Tre University and titled “Atlante Storico di Roma Moderna e Contemporanea” (Baiocchi et al., 2005; Lelo et al., 2005), refers to different historical periods of the city: the Map of Nolli of 1748, the Pio-Gregoriano Cadastre Maps over the period 1818 to 1824 and the Census Maps of 1866. The second (Micalizzi et al., 2012; Descriptio Romae, 2015) is developed by DIPSU department of Roma Tre University and brought in 2007 to the realization of a WebGIS of the Gregoriano Cadastre, integrated with other archival documents related to buildings features. For this reason, this informative system goes beyond the traditional use of GIS for urban and territorial analysis, making possible queries about the buildings.

In all these projects, open source technologies have been widely used in their realization. For GBHGIS the relational database was, on the first stage, managed with Oracle software, but during the second step of the project, the Vision of Britain system was moved to Postgres and PostGIS, as this open-source software had developed sufficiently to cope with the demands of the system.

Both the Southern Italy project and the Gregoriano Cadastre of Rome WebGIS were entirely developed with open-source technologies: in the former the conceptual model of the database was implemented with SQL Power Architect, the database uses PostgreSQL and the WebGIS prototype was developed with QGIS Web Client. The latter instead used Map Server to implement the cartographic engine of the WebGIS, Postgres/PostGIS as Data Base Management System and Apache Web Server and PHP5 to edit and develop the system.

This wide use of open source software, even in the management of complex systems such as the British project, testifies to the validity of these products, which are today more and more developed and can achieve performance comparable to commercial software but, in many frameworks, with higher flexibility.

The paper structure is developed as follows: first, a very brief description of the documents and maps used in the project, as well as the historical context in which they were developed is provided. Then a technical overview on the open-source database architecture of the HGIS system is described. In this context, probably, the project is not particularly innovative with respect to (w.r.t.) other cadastral HGIS. However, the reader should consider that the peculiar nature of historical sources (in particular as far as census and cadastral data are concerned) presents very specific and local features that make every single project innovative in such sense. In addition, being this the first technical contribution about the project, the main section of the paper addresses a new registration technique for the georeferencing stage, while further information about the implementation of the different historical periods and technical details about time-domain management will be provided in future works.

2 The Parma HGIS project

The Parma HGIS project is based on the computerization of cadastres and censuses, which, as seen before, are sources of primary importance for the knowledge of the city. Four main historical periods were identified. They correspond to key moments in the history of Parma and to the four main cadastral productions: Sardi Atlas (1767), which summarizes the reformist and illuminated period led by Minister Du Tillot during the first Bourbon government; Bourbon Cadastre (1853) realized during the government of Maria Luigia and considered as a continuation of Napoleon cadastre; Post-Unitarian Cadastre (1901) and Cadastre of 1940.

Although the database in general has been structured taking into account all the sources, providing a possible implementation with archival data as deeds, building permits, etc. of all the historic periods, at present, the work was completed entirely only for the Sardi Atlas (with the Bourbon Cadastre at 90% completion).

The Sardi Atlas is the first cartographic cadastre of Parma, made between 1765 and 1767 by the cartographer Gian Pietro Sardi. Its realization, promoted by the Minister Guillaume Du Tillot during the first Bourbon government of the city (1749-1804), can refer to the general program of knowledge and urban ornamentation of the city. The Atlas includes maps and registers with details about the particles drawn in the maps (see figure 1.).

Figure 1 - One of the Sardi Atlas tables.

As far as the internal structure is concerned, the Atlas has a table of location in scale 1:5000 that represents the city within the surrounding area and 28 detail tables, approximately in scale 1:850. The image depicted is objective, accurately surveyed and in strictly zenith projection. The tables are well-finished in the layout, thanks to the frequent presence of drawings on paper edge, and each element, according to the scale of representation, is drawn with the same degree of detail. There is distinction between built and unbuilt areas and particular attention is given to religious buildings, outlining the micro-plans, which are a section of the ground floor with the drawing of the roofing system. Each of these 28 tables is complementary to another one containing the list of holders of each parcel drawn in the map. The lists are drawn up in tabular form and contain the name of the holder and information on real estate. In this way, thanks to the precise correspondence between maps and lists, each graphic parcel can be associated to its owner. Each parcel is uniquely identified by the combination between its number and the number of block to which it belongs. The maps show, beside the distinction between covered and uncovered areas, also the presence of vegetated areas, place names, the canal system, and therefore are much more detailed than the ones produced by modern cadastre. Lists of owners instead provide information about the possessor, including name and surname, type of institution (if the possessor is not a natural person), title and, sometimes, degrees of relationship, patronymic, right of use of the land, intended use and share ownership of parcel. The connection between graphic tables and registers is through the number of parcel and block.

3 The Database structure

From the previous paragraph a complex and varied HGIS structure appeared. This, as far as the data and the relationship are concerned, is due to the coexistence of more historical periods, so distant in time. Approaching the contemporary, in fact, the informative content that can be deduced directly from the cadastral maps becomes more abstract (the maps of Sardi Atlas of 1767, for example, allow highlighting the types of vegetation, the plans of the main buildings, the canal system) but the information in the register associated with maps increases. The same consideration can be done w.r.t. censuses, which increasingly characterize in detail the individual inhabitants, allowing their precise placement within the building. Despite of these differences, which can be solved by a progressive increase of the attributes assigned to the same logical-conceptual entity, the type of data contained in the documents remains substantially constant over the time; the data derived from the registers are in fact always cartographic and descriptive, the one deduced from the census are, instead, only descriptive.

In other words, we evaluated which was the structure to be used for better overall data management. Among the various possible solutions, the focus has been placed in particular on two alternatives: a structure that would give preference to the internal coherence of the single source, thanks to the creation of specific tables for each, or a structure that, at the expense of the general compactness and consistency on individual sources, would make possible a more homogeneous database of all historical periods. It was decided to adopt this second solution, as it allows more easily the development of diachronic query, maintaining an overall structure rather simple.

One of the main system requests, being the HGIS initially devoted to historical study of the city social structure, was to enter data as reported by the source, without any interpretation and with specific archival reference so that the user, most likely an historian, can verify them in any moment. In this way, it is always possible to refer the data to a particular historical framework, with the possibility to separate it from the others. It is also possible to associate with the same object discordant data, gathered from different sources, leaving to the end user the ability to lean towards the greater reliability of one rather than the other. Finally, the precise reference to the source allows the reconstruction of the urban evolutionary process and the transformations occurred.

For the choice of the RDBMS (Relational DataBase Management System) to be used for the system, we opted for a platform based on PostegreSQL, with PostGIS extension for spatial data management. The choice was made for several reasons: first of all its open source nature allows to contain costs connected to operation and maintenance of the system, despite sometimes involves more technical complications and cannot always ensure timely technical assistance. Often, to solve some specific issue, a request on one of the many technical forum on the web receive a feedback in few hours, but the users have to faith in the support of the community.

While all major GIS software provide very efficient (and easy to use) graphical user interface, with PostGIS spatial database (albeit with different technical complications) it is much easier to expand as much as possible the level of interoperability with different GIS platforms (ArcGIS, AutoCAD Map, QGIS and Grass, etc.). Many of the previous packages (especially commercial ones), while providing enterprise multi-user architectures, require additional modules or complex data administration systems for specific procedure, increasing the final cost. In this context, since one of the main project objective is to provide data to a webGIS platform, to make the system usable by the wider spectrum of users possible, the opportunity not to be bound to a commercial solution provides greater flexibility, guaranteeing also a higher cost limitation.

The operation of vectorization of the maps was not particularly costly both in terms of time and work organization: the information contained in the maps are, indeed, easily classifiable within the structure of the information system thanks to simple decomposition in graphical units. On the contrary, the insertion of descriptive data contained in the lists of holders was more complex: thematic data are difficult to standardize because their complexity and heterogeneity. A preliminary phase of analysis of the document was therefore necessary to be able to extrapolate the invariant information, on the base of which organize a database structure that allows the inclusion of most of the cases also considering possible exceptions. However, before data insertion into the GIS, a phase of transcription and normalization of all entries in Sardi Atlas lists was carried out. The reason was that the original data were typologically different, written using a wide range of conventions and with a large number of exceptions, anomalies, inconsistencies, which obviously could not be contained in the database tables without negative impact on its functionality. The solution of these problems involved numerous historical researches to avoid oversimplification, distorting the original source. Again, in respect of the source, it was decided to leave traces in the database of the original form of data as reported in the Atlas, but without the insertion of special formatting or graphic signs that are not supported by the system.

As can be seen from this brief discussion, the step of transcription of the data has been very demanding in terms of time and difficulties in interpretation. Since they are historical data, their encoding is very complex for both the intrinsic differences and the necessity to keep an attitude of respect of the source. Before inserting the data, is in fact necessary to resolve the issues related to the interpretation of historical fact, sometimes not uniquely interpretable, resulting in a significant lengthening of database implementation. Such peculiarity are, in fact, specific of the Parma cadastres (and huge differences are anyway present between the different historical thresholds) and cannot be inferred or deduced by other HGIS projects.

4 Maps georeferencing

One of the main steps that should be carefully considered in developing a Historical GIS is the maps georeferencing stage. The integrated use of co-registered historical and modern maps is probably the first, and for most users the main, requirement for the system itself. At the same time all the different maps, along the centuries, were produced using different surveying and drawing techniques, and were commissioned often with different illustration and documentation purposes: unfortunately map metadata is a modern concept, and is quite rare to have any description about the restitution procedures adopted for the historical maps. Therefore, precision, reliability and correctness of the graphical information should always be examined and evaluated indirectly.

Probably, for many HGIS applications, a very accurate maps co-registration is not an essential feature since comparison with other historical periods or with modern cartography is a qualitative operation. The availability of old cadastral maps and historical cartography, are usually considered for analysing land and urban transformations along time. Nonetheless, such analysis commonly does not involve a thorough comparison at parcel boundary level, where errors higher than few meters cannot be accepted. Rather, the analysis often considers land and city development at a smaller scale and is focused on macroscopic transformations of areas and land use.

Nonetheless, even in those circumstances, where a perfect overlay of the different historical layers is not a severe requirement, it’s important inferring the correctness of the map, highlighting possible discrepancies and making possible to provide at least some hypothesis on the implementation of the map itself: e.g. at the very minimum, the result of the transformation provides information on dimensional parameters of the map and on the scale of original representation, even in cases where they are not known. Often, at the same time, analysing with proper carefulness the residuals of the registration transformation, especially if a comparison/overlap with modern cartography can be done, the accuracy of the map and the mistakes made during the survey and restitution operations can be evaluated. Finally, to critically evaluate the possible discrepancies, the analysis of the map should be considered with respect to its historical and cultural context. Such analysis should consider the scientific and geographical knowledge of the era and the progress of surveying techniques, but also the cultural background and political ideology that influenced greatly cartography in the past, determining types and methods of representation. It would be also very useful to know the author, date of creation, the equipment used, the contents and the objectives of representation, thus being able to provide for any errors (Brovelli et al., 2012). Unfortunately, this information is often missing, and surveying and restitution methodologies remain unknown. Finally uncertainties or lack of information about measurement units, map reference and projection system, as well as interpretation of semantic content, can make the georeferencing process much more difficult and tricky.

In the end all these uncertainties, as well as the choice of the registration transform model, affect the final result in terms of discrepancies between the points of the historical and modern map. The residuals, in fact, are the result of a summation of errors inherent to the historical map and errors related to the georeferencing process itself. As far as the historical data are concerned, the total graphical error arises from several causes including survey measures, graphic restitution and the deformations due to the preservation condition of the map. In addition, even if modern instrumentation is usually very accurate, instrumental errors due to the digitalization of the map should be considered. Finally for cadastral historical maps, where high scale factors are usually implemented (e.g. the Sardi Atlas has a 1:850 ca. scale), during co-registration the user should also consider the tolerances of modern surveying equipment or cartographic products used for georeferencing, and consider in addition the possibility that the position of homologous points not necessarily remained unchanged in time.

The Sardi Atlas maps were scanned at high resolution (400 dpi) while for acquiring the Ground Control Points (GCP) for map registration the official Topographic Regional Carthography (CTR – nominal scale 1:5000, reference system EPSG 23032) was used. During activity design, we decided to discard the idea of performing a thorough GPS survey of historical points considering not justified the higher cost for GCP materialization compared to the small budget for the project and low requests in terms of accuracy.

The double points were identified at corners of buildings or blocks of which the position has remained, with all probability, unchanged over time (monuments and historical buildings). While in the central areas of the city their detection did not involve a lot of problems, given the presence of many noble and religious buildings not destroyed during World War II, for many tables in the areas close to city walls the same proved to be more difficult. At the end of the XVIII century, those areas consisted mainly in gardens and cultivated fields and were used just in the next century for city expansion. For this reason while a good number of GCP can be found in the inner part of the city, a proper and reliable co-registration of outer drawings proved to be much harder.

For this reason a procedure derived from image block orientation in photogrammetry (and in particular in independent model adjustment) was adapted for the registration stage. Basically, in the same period, a similar approach was independently developed in (Barazzetti et al., 2014) so just few procedure details will be drawn in the following sections.

For outer plates, even if few historical GCP are present, many common elements between the maps can be found. For example, on each Atlas table there is a hint to the perimeter of surrounding blocks of neighbouring plates where common points can be easily tracked (in analogy to photogrammetry orientation such points will be defined Tie Point (TP) in the following) and used for mosaicking/registration (see figure 2).

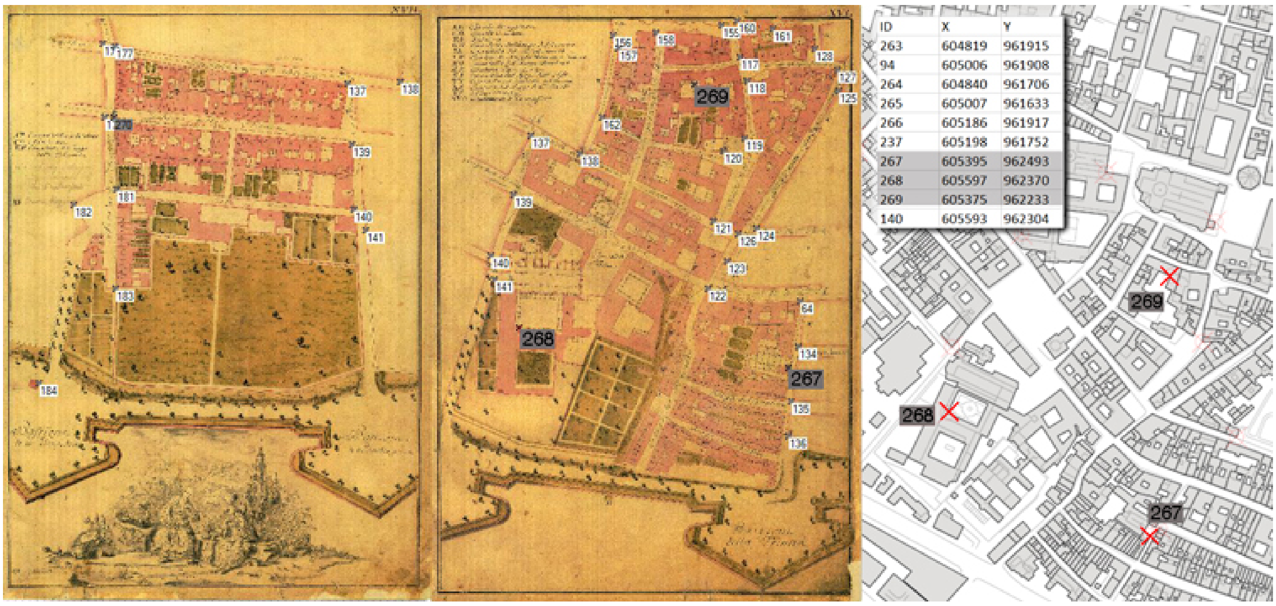


Figure 2 - Image measurements including the different categories of points: TP (white) and GCP (grey).

All the registration tools implemented in the most common GIS packages, both commercial (e.g. ESRI ArcMap, ENVI, ERDAS Autosynch, Autodesk Map, etc.) and freeware (e.g. Quantum GIS), have specific feature for map georeferencing. The procedure are usually quite straightforward, since the software allows loading different map sources or point coordinate lists, provides specific tools to allow the operator to identify and extract corresponding points and finally estimate the georeferencing transformation using a specific, user defined, geometric model. In the end the software can produce a World File definition (if the model is a translation or an Helmert transformation) and/or resample the original scanned map. Even the simplest software packages have the additional feature, at the end of the procedure, of reporting the results of the transformation, where double points residuals are commonly shown.

Nonetheless, as mentioned in (Barazzetti et al., 2014) a strong limitation of many of those software packages is that the procedure is commonly carried out as a progressive registration of maps, since the user can consider just two maps at a time. This means that, especially for those maps where TP are used, the registration errors accumulates along the georeferencing procedure and can make really hard the identification of possible discrepancies, in particular for those maps where no GCP are present and residuals are thus commonly low. For that reason a very simple C# software code was developed: being C# an object-oriented programming language its translation to other platform or languages, e.g. Python or Java, is straightforward.

In the proposed methodology all the maps are co-registered concurrently integrating GCP (or image-to-ground points) and TP (or image-to-image points) observations. As far as the transformation model is concerned, every parametric transformation can be taken into account in the processing framework. However, for the project, just two “not exact” (see Brovelli et al., 2012), parametric transformation model were considered: the simplest is a Helmert 2D model where, beside the rotational and translation parameters, the scale factor can be estimated or fixed (using, for instance, the value that can be evaluated using the graphical scale reported on the maps):

$$\begin{pmatrix} X = a_0 + \lambda\cos\theta\cdot x - \lambda\sin\theta\cdot y = X_0 + a\cdot x - b\cdot y \\ Y = b_0 + \lambda\sin\theta\cdot x + \lambda\cos\theta\cdot y = Y_0 + a\cdot y + b\cdot x \end{pmatrix}$$

Where X and Y are ground coordinates of object point, X_0 and Y_0 are the translation components, λ is the scale factor and θ represents the rotation angle of the transformation. Alternatively, an affine transformation can be used:

$$\begin{pmatrix} X = a_0 + a_1 x + a_2 y \\ Y = b_0 + b_1 x + b_2 y \end{pmatrix}$$

where a_i and b_i, are free parameters. The transformation model includes, besides a rigid rotation and translation, a non-isotropic scaling, and therefore lengths ratio along different directions and angles are not preserved.

Often, in similar cases, an affine transformation is able to correct some deformations that might occur during paper conservation or during the scanning stage (for example due to a not perfect orthogonality or scale coincidence of the scanning axis). However, if just few, not well distributed, double points were selected, the higher transformation degree of freedom can introduce unwanted deformation effects that are difficult to be identified. In this context, a conformal (Helmert) transformation is certainly a more conservative solution. The use of TP has proved essential to apply consistently the transformation to all the tables since, especially for outer areas, in many cases there were less than necessary GCP.

Both observation equations (1) or (2) are linear w.r.t. their transformation parameters and can be easily solved (the user might mix both transformation models) with a standard Least Squares Adjustment technique:

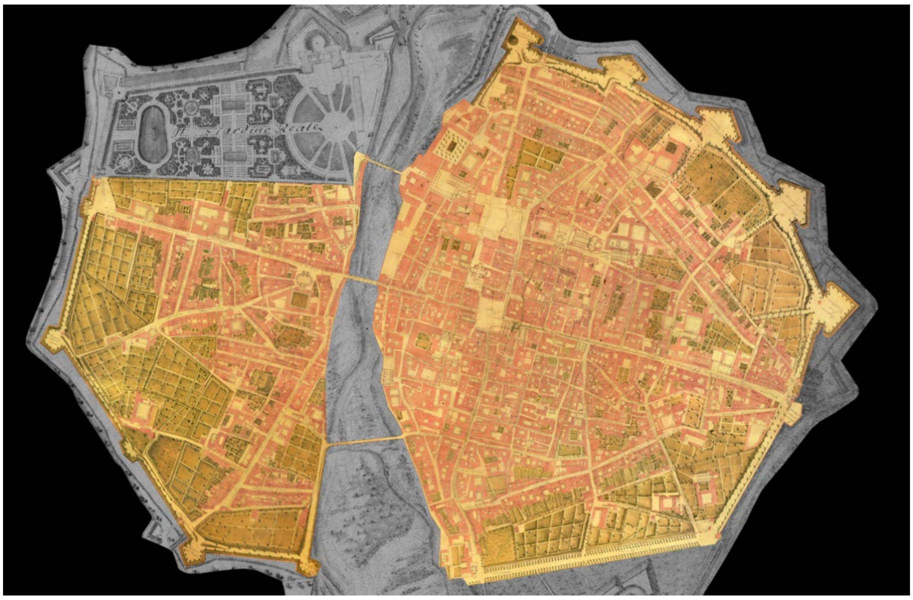
$$x = \left(A^T P A\right)^{-1} A^T P b$$

Where A is the design (coefficient) matrix, x the solution (unknown or parameter) vector, b the constant term vector, and $P = C_{yy}^{-1}$ is the weight matrix of observation that implements the stochastic model of the solving system. The latter allows specifying different weights, and thus different importance for the final solution, of the different observation. For instance, if the user can quantify, at least approximately, the expected error for GCP equation w.r.t. TP observation, the stochastic model can influence accordingly the final solution consistency. On the other hand, changing the weight of GCP and TP observation, the user might prefer to increase the adherence of the final registration solution with current cartography or, on the contrary, prefer a higher cohesion of adjacent plates, limiting the residual of TPs.

In this context, to provide an appropriate (or at least rational) definition of the weights, we considered several components that could affect the final result. First, we considered the factors that affect the correspondence mapping between current and historical maps, as survey errors, definition of the parcel boundary, deformation of the map, poor accuracy in the identification of the points on the CTR (which has a smaller scale compared to the Atlas), etc. The evaluation was empirical and extremely hard to be conducted in quantitative terms, considering the difficulties to clearly determine the various error components. Given the measurement tolerances of modern cartography technology at CTR scale (1:5000), and considered the measuring instruments used for the Sardi Atlas, it was considered for the equations relating to GCP a value of σ_{xy} equal to 2 m.

On the other hand, considering the precision of the design of Sardi Atlas, and the high coherence between adjacent plates obtained along the perimeter of city blocks, we found that the Atlas presented a very good internal consistency. Given the scale of representation of the Atlas, considered the drawing techniques and the good correspondence of the elements as a result of some tests, for the equations relating to the TP a σ_{xy} = 0.4 m was used. With these values the final global co-registration of the maps provided probably also the most appealing results. The solution shows very good consistency of the elements on the maps border, with little or no troubles in the mosaicking stage (see figure 3a), but still with a good correspondence of historical features with modern buildings (see figure 3b).

a.



b.

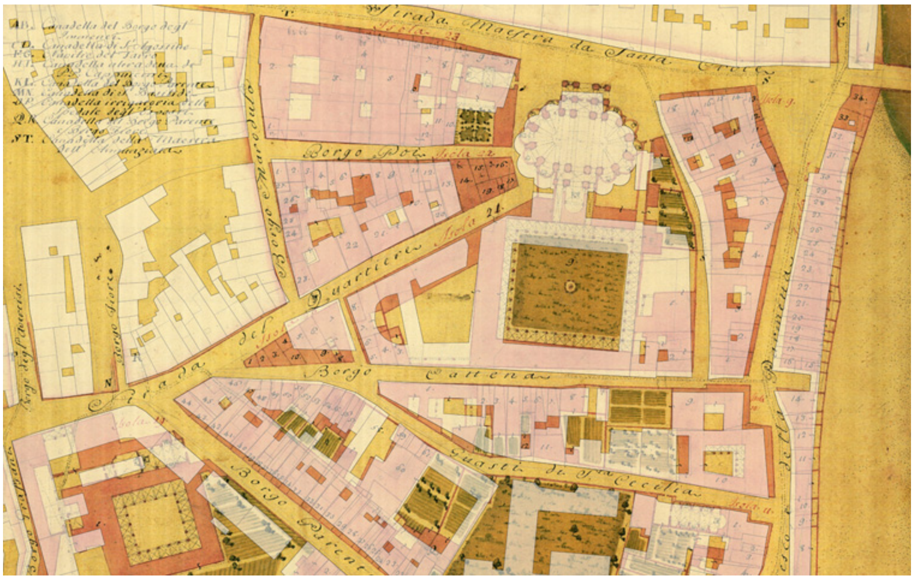


Figure 3 - a. Graphical representation of all the 28 Sardi Atlas tables after mosaicking.
b. Overlay of modern CTR with one of the Sardi Atlas table after registration.

As far as the transformation model to use is concerned, both Helmert and affine model were considered in our tests. However, from the comparison between the different georeferencing global solutions (we also tried to change the stochastic model to verify if different weights can influence the solution), we found that the results, using the same covariance matrix, were fully comparable and with residual errors very similar. Therefore, considering that the map support showed no particular deformation to justify the application of an affine transformation, it is preferred the use of a conformal transformation model.
For outlier rejection, instead of the RANSAC approach proposed in (Barazzetti et al., 2014), expecting few outlier, a much simpler data snooping procedure were implemented.

6 Conclusion and future development

The paper addressed the very early stages of an open-HGIS project regarding the historical cadastral documentation of the city of Parma. At the time of writing the first historical period, represented by the Sardi Atlas of 1767, is completed, and the next historical threshold concerning the Bourbon Cadastre of 1853 is almost done. The next project stages will involve, apart the completion of the Bourbon Cadastre data layers, the insertion of the later cadastral thresholds, and the further addition of other documental sources.
From the very beginning of the experimentation, the research pointed out that all the historical documentation are extremely complex in their structure and organization, with numerous ambiguities and intrinsic peculiarities which are difficult to standardize. Historical sources are in fact not homogeneous, both with respect to language and internal organization. Even the historical maps have often problems, especially related to difficult interpretation of descriptive content because of errors or omissions and, in some cases, inaccuracy of drawing or incorrect conservation of the paper support. Therefore, the organization in an encoded form, suitable for computing, of the heterogeneous reality reproduced by the Sardi Atlas was particularly onerous, but necessary for database structuring.
At present, it is possible to analyse the city referring to the original map source (see for instance the raster layers of figure 3). In addition, the first "beta-testers" of the system can make more complex queries, such as statistical and thematic analysis, related to the holders or to specific features concerning spatial data (see for instance figure 4).

Figure 4 - A thematic view of the Sardi parcel layer where areas are themed differently: built areas (brown), unbuilt areas (beige) and green areas (green).

The aim of the project is to make the system usable by the wider spectrum of users possible, also through a webGIS platform. In this regard, issues related to the dissemination of the results remain to be investigated, since the historical sources used are property of archives, libraries and museums that often pose restrictions on the disclosure and claim property rights on such sources.

Despite these difficulties under solution, in all these contexts of access to data, the choice of implementing the HGIS using open-source geospatial frameworks/software proved to be very successful. With commercial ("closed") GIS solutions, users are usually more confined in the specific software architecture/workflow, and interoperability between different researches is limited by the software platform itself (often the users must use the same software, with the same modules, to efficiently exchange all the information). On the contrary, an open-source architecture, with software and data infrastructure easily affordable at little or no cost by everyone, is the easier way to ensure that scalability as well as information and expertise exchange can be implemented in a data-driven environment. In this regard, at the end of this first phase of the project, an evaluation of the results can be made: the project involves a significant commitment of resources, since it requires specialized skills and many hours of work by the operators. However, the decision to use an open source platform and to organize the project in different stages delayed in time and executable by more (volunteers?) people, according to a protocol of intent and common behaviour, was taken wittingly to ensure that the project becomes fully realized and with a sustainable financial budget. Furthermore, we believe that the cultural value of the project, linked to the renovated, more efficient access of a documentary heritage of great value not very consulted, but fundamental for historical studies as well as in urban planning activities, is of primary importance for the community.

References

- ✔ Baiocchi, V., & Lelo, K. (2005, September). Georeferencing the historical maps of Rome between the seventeenth and eighteenth centuries. In *CIPA 2005 XX International Symposium-International cooperation to save the world's cultural heritage. Torino* (Vol. 26, pp. 114-118).
- ✔ Balletti, C. (2006). Georeference in the analysis of the geometric content of early maps. *e-Perimetron, 1*(1), 32-42.
- ✔ Barazzetti, L., Brumana, R., Oreni, D., & Previtali, M. (2014). Historical Map Registration via Independent Model Adjustment with Affine Transformations. In *Computational Science and Its Applications–ICCSA 2014* (pp. 44-56). Springer International Publishing.
- ✔ Brovelli, M. A., & Minghini, M. (2012). Georeferencing old maps: a polynomial-based approach for Como historical cadastres. *e-Perimetron, 7*(3), 97-110.
- ✔ Carrion D., Migliaccio F., Minini G., Zambrano C., 2013, Rappresentazione cartografica e condivisione di dati storici in ambiente GIS*. Atti 17a conferenza nazionale ASITA*, 5-7 novembre 2013, Riva del Garda – Italy*,* pp. 379-384
- ✔ Descriptio Romae (2015). http://www.dipsuwebgis.uniroma3.it. Last accessed March, 2nd, 2015.
- ✔ Gauthiez, B., & Zeller, O. (2009). Espace costruit, espace social à Lyon aux XVII-XIXe seècles: l'apport du SIG. In M. Panzeri, & A. Farruggia, *Fonti, metafonti e GIS per l'indagine della struttura storica del territorio* (p. 39-49).

Torino: Celid.
✔ Gauthiez, B., & Zeller, O. (2014). Lyons, the Spatial Analysis of a City in the 17th and 18yh Centuries. Locating and Crossing Data in a GIS Built from Written Sources. *Mapping Spatial Relations, Their Perceptions and Dynamics, Lecture Notes in Geoinformation and Cartography*, 97-118.
✔ Gregory, I. N., Bennett, C., Gilham, V. L., & Southall, H. R. (2002). The Great Britain Historical GIS Project: from maps to changing human geography. *The Cartographic Journal*, 39(1), 37-49.
✔ Gregory, I. N., & Ell, P. S. (2007). *Historical GIS: technologies, methodologies, and scholarship* (Vol. 39). Cambridge University Press.
✔ Knowles, A. K. (2002). *Past Time, Past Place: GIS for history A collection of twelve case studies on the use of GIS in historical research and education.* ESRI press. ISBN 1-58948-032-5
✔ Lelo, K., & Travaglini, M. (2005, September). The GIS-based historical atlas of Rome. In *XXth International CIPA Symposium* (Vol. 26).
✔ Micalizzi, P., Magaudda, S., Buonora, P., & d'Elia, L. S. (2012). A GIS for the city of Rome: archives, architecture, archeology. *e-Perimetron, 7*(1), 28-35.

Meteorological and Aviation Weather Open Data implementation utilising Open Geospatial Consortium (OGC) standards

Roope Tervo, Mikko Visa, Tero Koivunen, Jukka A. Pakarinen, Tarja Riihisaari

Finnish Meteorological Institute, Erik Palmenin aukio 1, Helsinki, Finland

Abstract

The purpose of the Finnish Meteorological Institute's (FMI) Open Data project was to make some of the institute's data freely available. Project was based on both Finnish government decision to improve availability of public sector information as Open Data and the directive 2007/2/EC of the European Parliament, 2007 Infrastructure for Spatial Information in the European Community (INSPIRE), due to this the FMI's implementation goes far beyond the requirements set by INSPIRE.

The implementation is based on OGC (Open Geospatial Consortium) standards. Latest developments during 2014 has been addition of the Meteorological Terminal Aviation Routine Weather Report (METAR) in System Wide Information Management (SWIM) compliant format to Open Data interface.

The FMI Open Data Portal is based on INSPIRE services. Catalog Service provides the catalog with both user and machine readable interface following Catalog Service for Web (CSW) standard. View Service provides Web Map Service (WMS) interface for browsing the data. Download Service provides the data in INSPIRE harmonised form via Web Feature Service (WFS).

Data formats provided follows OGC Observation and Measurement (O&M) guidelines and INSPIRE specifications. Beyond INSPIRE requirements, more compact data formats are provided for clients requiring better performance and more simple data formats are provided for older clients.

Registration is required in order to use View and Download Services. In the registration process only user information required is email address. After the registration, user is issued an identification key (APIKEY) which is required in every data requests. In case better service level or specific data is needed, clients can be directed to other service clusters based on APIKEY.

FMI Open Data API has been available since June 2013 and we have seen significant growth of amount of developers utilising FMI interface for their application. Latest information regarding FMI Open Data impact is presented.

Keywords

Open Data, Aviation weather, weather, INSPIRE, WFS, WMS

1 Data Sets and Aviation Weather Reports

Following data is shared in FMI Open Data interface; weather, marine and climate observations, weather radar images, lightning observations, data for national weather forecasting and marine models and aviation weather reports (METAR).

Real-time observations are: Station-specific observations (e.g. observations on

wind, temperature, humidity, atmospheric pressure, precipitation, sea level and waves), weather radar images and location data for lightning flashes in Finland.
Time series of observations are: Climate observations from 1959 onwards (station-specific daily and monthly values), Sea level observations from 1971 onwards and observations on waves and sea water temperature from 2005 onwards.
Forecast models are the latest forecast based on the national weather model, including surface weather data at one-hour intervals for 48 hours and the latest sea level forecast and pressure level data covering Europe.
Climate change forecasts for the 30-year periods 2010, 2040 and 2070 (model forecast; average changes for temperature and rain).
A typical METAR contains data for the temperature, dew point, wind speed and direction, precipitation, cloud cover and heights, visibility, and barometric pressure. METAR code is regulated by the World Meteorological Organisation in consort with the International Civil Aviation Organisation (ICAO).

2 FMI Open Data Implementation

The work was aiming to ensure interoperability of meteorological information as set in MetOcean Domain Working Group of OGC and in context of World Meteorological Organisation (WMO) Information System (WIS).
The data sets which have been made available since June/2013 are continuous observational and Numerical Weather Prediction (NWP) model data produced by FMI. Eg. Data consists of 200 weather stations, 13 sea water level stations, weather radar data, spatial data for lightning events, data from national forecasts models and currently also Aviation Weather Report data.
Implementation consists of three services: Discovery, View and Download. Discovery service stands for a high level catalog of data sets. The main purpose of the view service is to give an example of available data. Download service provides the data for users.

2.1 Discovery Service

Discovery service is based on GeoNetwork[1] software and follows Catalog Service for the Web (CSW). CSW provides an API for accessing the metadata records in addition to a web interface. Metadata from the catalog is automatically harvested by the National Land Survey of Finland (NLS) which is the INSPIRE authority of Finland. The NLS catalog is in turn harvested by the European Commission’s INSPIRE Geoportal.

2.2 The View Service

The most relevant meteorological parameters are illustrated in FMI View Service. The View Service is also used to provide radar data. The radar images are provided as 16-bit black and white GeoTIFF images and ready colored 8-bit PNG images.
View service is based on GeoServer[2]. The View Service is clustered with three physical servers and two GeoServer instances per server. The instances uses common database, NFS file system and configuration directory. The cluster is configured with a seprate configuration server (figure 1).

1 http://geonetwork-opensource.org
2 http://geoserver.org/

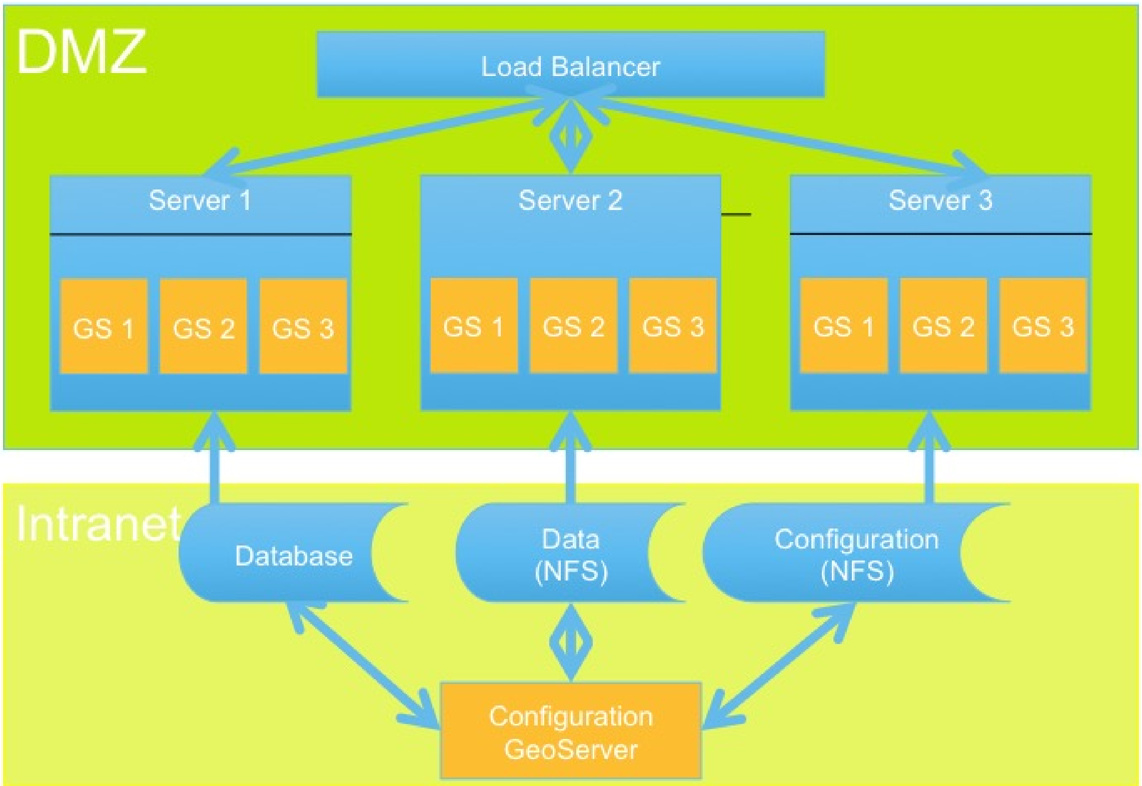


Figure 1: View Service clustering setup.

2.3 The Download Service

Download service is implemented as WFS 2.0 Simple Profile data server. The service is based on FMI data server. FMI data server uses memory-mapped files as a forecast data source and database as a observation data source.

As the service provides only WFS 2.0 Simple Profile, standard filters are not supported. Instead, data sets are fetched utilising stored queries with additional parameters such as spatial and temporal limitations, number of included measurement stations and measurement time inverval.

Basic WFS is currently under development. The biggest challenge in implementing standard filters are that INSPIRE specification does not distinguish technically different kind of data sets like observations and forecasts. Filtering need to be hence done for both database and file based data.

2.4 Provided data models

Data models are derived from INSPIRE specifications although some extensions were required. Whenever applicable, Observation and Measurement (O&M) (Earth Observation Metadata profile of Observations & Measurements, 2012) guidelines are followed. INSPIRE specifications require that data are represented as measurement time series[3]. The data model is quite intuitive and works well with standard XML tools like xPath. On the other hand it causes large file sizes and it is heavy for especially DOM-based parsers. As the FMI Open Data Portal implementation goes beyond INSPIRE requirements, additional data models were also provided for better performance and compliance.

Measurement time series (WaterML 2.0: Part 1- Timeseries, 2012) contains one

3 Schema available at http://inspire.ec.europa.eu/schemas/omso/2.0/SpecialisedObservations.xsd

member for one parameter and location. This relatively verbose data format does not provide good performance for meteorological large volume data sets. Thus grid series observation[4] with multipoint coverage was added for all meteorological data sets. Figure 2 represents the data format in multipoint coverage. The data model is compact and works well for any kind of data types. Still, it's not very intuitive for users and XML tools can not be used to quarry data values from the coverage.

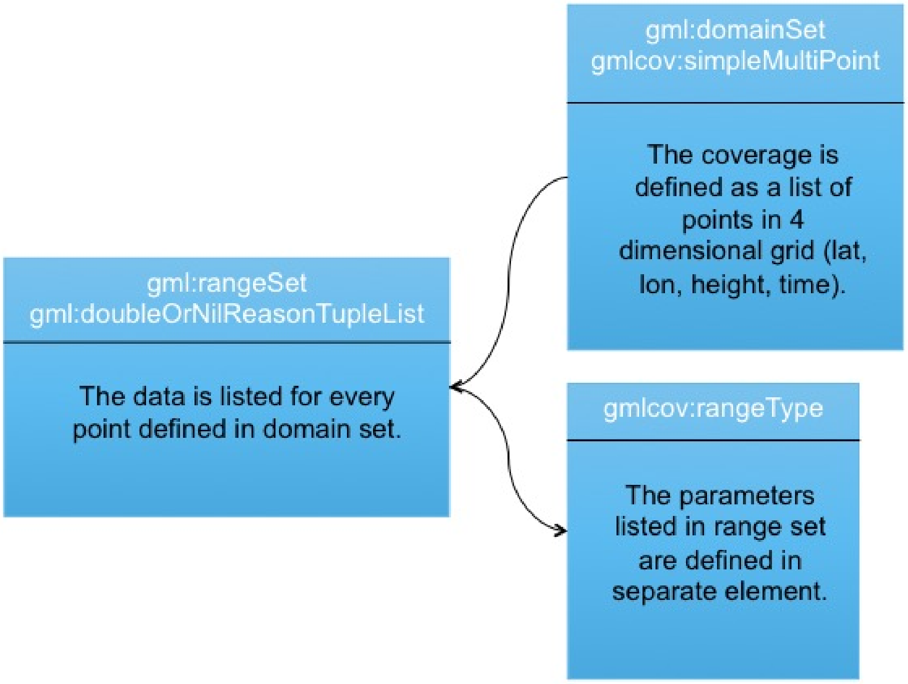


Figure 2: Multipoint coverage data model.

Despite the standard based services and data models, a significant gap between provided data model and clients' capabilities can be indicated. Hardly any clients supports complex features used in INSPIRE specifications described above. Thus, a simple feature[5] data model was added to FMI Download Service to support most common clients.
Figure 3 illustrates a file size comparison between different data models. The comparison is done with a quite typical meteorological data set with 138 observation stations (covering spatially approximately one third of Finland), 11 meteorological parameters and 12 hours of data. As can be seen, the differences are remarkable. It is however notable that XML based documents can be compressed very well and large file sizes causes hence only problems for the clients.

4 Schema available at http://inspire.ec.europa.eu/schemas/omso/2.0/SpecialisedObservations.xsd
5 Schema available at http://xml.fmi.fi/schema/wfs/2.0/gml/3.1/fmi_wfs_simplefeature.xsd

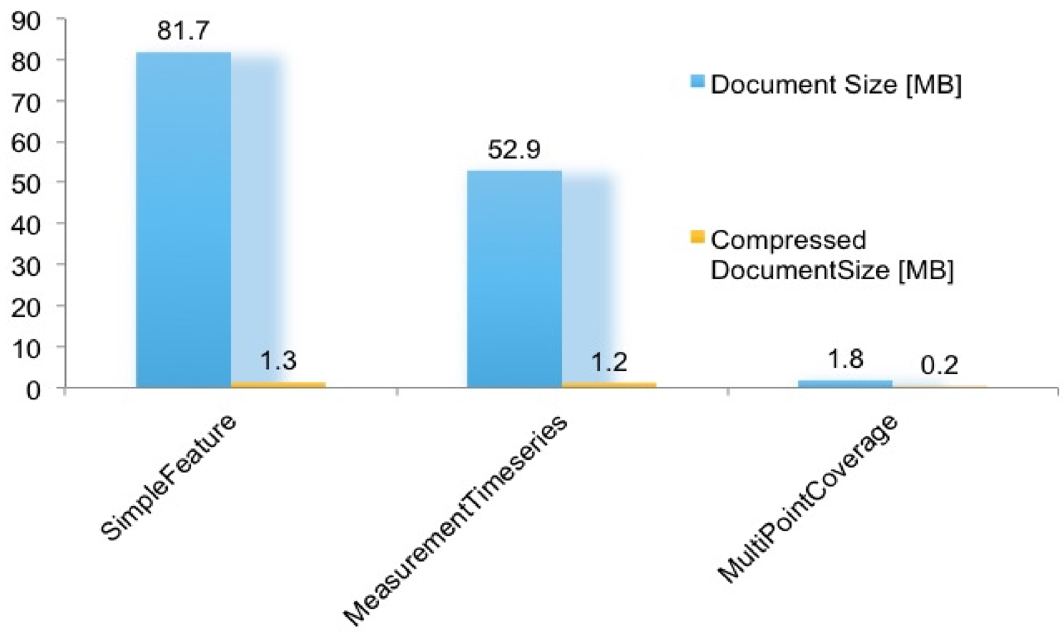


Figure 3: Multipoint coverage is the most efficent data model in compraison.

For very large data sets, ASCII-based XML representation is not applicable at all. In those cases WFS responses contains only 'a WFS envelope' with metadata required by O&M guidelines and link to actual data.
The METAR-messages generated by Finland's aviation airports are distributed in XML-format through WFS. The distribution format is ICAO Meteorological Information Exchange Model (IWXXM) version 1.0 which also offers the XML-format to SPECI-, TAF- and SIGMET-messages. IWXXM products are used for operational exchanges of meteorological information for use in aviation. The benefits being increase of flexibility for information sharing. IWXXM is step towards the envisaged System-wide Information Management (SWIM) environment. Within aviation the concept of SWIM – System Wide Information Management - covers a complete change in paradigm of how information is managed along its full lifecycle and across the whole European ATM system. In this context global interoperability and standardisation are essential and SWIM is expected to be an important driver for new and updated standards.

3 User Management

A registration function was implemented for the usage of View and Download Services. Working email address is the only required field for the registration. Upon registration users are issued an identification key (APIKEY) which is required to every view and download service request. Registration requires acceptance of the open data license. The licence is currently Creative Commons 4.0. One can register several APIKEYs with one email.
As the FMI Open Data Portal is mainly designed for data exchange, some usage limits are declared to prevent misusage of the services. The limits are described in table 1. Download Service limits are designed so that user can download all newly published data near realtime. View Service is designed only

for browsing the data. Usage limit for both services is to prevent flooding and is based both on APIKEY and ip-address of the user.

Download Service	20 000 req/day
View Service	10 000 req/day
Both Services	600 req/5 minutes

Table 1: APIKEY based usage limits.

It is notable that per day usage limits are not automatic. If user exceed allowed usage, her account is not immediately shut down but she is contacted personally (via email) and offered a premium account.
Better service levels and premium functionalities can be offered from other service clusters. User can be directed to different service clusters based on APIKEY as illustrated in figure 4.

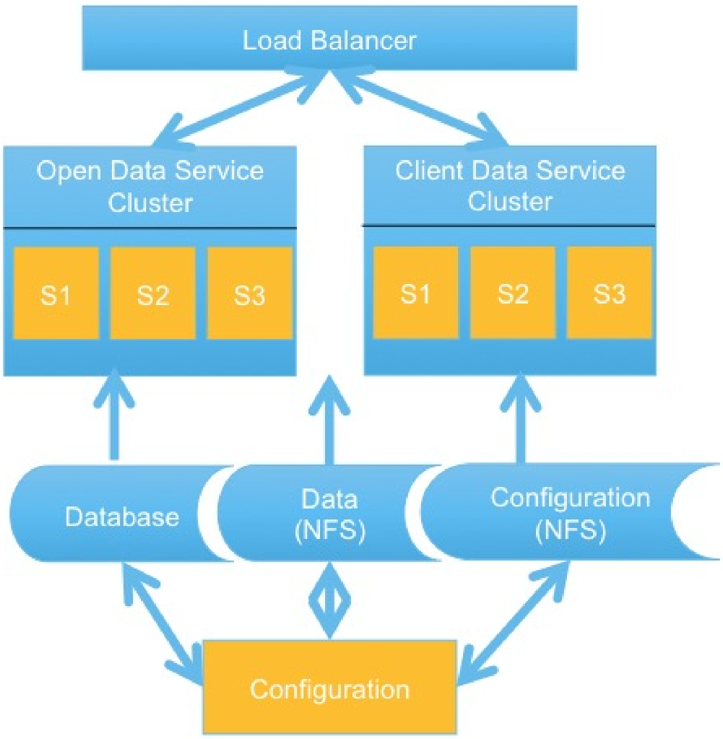


Figure 4: User can be directed to different service clusters based on APIKEY.

FMI offers advisory and consulting service related to Open Data interface and the datasets. A Help Desk function for open data was created as a part of the project.

3.1 MetoLib-Open Source Library

FMI produced also Open Source JavaScript library to help developers to load and use the data. The library contains functionalities to load data efficiently and provides client-side cache for the data. It also provides some easy examples how to use the library, the data interface and the data itself. Contribution to this is possible and it can be found from GitHub (https://github.com/fmidev).

Some additional recourses were published as well. Among them are in example open data logo, weather symbols for weather forecast data and SLD stylesheets for colouring radar images.

3.2 Open Data Impact

The most significant impact and benefit of the data opening is new business and innovation opportunities from the freely available open data from FMI. Other impacts of the open data release are enhanced use of spatial data, increasing inter-agency co-operation and creating new services for citizens. Technical co-operation between agencies and other parties will also be enhanced in the future, as data can be produced in a standardised format for applications.

The internal operating culture at FMI was also developed as part of the open data project, including improvements such as the development of FMI's in-house production system, the adoption of new types of operations such as consulting, co-operation with external developers and the production of open source code. Customer-related activities were also streamlined and enhanced in some respects.

The open data release created new forms of operation for the FMI. The new operating culture has got off to a good start, but development work will continue.

The FMI's reputation also benefitted from the open data release both in Finland and across Europe as the implementation was extensive and successful. The first application based on the open data was created in less than 24 hours of the beta version launch of the download service.

3.3 Apps4Finland

FMI was one of the organising partners of the 2013 Apps4Finland competition and a supporting partner in 2014. The aim of the competition is to promote the diverse and extensive use of released data sets in wider contexts. One objective to participate was also to make our data known to a more broader public and especially to software developers.

3.4 Application examples

Aaltopoiju.fi is a great example of clever data visualization. It show near realtime observation data from wave buoys located in the Baltic Sea. Other parameters include wind, sea level, temperature from other locations. Forecasts include wave, wind and sea level. A mobile site is also provided. Aaltopoiju.fi was the winner of FMI's challenge reward in Apps4Finland 2014 competition.

rOpenGov (https://github.com/rOpenGov/fmi) is an open data API R client for the FMI open data interface. It provides tools to access open data from FMI with the R programming language favored by the research community.

References

- ✔ Earth Observation Metadata profile of Observations & Measurements. (2012). Retrieved March 5 2015 from https://portal.opengeospatial.org/files/?artifact_id=47040
- ✔ COMMISSION REGULATION (EU) No 1253/2013 of 21 October 2013 amending Regulation (EU) No 1089/2010 implementing Directive

2007/2/EC as regards interoperability of spatial data sets and services. (2013). Retrieved March 5 2015 from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:331:0001:0267:EN:PDF

✔ OGC® WaterML 2.0: Part 1- Timeseries (2012). Retrieved March 5 2015 from https://portal.opengeospatial.org/files/?artifact_id=57222

Towards a Virtual Hub for a wider Open Data community

Mattia Previtali[1], Luigi Barazzetti[1], Raffaella Brumana[1], Branka Cuca[2]

[1] Politecnico di Milano, Department of Architecture, Built Environment and Construction Engineering, Via Ponzio 31, 20133 Milano, Italy

[2] Cyprus University of Technology, Dept. of Civil Engineering and Geomatics, Saripolou str. 2-8, 3036 Limassol, Cyprus

Abstract

The world of geographic information is extremely heterogeneous, especially in the field of Open data. System requirements are too various to be supported by a single unique system or technology limiting or impeding their interoperability. The proposed innovative approach facilitates the access to open data in a homogeneous way by means of broker connecting different data sources. An application integrating ancient cadasters and modern cartography is presented showing the flexibility of approach.

Keywords

Open data, brokering approach, interoperability, web services

1 Introduction

The world of geo-information (GI) is extremely heterogeneous. Different actors (Public Administrations, Research Centers, Citizens, etc.) with different roles (data creators, data providers, data users, etc.) are involved in the delivery and use of geo-information in different domains. This aspect is a significant barrier for the wide use of GI. Indeed, different users have different requirements in terms of communication protocols and technology, limiting a fruitful utilization of Open Data (OD).
In recent years, many important programs and initiatives have tried to address the ambitious objective of allowing a wider use of OD for production of services to the citizens in diverse communities and contexts. These programs recognize the needs of three main stakeholder types:

- Users;
- Data/Information Producers;
- Cyber(e)-Infrastructure Architects and Managers.

Examples of these initiatives are INSPIRE Directive established for geo-information of European Public Authorities, GMES/Copernicus core and downstream services for environmental and security applications from European Earth Observation data, the Global Earth Observation System of Systems (GEOSS) and the Eye on Earth and Shared Environmental Information System (SEIS) initiatives of the European Environmental Agency (EEA).
Generally, applications using OD are commonly achieved by implementing a stateless Client-Server (C-S) architectural style. In this style, server components offering a set of services, listen for requests upon those services, while client components desiring a service send a request to the specific server via its interface (Fielding, 2000). Almost invariably programs for OD sharing resemble the Service Oriented Architecture (SOA) archetype. In this architecture users play the role of service consumers; data producers represent service providers, while the cyber-infrastructure acts as a service clearinghouse or service registry, which is generally kept as simple as possible.
In a SOA approach, a cyber-infrastructure consists of a set of few selected common international standards (i.e. Web services protocols), common data models, and rules (e.g. best practices) adopted by both Data Producers and Users to publish available resources, discover

them and bind them. In other words, to enable interoperability C-S style, interfaces are objects of standardization at different levels, e.g. at the system, application, or infrastructure levels. The collection of these interface standards defines the interoperability arrangement (Ortiz, 2007; Schmidt, Hutchison, Lambros, and Phippen, 2005). Architectures implementing this pattern are known as Federated Architectures. In fact, they are based on the concept of "common/federal data model" and "common federal interoperability protocol(s)". Most Spatial Data Infrastructures (SDIs) around the world are based on this model. They use the SOA concept to mandate a limited set of standard specifications using a specific model to interconnect system components (i.e. clients and servers). Federated systems address interface heterogeneity by pushing common standards adoption.

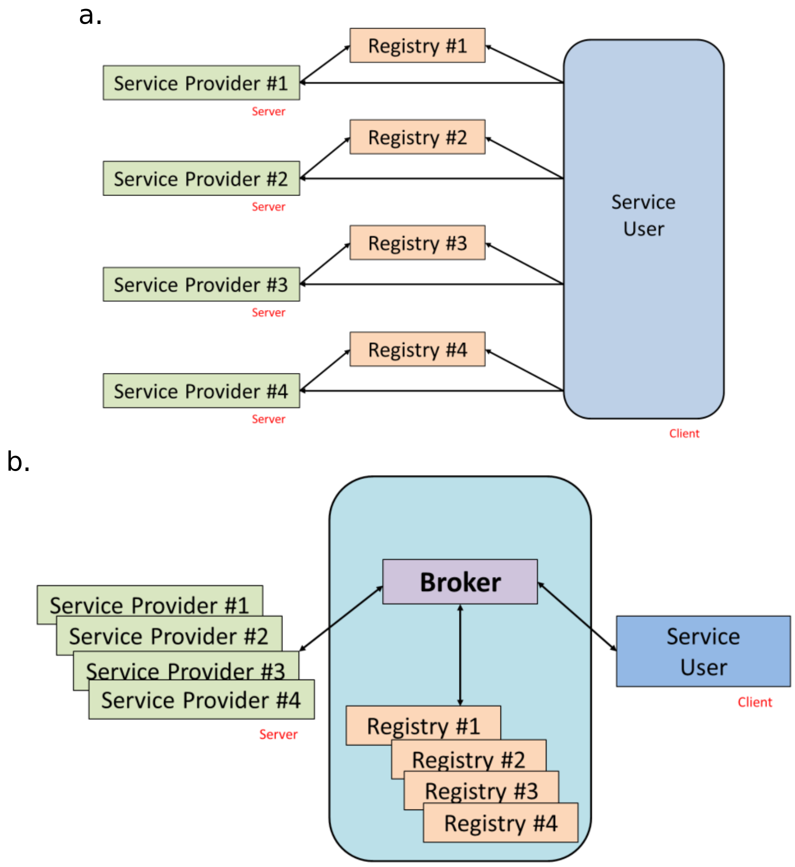


Figure 1: The difference between a Service-Oriented-Architectures-SOA approach(a) and a brokering approach (b).

The SOA approach works very well in consistent and controlled frameworks (e.g. enterprise environments), or where the approach is embedded in a strong legal framework that makes it mandatory for stakeholders to adopt the agreed standards and protocols (e.g. the NSDI for US federal government, and INSPIRE in Europe). However, each community generally tries to develop its own SDI using its own standard according to its specific requirements (Fig. 1). Indeed, it is noteworthy that the lack of agreed interoperability standards is not due to removable barriers, such as attitude, legal, financial or technological barriers. It is inherently impossible to define a common standard which is good for all the possible systems handling the geo-information. Even assuming that such a common standard suite would be developed, it would be too general for an effective utilization in real systems.
This is one of the reasons why, although several very good standards have been developed in the recent years (e.g. from OGC, ISO, etc.), they are not adopted by all the different domains. For this reason, the opportunity to access and integrate multiple sources for the development of new applications requires a work on different standards and protocols, which may be a quite complex and expensive task. The European project ENERGIC OD (European NEtwork for Redistributing Geospatial Information to user Communities - Open Data) tries to partially cope with the previously listed problems by developing a mediation system based on a brokering framework. Its aim is to lower the entry barriers, represented by the need to implement difficult and often unknown standards, to both providers and users for a wider use of OD.

In this paper the general concept of Virtual Hub adopted in ENERGIC OD is presented (Section 2) and an application of this approach integrating ancient cadasters and modern cartography will be presented (Section 3), showing the great flexibility of the brokering approach to access multiple and heterogeneous data.

2 ENERGIC OD: the brokering approach

As previously observed, the large variety of standards and protocols adopted by different data providers is an important barrier for an effective utilization of OD for production of services. A solution able to reduce the interoperability problems between data providers and applications requires extending the SOA archetype by introducing a new component that interconnects the different service protocols and standards, mediating their models and interface specifications. In addition, this component must work out all the necessary distribution and virtualization capabilities to lower the entry barriers for multidisciplinary applications. In ENERGIC OD, this intermediary role is played by an infrastructure called Virtual Hub, realizing a Brokering approach (Nativi, Craglia, and Pearlman, 2012).
The next subsections tries to define the primary requirements for an effective Virtual Hub implementation (Subsect. 2.1) and the main functionalities of the broker (Subsect. 2.2).

2.1 Interoperability issues

In recent years, several initiatives have been launched to improve the availability of geospatial information. These initiatives are developed not only by public agencies (e.g., INSPIRE, GMES/Copernicus, GEOSS) to provide a single point of access to global Earth observation resources and providing the framework for public authorities to publish themed data, but also informal "Volunteered GI" communities (e.g., OpenStreetMap with contributors; citizens and mapping parties), the professional geomatic community and open data networks (portals, "Hackathons" etc.). A valuable brokering framework should be aligned to this initiatives. This requires a connection with data sources that are already compliant with these initiatives in terms of standards and communication protocols. A trivial example is given by data formats: while in the word of professional geomatic community and PA a common data format for vector data are .dwg and .shp ESRI© files, in the word of mobile application development GeoJson and KML files are more popular. A brokering architecture should be able to discover both data sources to allow easy interoperability between different application fields and technologies. For the same reason, a format transformation module is needed. For example, accessibility to a data discovered as .shp should be given as KML or vice versa. When facing this problems the following issues should be addressed:

- Data harmonization, i.e. providing access to spatial data through network services that allow combining it with other harmonized data, using a common set of data product specifications;
- Interoperability of available public data, i.e. the possibility for spatial data sets to be combined and services to interact without repetitive manual intervention, so that the result is coherent and the added value of the data sets and services is enhanced ;
- Spatial inconsistency, i.e. different reference systems and their fragmentation and cross-border geospatial dataset integration at national, trans-regional and trans-national scales;
- Multi-scale data integration: consistency between data with varying levels of detail; multi-resolution data and various levels of processing.
- Consistency of sets of semantic: in many cases similar concepts are designed in a slightly different way by different communities and data sources. An homogeneous access needs to be ensured and guaranteed.

All these requirements needs to be taken into account in the definition of a proper brokering architecture. The following subsections shows how this requirements are taken into account in the definition of the architecture of the brokering service in ENERGIC OD.

2.2 The Brokering Architecture

In a traditional C-S architecture, a Server provides services to the Client and each Client sends requests to the Server. In the case the same interface(s), i.e. the same application, the need to access multiple services and the interaction with these services have to be implemented in a separate way. This means that multiple standards and protocols may be implemented by the

application developer.
The brokering-oriented architecture changes this architectural style by requiring the use of a specific intermediary: the brokering layer. The broker de-couples the server and the client layers allowing the C-S separation that is crucial for accessing multiple services and Open Data developed by specific user communities. On the other hand, the effort of the application developer to access heterogeneous sources is much reduced. In fact, the brokering architecture moves the implementation of standard and protocol mediation to the brokering layer.
The C-S decoupling simplifies the interoperability requirements for both server and client components, improving scalability and flexibility. Client and server components can evolve independently and interoperability problems are solved inside the mediation layer. In particular, clients can access to the OD in the same way regardless the Server interface changes. The brokering architecture, facilitates architecture evolution of both clients and servers since all protocols and interface changes are solved by the broker. Even if the interface of a Server changes significantly, the Client access of data remain unchanged. Giving always the same interface a Virtual Hub approach give the chance to solve many interoperability problems allowing also the reuse of OD, which is one of the major concern of Public Authorities producing and delivering OD. A broker encapsulates legacy interoperability arrangements lowering entry barriers for client and servers. The Brokering-oriented architecture is a solution to implement interoperability for large scale and complex systems. For those cases a complete knowledge of all sub-systems and their services would be prohibitively expensive for an application developer.
As previously anticipated, a broker is a mediator which implements some services (i.e. discover, access, distribution, added value, etc.) to facilitate the interconnection between client and server components in a Client-Server architecture. In particular, the broker which will be developed in ENERGIC OD project will present the following functional modules (Fig. 2):

- Client request distribution in asynchronous way. This implementation is performed to allow a client to distribute requests for searching resources across more data providers. One of the major concerns is the waiting time for a query. Indeed, one of the primary disadvantage of a brokering systems is the user-perceived performance reduction (Fielding, 2000). One of the strategies to increase performance is applying shared caching at the brokering to replicate the result of an individual request such that it may be reused by later requests. Discover of OD is enabled by specific services (e.g., search engines) and by appropriate ancillary information (i.e., metadata profile ISO 19115) associated to the resources, independently from their structure.
- C-S interface protocol adaptation. This includes functionalities to match many server protocols and client ones, and vice-versa. Access services support electronic data retrieval, usually based on spatial constraints and other criteria, in forms useful for client-side processing. Supported backend for data sources access includes: OGC standard services (e.g., WMS, WFS, WCS, SOS, CSW), ESRI ArcGIS Geoportal catalog service, THREDDS, etc. In the frontend the broker can be accessed by using a JavaScript API. This API exposes the brokering functionalities (semantic discovery and dataset access) hiding its complexity. In this way an easy programming for most common operations is allowed. However, full interfaces and options to the broker are available.
- Added-value functionalities at the C-S interconnection level: The ENERGIC OD broker will implement advanced/semantic discovery. In particular, a semantic broker will automatically looks for synonyms during a search so that researches are extended. In particular, two different strategies will be used: (i) automatic query expansion, keywords are expanded automatically by the broker by interrogating a set of aligned semantic instruments (typically, controlled vocabularies, thesauri, gazetteers, and ontologies). The final result consists of a set of semantically related queries that are all executed by the discovery broker. In the (ii) user-assisted query expansion, i.e. the second discovery style, we applies the same strategy, but the semantically-related terms, retrieved from the aligned semantic instruments, are presented to users who are allowed to browse a graph developed according to these terms and select the most pertinent terms.
  Transformation module will solve for spatial inconsistencies and access data having different CRS and perform data integration in the same CRS.
  Other application services may concern results clustering resources ranking according to unified metrics, etc.
- Configurability at the C-S component level. This part allows the customization of the brokering services for each brokered client or service component.

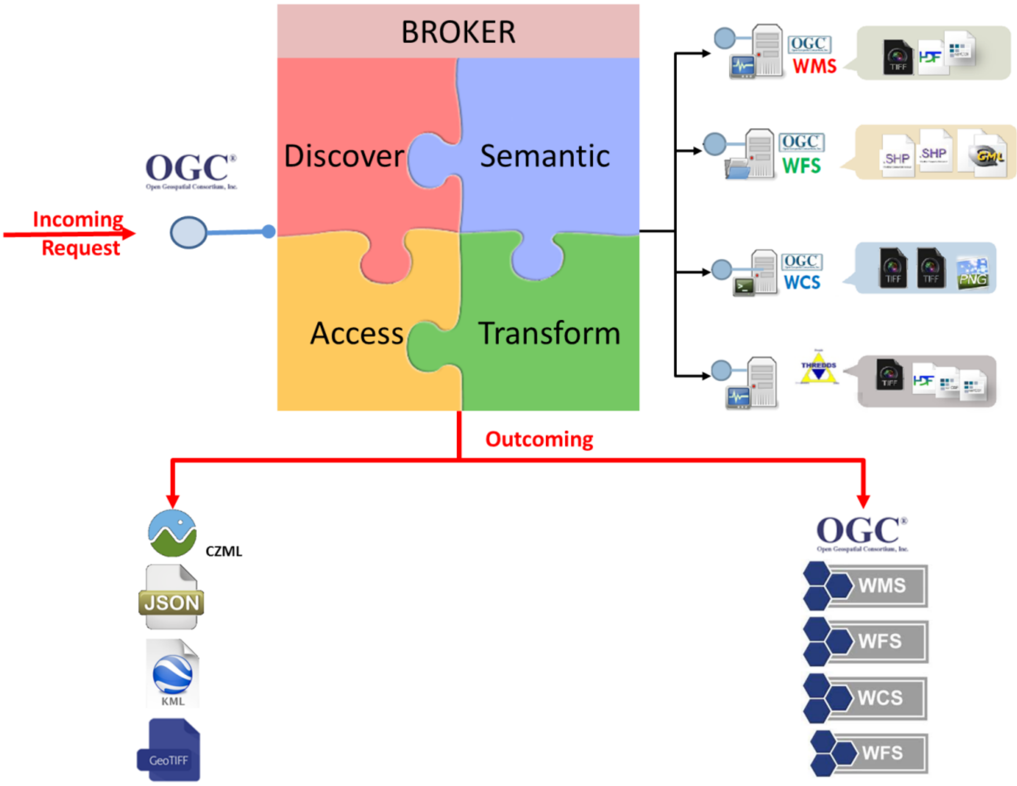


Figure 2: The broker architecture for heterogeneous data discovery and access.

As discussed, the introduction of the Brokering layer shifts the C-S architecture from a two-tier (Client-Server) to a three-tier (Client-Broker-Server) configuration. In this schema the broker may became a single point of failure. In particular, updating, governance and maintenance of the brokering are a significant problem to be evaluated. However, this is a common problem when complexity is shifted from Users and Resource Providers to the infrastructure/platform.

3 GeoPAN: development of an application based on the Virtual HUB

In the ENERGIC OD project the GeoPAN application is designed to prove the feasibility of a brokering approach for an effective wider utilization of Open Data. The purpose of this application is to combine management of historical maps in Lombardy Region (Italy) with local cartography delivered by local authorities (e.g., municipalities, regions, etc.).
Interest in the digital management of historical maps on the web has recently increased. Indeed, old maps can be of invaluable importance for a much wider range of applications, such as landscape change analysis (Azzari, Marcaccini, and Pizziolo, 1999), urban development studies (Bitelli and Gatta 2011), territorial planning (Oreni, Brumana, Scaioni, and Prandi, 2010), cultural and environmental heritage (Cuca, Brumana, Scaioni, and Oreni, 2011; Brovelli, Minghini, Giori, and Beretta, 2012) and archaeological research (Guarducci, Rombai, and Piccardi, 2011). As organizations started to upload their historical digital archives on the Internet, a series of problems arose regarding the ways data have to be processed, published and web-cataloged, leading to an established and shared process related to the utilization of ancient maps in digital format.
Modern developments in digital technologies allow a new and previously inconceivable utilization of historical maps allowing to exploit the old maps' metric information in a GIS or webGIS environment (Brumana, Oreni, Alba, Barazzetti, Cuca, and Scaioni, 2011).
The purpose of this study is to test the use of Free and Open Source software to achieve the following objectives: (i) create a webGIS for providing access to the historical maps, (ii) integrate OD in the application to enrich old historical map with modern cartographic information provided by local authorities and information for tourists (e.g., itineraries, historical information, information about point of interest), and (iii) validate the feasibility of the

brokering framework inside a Virtual Hub to access heterogeneous data. The innovation given by the brokering framework to access heterogeneous data sources in an easy way is of major importance for the GEOPAN application. Indeed, in Lombardy local authorities shares OD of local cartography using different formats and services. Indeed, while some municipalities rely on WMS and WFS, others publish ESRI© shapefiles. The Virtual Hub may give a single point of access to these data in a homogeneous way.
The results of this research should provide common map users with a new and easy way of consulting historical maps, getting the related information and stimulating tourism. The high modernization level of tools and the archive availability through the web should also enlarge the potential public with respect to the current one.

3.1 Data description

The historical cartography used to perform and evaluate the previously mentioned procedure comes from the huge collection of the Archivio di Stato di Milano and partially Archivio di Stato di Como, whose maps belong to four distinct cadastral productions: (1) the Theresian Cadaster of XVIII century; (2) the Lombardo-Veneto Cadaster of mid-XIX century and (3) its updates of 1898 (Signori, 2002); together with samples of (4) 'Impianto in conservazione' by AdT (Italian Cadastral Administration). For each cadastral series, maps are divided in sheets and are related to single municipalities. Other historical map series considered here are the 'Corografie delle Province del Regno Lombardo-Veneto'(from 1836, in scale 1:115,000), the 'Carta del Territorio del Milanese e Mantovano' (1788-1796, in scale 1:86,400), the 'Carta del Regno Lombardo Veneto' (ITM 1933, in scale 1:86,400).
Local authorities (e.g., municipalities and region) are also producing cartographic information. However, while some authorities share them as Web OD by using either OGC services or ESRI© shapefiles, others do not share cartography online. For this reason, in ENERGIC OD a survey of available cartographic information will be performed to derive a catalogue of geographic OD in Lombardy Region. This catalogue will be used as input for the Virtual Hub.

3.2 Map georeferencing

Maps were georeferenced using the procedure described in (Oreni et al. 2010). Here only the main aspects are recalled.
Each sheet was georeferenced individually with respect to the chosen grid mapping system, (Gauss-Boaga – datum Roma40 – for local cadastral maps, and UTM – datum WGS ETRF89 - for regional-territorial maps) by using an affine model. This model is a global approximate interpolator linking ground coordinates (X,Y) to image coordinates (x,y). Ground coordinates were derived from the Spatial Databases of current municipalities, which are composed of a large number of vector layers and are available at scales between 1:1,000 and 1:5,000. Ground points should be well spread in the entire raster images (Peck, and Devore, 2011). An example of point collimation on the historical cadastral map and the corresponding current cartography is shown in Figure 3.
It is very difficult to estimate and evaluate the analytical reliability, due to the non-homogeneous distribution of the error in the georeferencing process. Indeed, the georefencing process is not only influenced by the accuracy of the two data sets but also to the scanning of maps, the deformation occurred overtime in the ancient maps, and the individuation of the persisting ground features overtime.

a.

b.

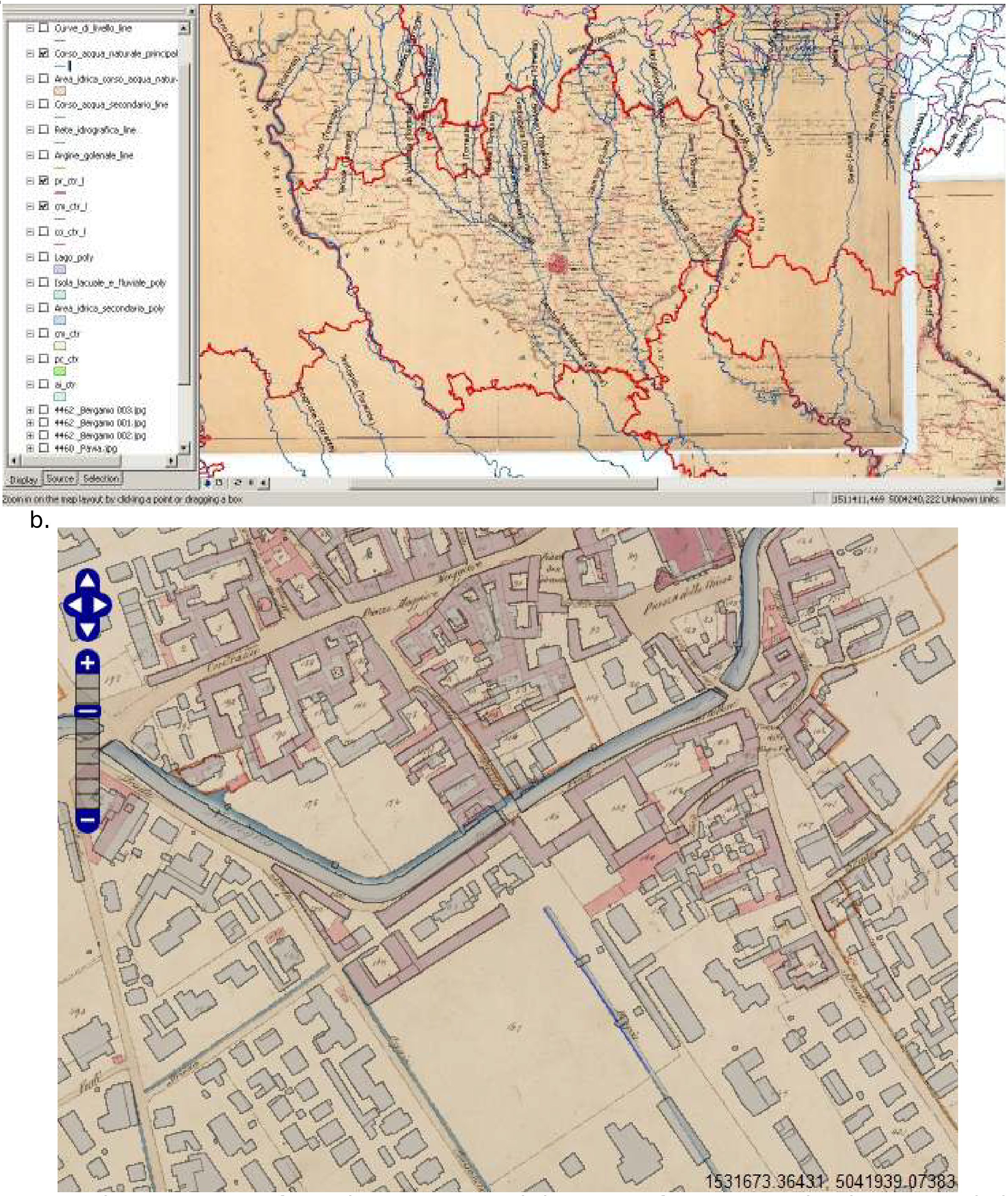


Figure 3: Ancient map georeferencing process and the vector features used to locate persisting control points.

Obviously, quantitative results of georeferencing depend on the selected number and distribution of ground points on each map (Boutoura and Livieratos, 2006). However, in the case points are well distributed in the area, the quality can be evaluated by the obtained RMSE. To have a general idea about the performance of the method it is useful to consider two extreme cases: the city center of Milan represented by the Astronomi di Brera map (39 sheets), and the 'Corografie delle Province del Regno Lombardo-Veneto'. The difference between the two is given by the fact that while in the first case it is possible to identify a large number of ground points in a very small area, while in the second case less points could be selected. In this second case, it was evident that it is impossible to use the center of the inhabited places, whose position was simply indicative, or rivers and lakes, for which it is impossible to individuate with sufficient precision correspondent ground control points, since the ancient river beds could have varied in position. The only elements that were represented in a rather precise way, being this the scope of the map, were the provincial and municipal boundaries.

3.3 Implementation of GeoPAN

A webGIS platform will be created to publish the georeferenced historical maps along with local cartographic data. As things stand at the moment, some tests and a proof of concept of the application were developed.

During the developing the web application, choices were made by paying attention to the needs and peculiarities of the project: visualization and management of large dimension raster maps, organization of cartographic resources for easy consultation, and the need of a visual impact to attract people and tourist.

Different FOSS solutions were considered for both server and client-side implementation. The server-side software used is GeoServer (http://geoserver.org), a powerful Open Source platform for publishing spatial data and interactive mapping applications on the web. GeoServer allows the visualization of produced Open Data (like touristic itineraries) through the use of an OGC standard Web Map Service (WMS). The time needed for loading and navigating these data represents the biggest challenge in accessing and displaying large raster maps in an efficient way. For this reason historical maps are released by using WMTS which is a standard protocol for serving pre-rendered georeferenced map tiles on the Internet. Indeed, for most WMS services it is not uncommon to require 1 or more CPU seconds to produce a response. In the case of massive parallel use, such CPU-intensive service is not practical. To overcome the CPU intensive on-the-fly rendering problem, WMTS allows for using pre-rendered map tiles.

In the general architecture of the service the Virtual Hub will be used to access local cartographic information. Figure 4 shows two possible architectures of the GeoPAN app with and without the Virtual Hub. The architecture with the VH gives a simpler access to geographic OD giving a single point of access with a standard protocol and interface.

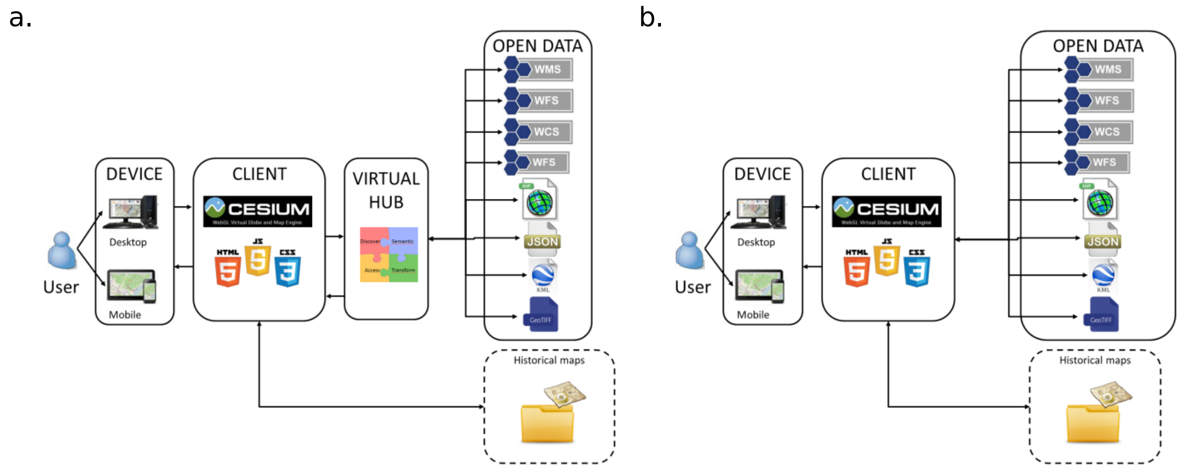


Figure 4: Two possible architectures for the GeoPAN app: (a) with and without (b) the Virtual Hub. An architecture with the Virtual Hub all interoperability problems are solved by the Virtual Hub. An architecture without the Virtual Hub would require to solve several interoperability problems at client side.

To develop the client-side of the system Cesium (http://cesiumjs.org). Cesium is a JavaScript library for creating 3D globes and 2D maps in a web browser without a plugin. It uses WebGL for hardware-accelerated graphics, and is cross-platform, cross-browser, and tuned for dynamic-data visualization. Cesium is supported by Android, iOS and Window tablet. It is delivered under the Apache 2.0 license, which means Cesium is free for commercial and non-commercial uses. Today Cesium support the following geospatial features:

- Map layers: WMS, TMS, WMTS, OpenStreetMap, Bing Maps, ArcGIS Map Server;
- Global terrain: quantized mesh, ArcGIS Image;
- Vector data: CZML, GeoJSON, TopoJSON, KML (announced for March 2); and
- 3D models: glTF (on-line and off-line converter from COLLADA available).

The great advantage of Cesium is the opportunity to visualize maps and contents over a Virtual Globe. It gives an interesting approach to the old maps and can attract more the people's attention in a way similar to Google Earth©. Cesium allows also great freedom in developing

the web application structure and layout (Figure 5).
As the proof of concept is the actual implementation, the graphical interface of GeoPAN is under development. The GeoPAN web client consists of two parts: a layer menu and a main map panel. The layer menu itself is additionally divided into two parts: the first one allows users to view and select all the georeferenced historical maps which are ordered in two separate trees, the second one will be developed to interact with the Virtual Hub formulating queries and displaying results.

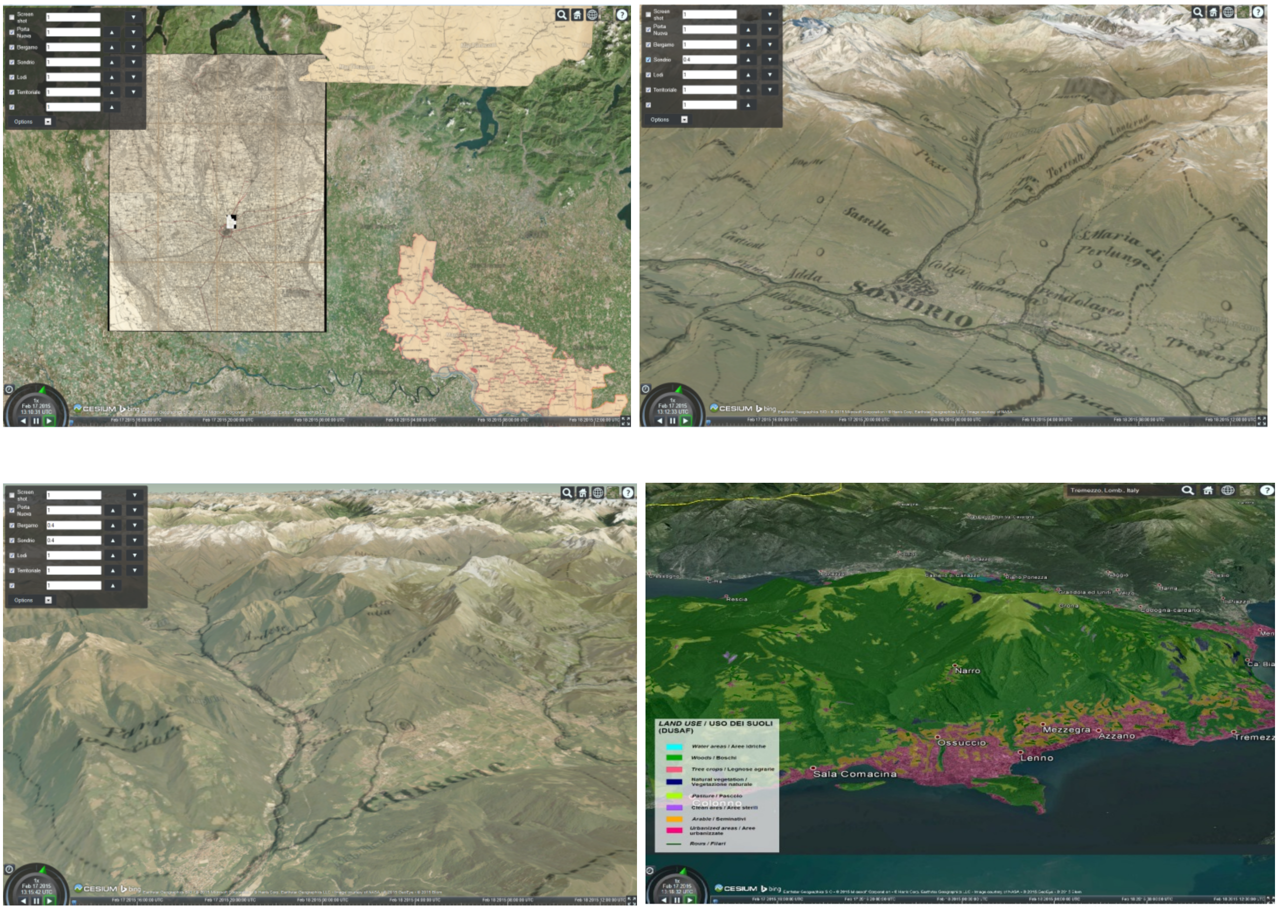


Figure 5: Some ancient map georeferenced in GeoPAN interface and a land use layer.

4 Conclusion

ENERGIC OD tries to lower User and Data Producers entry barriers for fruitful exploitation of OD for service production to citizens and Public Administrations. To achieve this goal, the brokering approach adopted by Virtual Hubs can offer the opportunity to support existing and independent infrastructure and OD initiatives (supplementing and not supplanting them). In the brokering approach the data heterogeneity is faced introducing a mediator which separates the client and server layers implementing brokering services (i.e., mediation, distribution, added value, etc.) and facilitating the interconnection between client and server components. What is most important in moving toward a Broker-based infrastructure is the greater level of flexibility than other architectural solutions. This gives the possibility to integrate not only traditional geographic information, but also cultural, social and economic information. The brokering-oriented style was successfully implemented and demonstrated by the EuroGEOSS project. This experience has shown that the Brokering approach is mature to discover and access information useful in the scientific world without the need of scientific users to implement standard and protocols. The aim of the ENERGIC OD project is to demonstrate that the same approach can be used also for the wider community of web service developers. In particular, this architecture may help them in accessing OD using a single interface. GeoPAN tries to prove this concept. The decision of publishing historical maps was taken with the aim of providing a synthetic and continuous vision of the territory, at different historical levels, in order to allow different users to make immediate comparisons between the shapes and the characteristics of the landscape of yesterday and today. To this aim, current cartographic information accessed with the broker is of primary importance. This service may also convey touristic information along with ancient maps giving the opportunity to attract tourist to the different landscape levels and providing virtual and on site itineraries. Unfortunately, only few municipality allows for publication of geospatial OD.

Acknowledgements

The research leading to these results has received funding from European Commission Competitiveness and Innovation Framework Programme (CIP) under grant agreement n°620400 for the project ENERGIC-OD (European NEtwork for Redistributing Geospatial Information to user Communities - Open Data).

References

- ✔ Azzari M., Marcaccini P., and Pizziolo G. (1999). A geographical information system in Tuscan coastal wetlands. In Actes du II Congres International sur la Science et la Technologie pour la Sauveguarde du Patrimoine Culturel dans les Pays du Bassin Méditerranéen. Paris, Elsevier: 1189–92.
- ✔ Bitelli G., and Gatta C. (2011). Digital processing and 3D modelling of an 18th century scenographic map of Bologna. In Ruas A (ed) *Advances in Cartography and GIScience* (Volume 2). Berlin, Springer: 129–46.
- ✔ Boutoura, C., & Livieratos, E. (2006). Some fundamentals for the study of the geometry of early maps by comparative methods. *E-Perimetron*, 1(1), 60-70.
- ✔ Brovelli, M. A., Minghini, M., Giori, G., and Beretta, M. (2012). Web geoservices and ancient cadastral maps: the Web CARTE Project. *Transactions in GIS*, 16(2), 125-142.
- ✔ Brumana, R., Oreni, D., Alba, M., Barazzetti, L., Cuca, B., & Scaioni, M. (2011). Panoramic UAV views for landscape heritage analysis integrated with historical maps atlases. In XXIIIth CIPA Symposium, Prague, Czech Republic, September (pp. 12-16).
- ✔ Cuca B., Brumana R., Scaioni M., and Oreni D. (2011). Spatial Data Management of Temporal Map Series for Cultural and Environmental Heritage. *International Journal of Spatial Data Infrastructure Research (IJSDIR)*, Selected Articles from INSPIRE 2010, Vol. 6, pp. 95-127.
- ✔ Fielding, R. T. (2000). Architectural styles and the design of network-based software architectures (Doctoral dissertation, University of California, Irvine).
- ✔ Guarducci A., Rombai L., and Piccardi M. (2011). Mare Oraque Tusciae. *E-Perimetron*, 6(2), 114–21.
- ✔ Nativi, S., Craglia, M., & Pearlman, J. (2012). The brokering approach for multidisciplinary interoperability: a position paper. *International Journal of Spatial Data Infrastructures Research*, 7, 1-15.
- ✔ Oreni D., Brumana R., Scaioni M., and Prandi F. (2010). Navigating on the past, as a bird flight, in the territorial scale of historical topographic maps: WMS on the"Corografie delle Province del Regno Lombardo-Veneto", for accessing cadastral map catalogue. *E-Perimetron,* 5(4), 194–211.
- ✔ Ortiz, S. (2007). Getting on board the enterprise service bus. *Computer*, 40(4), 15-17.
- ✔ Schmidt, M. T., Hutchison, B., Lambros, P., & Phippen, R. (2005). The enterprise service bus: making service-oriented architecture real. I*BM Systems Journal*, 44(4), 781-797.

- Peck, R., & Devore, J. (2011). *Statistics: The exploration & analysis of data*. Cengage Learning.
- Signori M. (2002). I lavori del Catasto Lombardo-Veneto nei territori lombardi. In Ricci G. and D’Amia G. (eds). La cultura architettonica nell’età della Restaurazione. Milano, Mimesis Edizioni: 85–100.

Development of the open cadastre of protected areas in Ukraine

Daria Svidzinska[1], Oleksij Vasyliuk[2, 3], Oleg Seliverstov[4], Daria Shyriaieva[1, 3], Anton Biatov[5, 3], Dmytro Diadin[7], Alevtyna Ponomarova[4], Oleh Sklyar[6], Svitlana Vinokurova[2], Ievgeniia Luchnykova[8, 9], Alexander Kleshnin[8]

[1] Taras Shevchenko National University of Kyiv

[2] Institute of Zoology of National Academy of Sciences of Ukraine

[3] Ukrainian Nature Conservation Group

[4] V. N. Karazin Kharkiv National University

[5] National Nature Park "Slobojanskiy"

[6] National Nature Park "Hetmanskyi"

[7] O. M. Beketov National University of Urban Economy in Kharkiv

[8] Intetics-Geo Company

[9] National Scientific Center "ISSAR"

Abstract

Ukraine does not have a reliable and publicly available source of information and data about protected areas (PAs). Therefore crowdsource development of the open cadastre of PAs was jointly initiated by non-governmental organizations (NGO) activists, GIS-specialists, conservation researches and practitioners. As a result, primary data has been collected for all regions of Ukraine, and for 10 of them the process of borders digitizing and data import into the OpenStreetMap (OSM) has been started. Alpha-version of the online service 'Daily Ukraine OSM Extracts' for the rapid data access and validation has been launched and available from http://opengeo.intetics.com.ua/osm/pa/.

Keywords

protected areas, open geodata, OpenStreetMap, crowdsourcing

Introduction

PAs are widely recognized as a cornerstone of biodiversity conservation and environment sustainability (European Environment Agency [EEA], 2010). Aiming to catch up with the European guidelines and best practices (EEA, 2012; NatureSDIplus, 2015) Ukraine has set up an ambitious goal to raise the PA from 6,08% to 15% from the country area by 2020 (Law of Ukraine on the basic principles of governmental environmental policy of Ukraine till 2020, 2010). Despite the goals and requirements of national and international environmental policy and legislation Ukraine still does not have a single common, up-to-date, reliable and publicly available source of information and data about PAs.
Previous attempts to develop national database of PAs have been scattered

and incomplete in their efforts to provide exhaustive and publicly available results. Also most of them were abandoned with time and currently contain outdated and incomplete information. Failure of these attempts can be attributed to several factors.

Firstly, Ukraine has a complex hierarchy of PAs types represented by 11 categories that on the highest level can be divided into PAs of state and local importance. Borders and location of the PAs of state importance (nature and biosphere reserves, national nature parks etc.) are well known, and fixed in the field, because they have their own administrations, encompass wide areas, and popular tourist destinations. At the same time, their share in the total amount of PAs is only 8.88%. Greater proportion is represented by PAs of local importance (nature monuments, protected sites, conservation stows etc.) which constitute about 91.12%. These categories are usually small areas, that do not have their own administration, and protection service. Often, their location is badly documented and known only approximately with the borders that are not strictly fixed.

Secondly, information about PAs is typically stored in inaccessible paper reports scattered amongst different local and regional state administrations. Moreover, there have been already known accidents of documentation lost during social unrest due to state of emergency, negligence, or criminal activity.

And last, but not least is that development of the national cadastre of PAs is the kind of project that should be financed and supported from the state sources. This kind of support is not currently available in Ukraine which struggles with economical and geopolitical crisis, and cuts all possible expenses.

Scarcity and inaccessibility of information and data significantly restricts PAs research, monitoring, and management. It also provokes violations of territorial integrity and land use regimes that put under threat protected species and habitats. As a result the effectiveness and integrity of regional and local nature conservation policy is seriously degrading. In case of state inaction civil society should become a driving force in the implementation of environmental initiatives.

Understanding the urgency and importance of the issue, the project aimed at the development of the open cadastre of PAs of Ukraine has been initiated in July 2014. This is a horizontal crowdsourcing project implemented by two groups of experts in the fields of nature conservation and geodata respectively. Members of the project are represented by NGO, academics and researchers, conservation practitioners, enterprises, volunteers. Currently the project is leaded by the representatives of the Ukrainian Nature Conservation Group, V. N.Karazin Kharkiv National University, and Taras Shevchenko National University of Kyiv.

The cadastre is expected to be characterized by the features as follows:

- actual – corresponding to the current state of the PAs in Ukraine;
- reliable – using credible and trustworthy sources of original information;
- exhaustive – providing comprehensive and complete information;
- publicly available – freely open to all relevant stakeholders;
- updatable – open to changes and modifications to maintain up-to-date state and completeness.

Implementation practices and results

The open cadastre of PAs aimed at improvement quality of PAs research, management, and monitoring through the achievement the following objectives:

- to establish free and open public access to the actual and updatable data and information about PAs;
- to support an innovative participatory approach gathering conservationists, researchers, technical staff, policy, and decision makers in creation, dissemination, and usage of publicly available data and information;
- to build knowledge and capacity in the use of geospatial data and software, by improving the professional level of technical and managerial personnel involved in PAs management and research;
- to initiate a process aimed at filling the gap between Ukrainian and international best practices and models for PAs management and monitoring.

As the PAs research, monitoring, and management affects the interests of the numerous parties project is a cooperative effort among the various stakeholders represented by NGO activists, conservation researchers and practitioners, technical GIS-specialists, and members of the OSM community. Their primary effort is aimed at the original data collection and development of the initial spatial data infrastructure based on the OSM framework.

The project is implemented through the following stages (Figure 1):

1. Primary data acquisition: original hardcopy documentation is scattered amongst various administrations (local, regional, national). The goal of this stage is to get legal access to the documentation and create its digital scanned copies.
2. Primary data sorting and completeness analysis: scanned copies are sorted and organized. In case of documentation lost or incompleteness request letters are prepared to inquire missing information.
3. Borders digitizing and validation: scanned maps of PAs are georeferenced, borders digitized and with some basic tags are uploaded into the OSM. Uploaded data is validated for the basic inconsistencies as follows:
   - naming errors (incorrect or invalid characters, extra spaces, unnecessary words);
   - multipolygon geometry errors (overlapping polygons, missing multipolygon relation);
   - tagging errors (empty or invalid tags).

   Error reports are generated automatically, and summarized in maps and tables. Open-source software is used for work: QGIS for scanned maps georeferencing, JOSM for data creation and editing, PostGIS and Leaflet at the stage of quality control, and GeoServer for publishing.

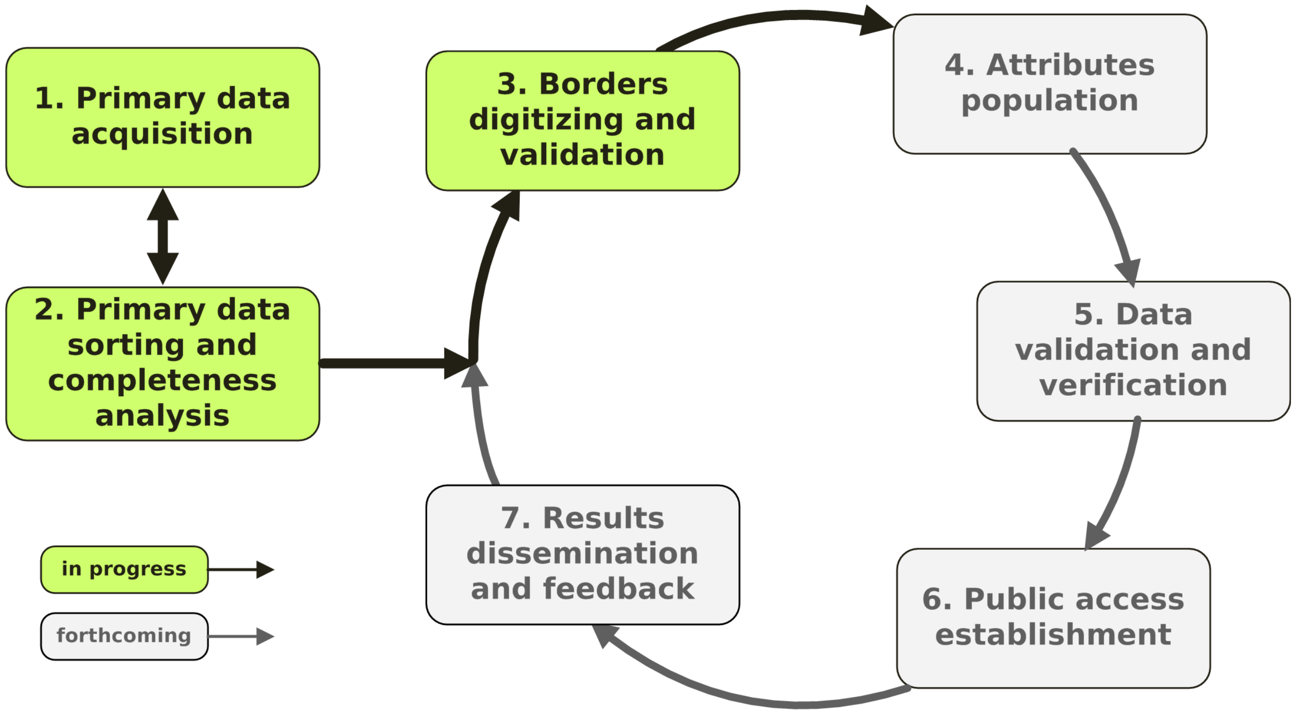


Figure 1: Main workflow stages.

4. Attributes population: available supplementary information from paper documents namely cards of primary account that are supposed to be a reference frame for the national PAs cadastre will be collected and aggregated in the uniform digital attributes.
5. Data validation and verification: digitized borders will be compared against public cadastral map of Ukraine to identify potential inconsistencies, land use and territorial violations. Also field surveys will be performed to update and correct the data.
6. Public access establishment: online public access repository is supposed to be a web-site that provides a web-interface to the initial information sources, developed cadastre database, and machine-readable data download (ESRI shape-files, GeoJSON, KML). Derivative maps are expected to become available online and offline for field survey navigation and mobile devices.
7. Results dissemination and feedback: information about the cadastre will be distributed through social and public media, workshops for relevant stakeholders. Publications that describe project methodology, results, and best practices will be prepared and distributed online and offline. Accumulated feedback is expected to define directions for further improvement.

The OSM was decided to be used as a common framework for the primary data creation and consolidation efforts due to the reasons as follows:

- complexity and flexibility of the OSM data model that allows to create, store, and manipulate various objects in nodes and ways, establish their spatial affiliation through relations, and provide well-structured description through the key=value tagging system;
- free and open access to the high-resolution imagery provided by

Microsoft Bing and Mapbox and available for tracing in the OSM as a basis for the digitizing of the PAs borders;

- access to the public GPS-data that is of great use for the positional accuracy control and enhancement;
- according to the Ukrainian legislation the data about PAs should be public, and once added to the OSM it inherits non-restricted access under the Open Data Commons Open Database License (ODbL).

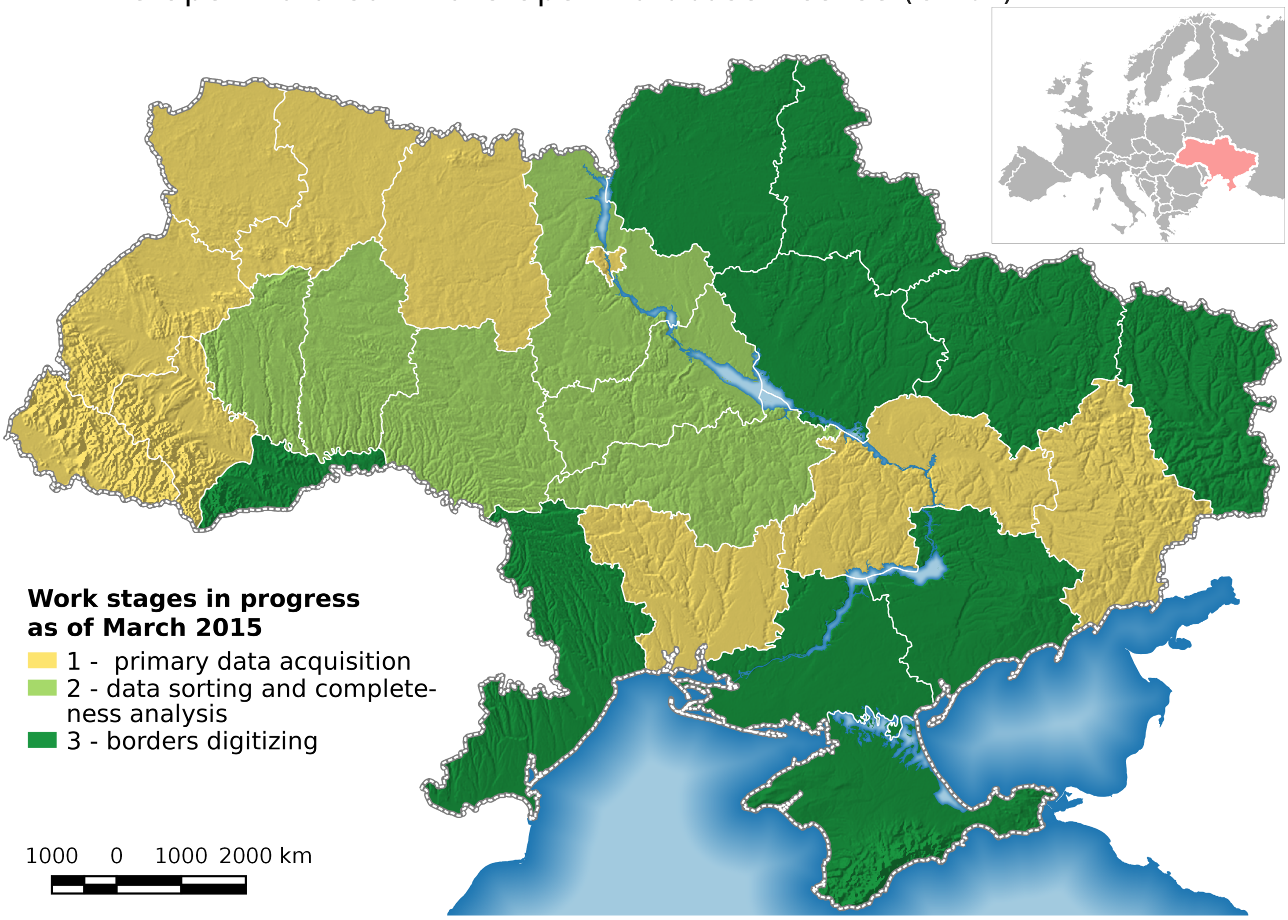


Figure 2: Work progress as of March 2015.

It is expected that work stages 1-2 to be performed once while stages 3-7 are connected into the cycle which means that sequence of steps can be performed multiple times to obtain the best result and keep the cadastre up-to-date and reliable. Currently we are at the beginning of the project and stages 1-3 are performed simultaneously (Figure 2). Mailing list, OSM Forum and Wiki, Mapcraft, and Sharepoint are used for community coordination and communication.
Also there is an alpha-version of the online service 'Daily Ukraine OSM Extracts' has been launched for the project purposes and available from http://opengeo.intetics.com.ua/osm/pa/ (Figure 3) that is currently being used to:

- monitor the overall progress;
- provide quick access to the data created (downloads available as ESRI shapefiles, KML, ESRI file GDB);
- rapidly assess data quality and edit errors;

- prototype and test services for online data access and distribution.

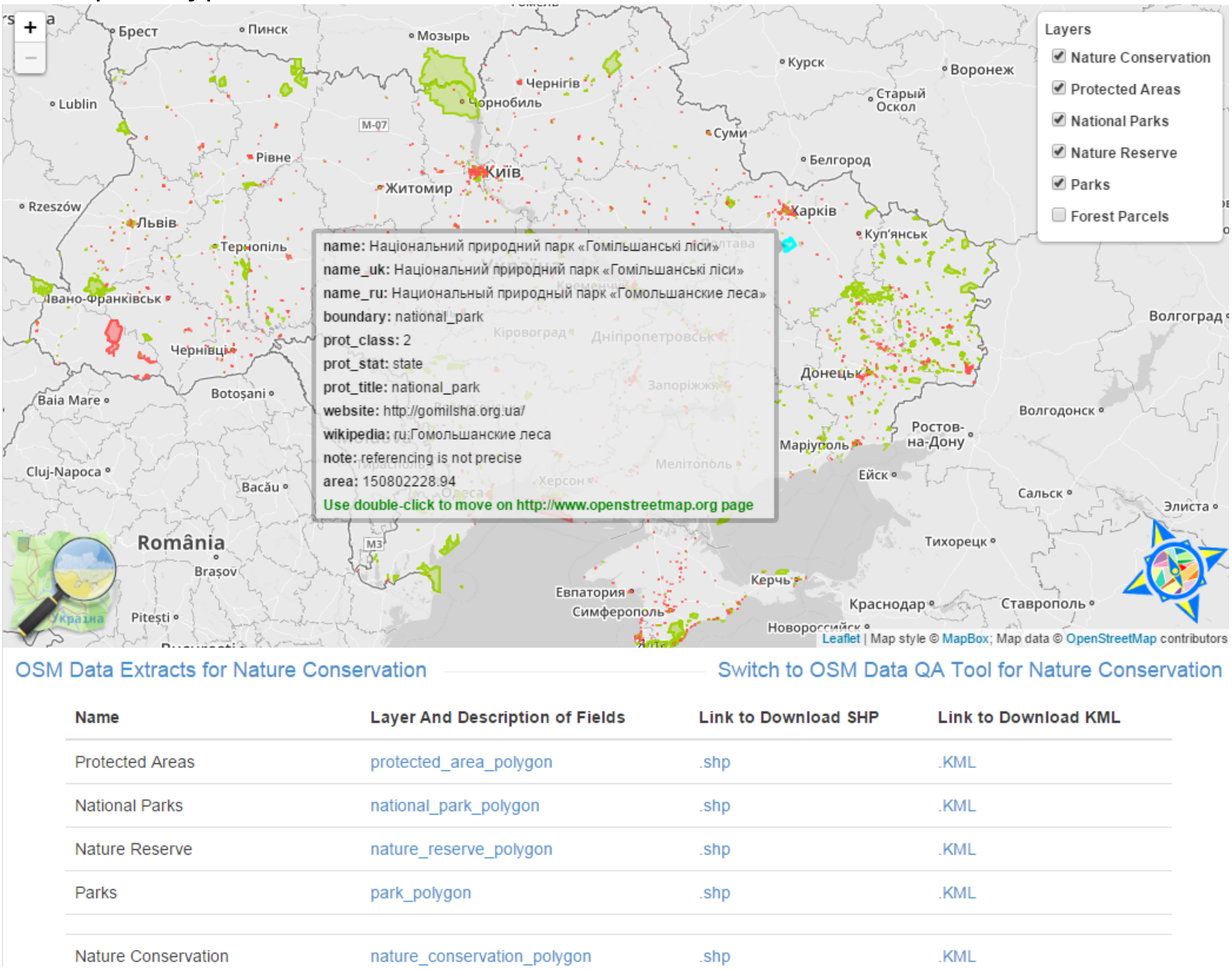


Figure 3: Web-interface of the online service 'Daily Ukraine OSM Extracts' available from http://opengeo.intetics.com.ua/osm/pa/

3 Conclusions and perspectives

As much of the work done by volunteers in their spare time, flexible tasks scheduling based on changes in capacity is used. It is planned to move to a more uniform planning as one of the organizations involved into the project will be able to allocate funding and resources on a regular basis. Nevertheless, application of the crowdsource data collection and processing has proven to be of great use for the joint development of the best practices and enhancement of the flexibility and scalability of their implementation. According to the rough estimates by the end of 2015 the following tasks are planned to be completed:

- development of the datasets for 5 regions and 1 city;
- initiation of the procedures for keeping the data up-to-date;
- publishing of the project's official web-site supplemented by the presence in the social media.

Provisional key deliverables include items as follows:

- consistent and updatable spatial database of PAs in Ukraine;
- a website that contains project news, documentation, maps, data quality

assessment tools (validators), spatial data, and provides non-restricted public access to its browse and download;
- all the data, documentation, and codes developed within the project will be available to the community under open licenses (e.g. ODbL, CC BY-NC-SA, GNU GPL respectively);
- developed and well-documented project approach and methodology available for further adoption;
- workshops aimed at policy and decision makers to cover project’s results and collect feedback;
- workshops aimed at researchers, PAs managers, NGO activists etc. to cover project’s results and collect feedback;
- community grown around the project (both OSM, and non-OSM) through social media and offline events.

As the project is based on open standards and technologies, its expertise and processes could be easily adapted to new regions and scaled to a higher level. In the future the cadastre will be able to be merged into the World Database of Protected Areas supported by IUCN and UNEP (Protected Planet, 2015). On the national level the cadastre database is expected to be harmonized with the Ukrainian SDI requirements that is currently under development and mostly relies on the INSPIRE guidelines and practices.

References

✔ European Environment Agency. (2010). *10 messages for 2010 Protected areas*. Copenhagen: Author. doi:10.2800/56932

✔ European Environment Agency. (2012). *Protected areas in Europe — an overview (No 5/2012)*. Copenhagen: Author. doi:10.2800/55955

✔ Law of Ukraine on the basic principles (strategy) of governmental environmental policy of Ukraine till 2020. (2010). Retrieved from http://zakon1.rada.gov.ua/laws/show/2818-17

✔ NatureSDIplus — Best practice network for SDI in nature conservation. (2015). *NatureSDIplus Good Practices*. Retrieved February 25, 2015, from http://www.nature-sdi.eu/index.php?option=com_fabrik&view=table&tableid=3&calculations=0&resetfilters=0&fabriklayout=default_grouped&Itemid=73

✔ Protected Planet. (2015). *The latest initiative harnessing the World Database on Protected Areas*. Retrieved February 25, 2015, from http://www.protectedplanet.net/about

How is Prague opening geospatial data

Jachym Cepicky[1], Martin Landa[2]

[1] Geosense, s.r.o

[2] CTU in Prague - Geomatics

Abstract

Prague Institute of Planning and Development started to open their (not only) geospatial data. But before this was technically implemented, potential users were asked (with help of questionnaire), which type of data they want, how do they want to distribute it, which formats do users prefer and other questions to help with the clarification of optimal way for geospatial data opening.

Based on internal discussion and external consultation, choosed solution is based on INSPIRE download service using ATOM format, with pre-generated raster and vector geospatial data for various formats and coordinate reference systems. Custom script, written in Python programming language was build, in order to be able to publish the datasets from internally used database to web servers.

In the presentation, the results of the questionnaire will be presented as well as final solution.

A FOSS4G-based procedure to compare OpenStreetMap and authoritative road network datasets

Maria Antonia Brovelli[1], Marco Minghini[1], Monia Molinari[1], Peter Mooney[2]

[1] Politecnico di Milano, Como Campus, DICA, Via Valleggio 11, 22100 Como, Italy

[2] Maynooth University, Department of Computer Science, Maynooth, Co. Kildare, Ireland

Abstract

OpenStreetMap (OSM), the most popular VGI crowdsourcing project, is an excellent example of an open-license spatial database. But what is the quality of OSM road network datasets compared to authoritative counterparts? Several comparisons of this kind are detailed in literature but these cannot be easily adapted to other scenarios. Developing a generic automated procedure is very challenging. This paper proposes a FOSS4G-based procedure for automated quality comparison of OSM and any authoritative road network datasets. We detail work-in-progress which has great potential. Our procedure is currently implemented into a GRASS command with future plans to extend this to a QGIS plugin and a FOSS4G-based WPS.

Keywords

Road network, FOSS4G, GRASS, open geo-data, OpenStreetMap, quality.

1 Introduction

Comparing open datasets is an interesting computational and societal challenge. Open geographic datasets from Volunteered Geographic Information (VGI) projects like OpenStreetMap (OSM) are excellent examples of crowdsourced geographic data of real world features such as buildings and roads. Recently some National Mapping Agencies (NMAs) have been making their geographic datasets available as open data. This greatly increases their distribution and opportunities for access for citizens. Comparing OSM data and open data from NMAs has been studied in the literature. There has been some work published reporting comparison of OSM and other official datasets displaying different levels of automation. Mooney, Corcoran, and Winstanley (2009) developed an automated quality assessment measure to compare OSM with authoritative datasets containing natural water features. Fairbairn and Al-Bakri (2013) compared OSM and authoritative large-scale databases in the UK and Iraq to address possible integration of these datasets. Ludwig, Voss, and Krause-Traudes (2011) developed a fully automated approach to matching street objects in Germany contained in OSM and in the commercial Navteq database. Haklay (2010) compared the OSM dataset and Ordnance Survey dataset in the UK, while Girres and Touya (2010) compared the quality of the

OSM dataset in France with the reference database from the French National Mapping Agency. Siebritz (2014) performed a quantitative and qualitative comparison between OSM and national mapping agency data in South Africa. In the studies outlined above (and some others in the literature) the authors have designed and developed software implementations of comparison methodologies which are specific and tightly coupled to the OSM and authoritative datasets involved. We believe that this makes the comparisons more complex and can prevent other authors from replicating them.

2 Our proposed comparison procedure

In this paper we propose a novel and generic procedure to perform comparisons of OSM and authoritative road network datasets in terms of spatial accuracy and completeness. The procedure is designed to a comparison framework which is general and applicable (in principle) to any pair of comparable open datasets while carefully considering their specific characteristics. When executing this procedure users can supervise the computation by manipulating parameter values to reflect the specific features of the authoritative dataset under investigation such as its scale and nominal accuracy. These characteristics make the procedure effectively adaptable to most comparisons of this kind.

2.1 Procedure implementation and workflow

The procedure is developed as a GRASS module written in Python. Its main steps are as follows: preliminary comparison of the datasets and computation of global statistics; geometric preprocessing of the OSM dataset to extract its networks representing the same road features as the authoritative dataset; and evaluation of the OSM dataset's spatial accuracy through a grid-based approach. These steps are separately described below by outlining their purpose, input/output data and GRASS modules used. REF (meaning “reference dataset”) is used to identify the authoritative road network dataset being compared with OSM.

2.1.1 Preliminary comparison of the datasets

The first step prepares the OSM and REF road datasets in addition to performing an initial comparison of their spatial coverage similarity. The key operations are as follows:

- import the OSM and REF datasets; if spatial clipping is required the user can import a vector layer to be used as the clipping mask (*v.in.ogr*, *v.overlay*);
- compute the total length of the OSM and REF datasets and their length difference, both in meters and percentages (*v.to.db*);
- apply a user-specified buffer around the REF and OSM datasets and compute the length and the length percentage of the OSM and REF datasets included in the buffer (*v.buffer*, *v.overlay*, *v.to.db*).

2.1.2 Geometric preprocessing of OSM dataset

This step prepares the OSM dataset so that only features which have a correspondence in the REF dataset are extracted. This correspondence is computed using angular coefficients as follows:

- apply a user-specified buffer around the REF dataset to extract only the OSM features included (*v.db.select*, *v.extract*, *v.buffer*, *v.overlay*);
- compute the angular coefficient of each feature in the REF dataset (*v.db.addcolumn*, *v.to.db*, *v.db.update*). This is compared to the angular coefficient of all OSM features falling inside the buffer around that REF feature. If the difference between the two angular coefficients exceeds a user-specified threshold (e.g. 30°), then that OSM feature is deleted as it does not correspond to the REF feature considered (see Figure 1). Corresponding OSM features are instead added into a new vector layer (*v.edit*, *v.patch*).

Before these operations the REF line features must be split into segments (*v.split*) allowing the angular coefficient be computed on each segment. This considerably increases the number of REF features for comparison and consequently the computational time required. For this reason the splitting operation is preceded by a generalization of the REF dataset (*v.generalize*) using the Douglas-Peucker algorithm (Douglas and Peucker, 1973). Users specify the threshold value for the line feature generalization (see Figure 1).

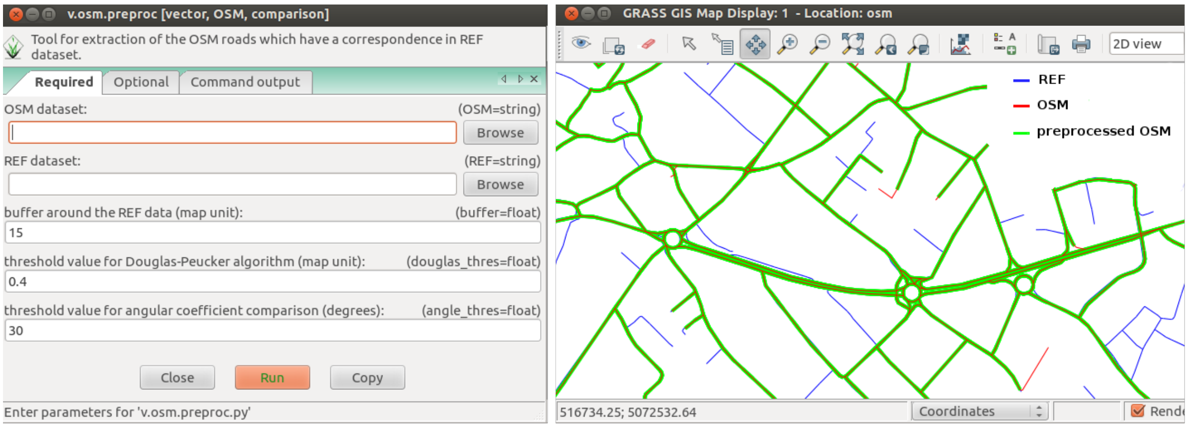


Figure 1: Geometric preprocessing of OSM dataset showing GRASS interface for user inputs (left) and REF, original OSM and preprocessed OSM roads (right).

2.1.3 Evaluation of OSM spatial accuracy using a grid-based approach

This step performs the comparison between the REF and the preprocessed OSM datasets and evaluates the spatial accuracy of the latter. This process is based on a grid approach which takes into account the possible heterogeneous nature (translating into an heterogeneous accuracy) of the OSM dataset. The single operations are the following:

- define a grid, asking the user to either build it in real time (by setting the bounding box and the grid step in the two directions (*v.mkgrid*)) or uploading a predefined polygon vector layer (*v.in.ogr*). The use of a grid is optional.
- apply one or more user-customized buffers around the REF dataset and compute the length and the length percentage of the OSM dataset included in the buffer (*v.buffer*, *v.overlay*, *v.to.db*).

The final operation is executed to obtain a number of different outputs according to user requirements. For example users can retrieve the grid cells where the deviation of OSM dataset from the REF dataset is lower than a fixed threshold or within a fixed interval of thresholds. Moreover they can evaluate which is the maximum deviation from the REF dataset for each grid cell.

3 Conclusions

The proposed procedure is under active development and research. We are working to achieve improved computational performances. In addition to the geometrical operations described here an analysis of the correctness of OSM road attributes (i.e. the values of the *highway* key) compared to the REF dataset is planned for future development. The procedure has been tested to compare the OSM road network dataset with those authoritative open datasets of the municipalities in the Lombardy Region in Northern Italy (scale 1:2000). Our immediate future work should, at first, confirm the suitability of the procedure – which, as shown, is heavily supervised by the user input values – on the input datasets. From a FOSS4G viewpoint the authors will develop this procedure to become a QGIS plugin before being provided as a Web Processing Service (WPS) available for the whole community.

Acknowledgements

The authors grateful acknowledge support from COST Action TD1202 “Mapping and the Citizen Sensor” and specifically the Short Term Scientific Mission (STSM) COST-STSM-TD1202-24885.

References

- ✔ Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *10*(2), 112-122.
- ✔ Fairbairn, D., & Al-Bakri, M. (2013). Using Geometric Properties to Evaluate Possible Integration of Authoritative and Volunteered Geographic Information. *ISPRS International Journal of Geo-Information*, 2(2), 349–370.
- ✔ Girres, J.-F., & Touya, G. (2010). Quality Assessment of the French OpenStreetMap Dataset. *Transactions in GIS, 14*(4), 435–459.
- ✔ Haklay, M. (2010) How good is OpenStreetMap information? A comparative study of OpenStreetMap and Ordnance Survey datasets for London and the rest of England. *Environment and Planning B*, 37, 682-703.
- ✔ Ludwig, I., Voss, A., & Krause-Traudes, M. (2011). A Comparison of the Street Networks of Navteq and OSM in Germany. In S. Geertman, W. Reinhardt & F. Toppen (Eds.), *Advancing Geoinformation Science for a Changing World* (pp. 65-84). Springer Berlin Heidelberg.
- ✔ Mooney, P., Corcoran, P., & Winstanley, A. C. (2010). Towards Quality Metrics for OpenStreetMap. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (pp. 514–517). New York, NY, USA: ACM.
- ✔ Siebritz, L.-A. (2014). Assessing the accuracy of openstreetmap data in south africa for the purpose of integrating it with authoritative data. PhD Thesis, University of Cape Town, SA.

Fentrol.hu – The digital aerial photo archive

Péter Braunmüller[1], Krisztián Takács[1], Csaba Sándor[2]

[1]FÖMI (Institute of Geodesy, Cartography and Remote Sensing - Hungary)

[2]Viamap Ltd.

Abstract

FÖMI has successfully completed the Digital Aerial Photograph Archive project funded by EU, on 30th June 2014. The project had been executed by the support of EKOP 2.A.2.

In accordance with the Hungarian law the goal was to digitize more than 60,000 analog photos in the archive, organize them into a database and present them online to all kinds of users. The project also aimed at crowd-sourcing the georeferencing, tagging of the digitized images as well as inspiring the community to discuss them.

Keywords

Aerial photo, Archive, Digitization, Geospatial, Open data, Crowd-sourcing

1 Introduction

The consistency of the analog aerial films deteriorate continuously. For this reason, Digital Aerial Photograph Archive project, the worst conditioned aerial photographs were digitized mostly from the period 1959-1966. There are almost a half million frames in the archives from the period of 1959 - 2005. By the end of the project we have finished scanning 61,800 archive images.

2 Processing of the Aerial Photo Archive

2.1 Georeferencing

First, metadata were recorded from the logbooks of the flights; then the photographic frame centres were digitized. In case flight plans (containing the frame centres of the images on long papers) were available, the tracing papers were strengthened on a flat surface and a digital map was projected on them. The necessary transformations were made according to the relevant points and the frame centres were digitized (Figure 1).

The films dated before 1969 don’t have tracings, so their frame points were determined by visual identification from the single aerial photographs.

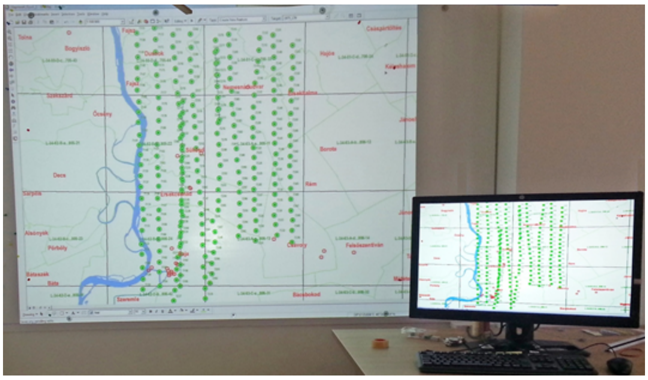


Figure 1: Digitizing the frame centres.

2.2 Scanning

Many photographs had adhesive tapes on the surface. The encrypted images were often covered with glue. To clean and scan the films, preparatory pressurized rooms were established. The scanning was performed with two Leica DSW 700 aerial film scanners.

2.3 Preprocessing for the web framework

To initialize the website mass-calculation of initial georefencing parameters was required from the position and flight height. The images were also preprocessed using semi-automated workflows (resize, fill metadata in the database etc.). The website features a georeferencing module in OpenLayers and Single Sign On facility with Facebook, Google, geoshop.hu and the Hungarian official administrative portal magyarorszag.hu.

3. Fentrol.hu – Open archive data

3.1 Architecture

The web application utilizes a FOSS4G technology stack in Linux operating environment on 1 physical server: linux-apache-mysql-php (LAMP) web framework, Postgresql+PostGIS database engine with foreign data wrapper interface to mySQL, Mapserver in FastCGI mode, OpenLayers map client with custom interface components.

3.2 Viewing and download

The digitized aerial images can be freely browsed on the Internet at www.fentrol.hu (Figure 2).

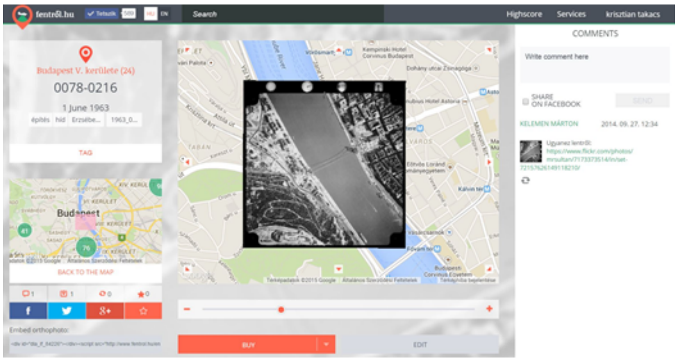


Figure 2: Preview of an aerial photograph at fentrol.hu.

The photographs can be bought online (with georeference files also), or could be downloaded free in reduced resolution (80%). Fentrol.hu allows searching aerial photographs by location or by attributes also.

3.3 Interactions

Every registered user can tag, comment or annotate a photo. The clients have possibility to georeference the aerial photographs (Figure 3). As an advantage of the power of the crowd (crowd-sourcing) until now nearly 10 % of the uploaded images have been edited to have the correct georeference information.

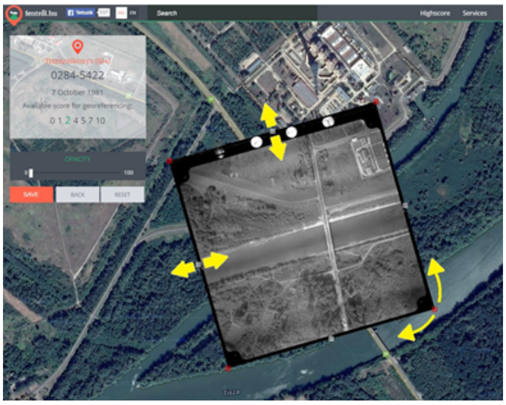


Figure 3: Georeferencing of an aerial photograph.

3.4 Typical user activities

Clients use fentrol.hu to research local history. They are monitoring the changes of railways and sports facilities, airports and military installations. Archaeologists research ruins, fortresses, lines of trenches.

4 Development of the database

Since the end of the EKOP 2.A.2. funded project, FÖMI have continued to digitize the archives. By the end of February 2015, further 9.885 archive images were scanned and uploaded to fentrol.hu website so currently there are 78.114 aerial photos available online.

5 Conclusions

This work is one of the first attempts made by a public administration organization in Hungary to improve public access to geospatial content in large quantities. The fentrol.hu project besides contributing to preservation of national heritage, also fostered OpenData and Crowd Sourcing concepts by providing data to interested parties and also involving them in increasing its quality and information content. The technical lessons learned during the realization of the web application showed that such amounts of image data can be economically and efficiently presented using opensource geospatial software stack.

Improving environmental monitoring against the risk from uncontrolled abandonment of waste containing asbestos. The DroMEP project.

Carmine Massarelli[1], Maria Rita Muolo[2], Vito Felice Uricchio[1], Nicola Dongiovanni[2], Ruggero Palumbo[2]

[1] Water Research Institute, National Research Council. Via De Blasio n. 5 - 70132, Bari

[2] Servizi di Informazione Territoriale S.r.l., P.zza Papa Giovanni Paolo II, 8/1, 70015 Noci (BA)

Abstract

DroMEP is the acronimous of the project *Drones for Monitoring and Environmental Protection* that aims to significantly contribute to the realization of a sustainable smart monitoring system regarding to the management of environmental issues related to the presence of illegal dumping and uncontrolled abandoned or burial waste, especially if containing asbestos. The aim will be pursued through the creation and implementation of an integrated Geo-Database with heterogeneous data from different sources, including those acquired through UAV and a smart app, that would provide tools for extraction, analysis and post-processing of the data for research purposes, environmental characterization and monitoring. The WebGIS system to support the coordination of the smart monitoring method will allow to view and consult of all the data and information and it will act as DSS to suggest the adoption of operational strategies and tactical decisions in critical situations. Tools to facilitate the consultation of data from different sources relating to entities with their own autonomy and role and everyone involved in the process of protecting the environment and health of citizens, will be implemented to achieve the interchange of data between institutions. This will allow a more efficient and integrated information management between all actors involved in optimizing processes of evaluation and planning.

Keywords

Sharing, interoperability, monitoring, asbestos, drones

1 Introduction

DroMEP was created with the intention of making more widespread the monitoring process of the territory by relying on the collaboration of all citizens, an operation definable *bottom up*.
In addition to the ICT platform made it aims to simplify the management of the *big data* that the Public Administrations have to manage every day.
The government portal on open data has changed ownership in march, 2015;

so this period is crucial to ensure that the resulting takeover of open data make sure that they are truly accessible, understandable and usable. The open data enable democracy, making citizens more informed and aware, and the new entrepreneurship, releasing data to developers and civic hackers to create new apps and services.
Just think of the importance of environmental information by which it would be possible, for example, manage more efficiently its territory: from control to monitoring of protected areas to interventions for sustainable agriculture.
It is no coincidence that one of the dataset from the website *dati.gov.it* surveyed, the majority of those made accessible to citizens and businesses are linked to the trend of the population and environmental phenomena.
In this view of data sharing and participation for the monitoring and control of environmental risks due to the presence of asbestos abandoned in agro-ecosystem project DroMEP was born.

2 Material and methods

2.1 Study area

The study area is the *Gravina di Leucaspide* (Statte Municipality), area of difficult access to the usual tasks of monitoring (in fact, this area is expected to use a drone), the natural oasis of *Salina Grande* of Taranto and the farmland of Capurso Municipality.

2.2 Open source tools

The basis of the research work was placed a methodological multidisciplinary marked by the creation of a spatial data infrastructure, publishing and sharing services with attention to the statement of environmental elements, it can also be found within the context of protected areas which may be threatened by illegal littering. The aim is to support (thanks to applications developed) operations management, study and communication between citizens and Authority to face critical issues. The achievement of results is pointed to an ICT architecture that starting from the relational database, through the geodatabase and reaching dynamic and interactive webgis platforms act as support to the processes of coordination, analysis and dissemination of selected elements related to the territorial area and its main planning tool facilitating and supporting management and decision processes so be more aware.
Here there are some of the components: Apache Tomcat server (Vukotic & Goodwill, 2011); Geoserver, independent platform server side software widely used in applications for management of data from forest areas (Yu & al., 2013); GeoNode for sophisticated web browser spatial visualization and analysis; GeoExt JavaScript Toolkit for Rich Web Mapping Applications; GeoExplorer a browser-based interface; OpenLayers a modern library that adopted the improved JavaScript-rich browser environment (Schmidt, 2008); PostgreSQL open source object-relational database system.

2.3 Open data

The following list shows the steps of the proposed approach to integrate and use open data into the project: 1. identify and select the dataset; 2.

remediation and cleaning of corrupted or unusable data; 3. dataset enrichment by metadata; 4. analysis and modeling; 5. linking with external datasets; 6. validation and publication.

We are considering the possibility of cataloguing by using the method of the stars[1]. The platform (point 2.2) includes the essential building blocks to deploy an OGC-compliant spatial data infrastructure (SDI). The open source server GeoServer enables to publish spatial data from a wide variety of data sources using open protocols from the OGC such as Web Map Service (WMS) and Web Feature Service (WFS). In fact, the data managing platform uses WMS/WFS services[2] which implies two important and indisputable advantages for end users of geographic data:

1. uselessness of data duplication in local machines;
2. certainty to consult/use a certified version of the requested data because they are directly distributed from those who produce them or certifies their reliability.

These services, therefore, allow the creation of a distributed map servers network from which individual clients can create customized maps and then share them in several formats, such as csv, gml, kml, xml, jpg, pdf, etc.

In addition, the platform is capable of metadata and catalogue-system management and also enables publishing metadata in a variety of standards (ISO, FGDC, NASA DIF, Dublin Core).

For the purpose of the project metadata standard ISO 19139:2007 will be edited using a Catalogue Service of the Web (CSW) interface (users may also upload a metadata XML document (in ISO, FGDC, or Dublin Core format) to fill elements automatically. The catalogue implementation, pycsw, is certified OGC-compliant. In the platform all the available service are: Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Catalogue Service for Web (CSW), Web Map Context (WMC), Tile Map Service (TMS).

Harmonization of data will be possible by providing information from a geodatabase in OGC standard formats such as WMS architecture supported by open source software.

The levels of accuracy of geospatial data captured through the drone are very high: they have a resolution of 4-5 cm to the ground; geospatial data captured by airborne sensors, have a resolution of 20-25 cm;

Refresh rate is related to acquisitions by UAV and to any progress in the operations of securing and remediating of identified sites; refresh rate is immediate in the case of new reports acquired by citizens.

2.4 Smart app

DroMEP also includes the development and release of an App for last generation mobile devices, in order to directly involve citizens and the operators of organizations and associations responsible for environmental risks control and management, in the field collection of geospatial data and information, related to specific project issues.

[1] Open data: how to make public authorities data, opened – Guidelines for the Web sites of Public Administration Vademecum.Beta 2011- www.funzionepubblica.gov.it/media/982175/vademecumopendata.pdf

[2] http://webgis.sit.puglia.it/sit-help/SIT-Puglia/Guida/Sit-Cittadino/Standard-Ogc.html

2.5 The UAV

Within DroMEP project, spatial data acquisitions has been planned through the use of unmanned aircrafts (or Drones). The gain is to take advantage of low altitude surveys that are more accurate and detailed, and that become necessary to identify, characterize and monitor authorized or unauthorized landfills and uncontrolled release of waste into the environment, both in terms of quality and quantity.

2.6 The environmental-aware community

The environmental-aware community will be created through specific phases foreseen in the project. In short we can summarize them in: 1. communication, for better understanding the use of open data; 2. training, involving operators in the public sector and stakeholders; 3. dissemination, by collecting and publishing examples of good practices on a national catalog (eg. Best Practice Catalogue smeSpire 8) useful to the activities of soil characterization and reclamation; 4. making homogeneous and harmonized data open and accessible with a web browser (some information might be considered sensitive as those related to the owners of the sites where asbestos is found).

2.7 The collaborations

Plans for collaboration in the monitoring activities are already active, to the 2015 there are conventions of CNR at national level with the *Guardia di Finanza* for environmental crimes and at the regional level with the *Corpo Forestale dello Stato* (for the monitoring of forest areas) and the *Guardia Costiera* (for the monitoring of coastal and humid areas).

3 Partial results

The project is still in progress, the ICT platform has been realized all the *actors* begin to share information through it. The DroMEP data are already available through a web browser and the information are also useful in terms of environmental controls.
Aspects that will affect future activities are those related to the improvement of the data, by clumping them with search semantic technologies, that is by transforming and enriching them with language or logical reasoning analysis techniques.

References

- ✓ Schmidt, C. (2008). Open source #15: OpenLayers. *GEO: connexion*, 7 (10), pp. 36+38.
- ✓ Vukotic, A, & Goodwill, J. (2011). *Apache tomcat 7*. New York, NY. Apress
- ✓ Yu, X., Liu, H., Yang, Y., Zhang, X., Li, Y. (2013). GeoServer based forestry spatial data sharing and integration. *Applied Mechanics and Materials*, 295-298, pp. 2394-2398.

Towards Open Big Geospatial Data for geodata.gov.gr

Angelos Tzotsos, Michail Alexakis, Spiros Athanasiou, Yiannis Kouvaras

Institute for the Management of Information Systems / "Athena" Research and Innovation Center

Abstract

Open data provided by the public sector constitute a significant opportunity for growth. PublicaMundi (http://publicamundi.eu) is an EU FP7-ICT project aiming to make open geospatial data easier to discover, reuse, and share by fully supporting their complete publishing lifecycle in open data catalogues. To achieve this, PublicaMundi extends and integrates leading open source software for open data publishing and geospatial data management. During the first year of the project, the main focus of the development was on extending the CKAN project with new geospatial capabilities through integration with leading OSGeo software, such as pycsw, rasdaman, ZOO-Project, GeoServer and GDAL. CKAN was extended to support various established metadata models (including ISO-19115 and INSPIRE schemata), through integration with OWSLib and pycsw. Additionally, CKAN has been succesfully integrated with the rasdaman raster data engine, adding big data capabilities to the PublicaMundi software stack, providing the ability to serve hundreds of TBs of Earth Observation data through the WCS OGC API or through PublicaMundi's Data API. A first integrated prototype is already available on labs.geodata.gov.gr, providing beta access to data publishers and developers for Greek open geospatial data.

Keywords

PublicaMundi, CKAN, pycsw, Open Data Catalogues, OGC

1 Introduction

Open data provided by the public sector constitute a significant opportunity for growth. Geospatial data account for an estimated 80% of public sector information and are the most significant category of open data due to their high production, procurement, and update costs, as well as their relevance in multiple domains. Despite their importance, they are increasingly difficult to reuse, especially in a cross-boundary multilingual context. The vast majority of open data catalogues in the EU have limited support for geospatial information with insufficient capabilities in publishing methodologies and tools, limited technical foundations to support value added services, and simplistic non-scalable support for geospatial data visualization.

PublicaMundi (http://publicamundi.eu) is an EU FP7-ICT project aiming to make open geospatial data easier to discover, reuse, and share by fully supporting their complete publishing lifecycle in open data catalogues. To achieve this, PublicaMundi extends and integrates leading open source software for open data publishing and geospatial data management. In particular, PublicaMundi extends CKAN, the leading open data catalogue, into treating geospatial data as “first-class citizens” and providing automatic OCG and INSPIRE access to geospatial data.

A first integrated prototype is already available on labs.geodata.gov.gr, providing beta access to data publishers and developers for Greek open geospatial data. In the next few months, PublicaMundi is being deployed to geodata.gov.gr to serve as the main open geospatial data catalogue of the Greek government.

This paper is structured as follows: in Section 2, the system architecture is briefly described. The integration environment of the PublicaMundi project is presented in Section 3, while geospatial extensions of CKAN and open data publishing workflow are described in Section 4. In Section 5 the support for big geospatial data is presented. In Section 6, the processing engine is shown and Section 7 describes future work.

2 System Architecture

This chapter describes the System Architecture of PublicaMundi, which consists of several application levels.

PublicaMundi project was designed based on seven application levels, from the high level client applications to low level data storage components. The system follows a multi-tier, multi-level and multi-user approach aiming to provide tools for all stages of the open data lifecycle.

The goal of this multi-layer architecture was to address specific deployment needs (e.g. different map servers, deploy other CKAN extensions if needed, deploy on top of a database cluster, or change a specific component of the system without affecting the rest of the components). At the same time, another goal was to make the system architecture extensible for future improvements. The system was envisioned to be built on many loosely coupled layers in order to be able to scale and be deployed in a modular way, according to the end-user needs.

More specifically, the application levels are:

- **Data Storage.** This layer includes all the storage units, database clusters and cloud storage infrastructures that are used or developed in the PublicaMundi system. Currently, this layer of the system is based on PostgreSQL/PostGIS, rasdaman (Baumann, 2014) and cloud storage provided by Synnefo cloud stack (Koukis et al 2013).
- **Data Processing.** This layer includes all the spatial software that is responsible for transforming and processing spatial data (vector, raster and their metadata) before the core CKAN application publishes them on the web. Currently, this layer of the system is based on GDAL, PROJ, OWSLib and ZOO WPS.
- **CKAN (core).** The core application is based on the widely used open data catalogue CKAN. CKAN is based on the Pylons Web Framework, which follows the *Model, View, Controller* approach (MVC). This approach

has the advantage of loose coupling of resources and software components, which leads to great re-usability of the source code and better management of the stack. This layer of the system is based on CKAN application in stable version 2.2.

- **Application Modules.** This layer consists of CKAN extensions that were developed in order to spatially extend CKAN but also to provide new functionality such as multilinguality. Currently, this layer of the system is based on pycsw, ckanext-schematic, ckanext-publicamundi, ckanext-harvest and ckanext-spatial. The PublicaMundi extension also includes modules to enhance scalability of the application (caching, proxy and analytics).
- **Web Services.** This layer consists of several geospatial servers that are integrated with CKAN providing geospatial data support for the web. These servers are compatible with OGC and come from the OSGeo stack. Currently, this layer of the system is based on GeoServer, rasdaman, ZOO WPS and pycsw. Support for MapServer and MapProxy is being implemented.
- **Developer APIs**. This layer consists of all the development tools and APIs that are provided to developers in order to: create maps, re-use data, process data and publish/search/harvest metadata. On top of CKAN API and several OGC APIs, PublicaMundi provides its users with two new APIs: the Mapping API and the Data API. The Mapping API is based on integration of OpenLayers and Leaflet. The Data API is based on PostGIS and ZOO-WPS.
- **Client Applications**. This layer consists of external applications that use PublicaMundi’s APIs in order to interact with open data. Currently, this layer of the system is based on QGIS, CKAN harvesters and other CSW clients.

Figure 1 provides a detailed overview of the software stack used to implement the system.

3 Integration environment

For the purposes of the system development and the first deployment to labs.geodata.gov.gr, an integration environment was established. On this environment, beta testing of the software takes place and the development team gathers feedback from the early users/testers of the system. For this environment, a set of high-end servers was acquired and the software components were deployed. Also, new versions of the software components are periodically integrated from the project's git repository and are available for evaluation. Labs.geodata.gov.gr was initialized with real-world geospatial datasets from geodata.gov.gr (geospatial datasets plus *metadata*) to enable user to experiment with all implemented data publishing/reuse services.
The integration environment of PublicaMundi is deployed on top of the Synnefo cloud stack (Koukis et al 2013), within a number of virtual machines. Synnefo is a complete open source cloud stack written in Python that provides Compute, Network, Image, Volume and Storage services, similar to the ones offered by Amazon Web Services (AWS).

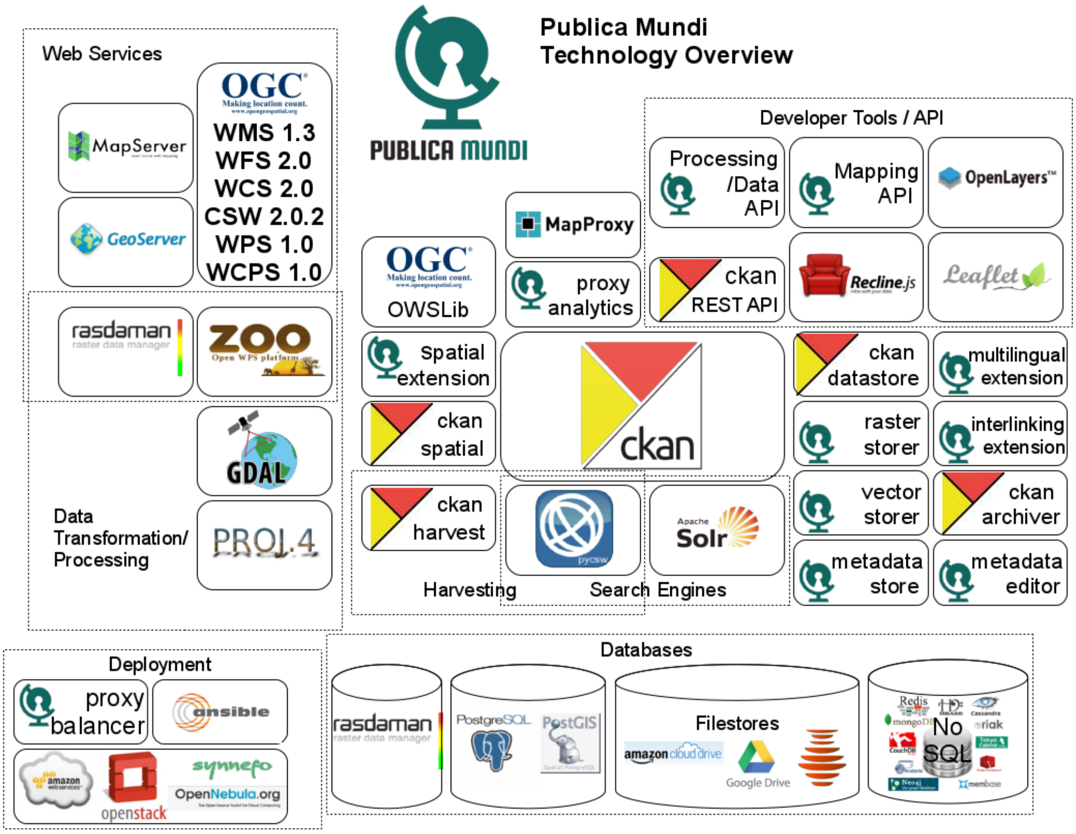


Figure 1: PublicaMundi software stack.

Synnefo manages multiple Ganeti clusters at the backend for handling of low-level VM operations and uses Archipelago to unify cloud storage. To boost 3rd-party compatibility, Synnefo exposes the OpenStack APIs to users. Synnefo keeps a clear separation between the traditional cluster management layer and the cloud layer. This unique design approach leads to a completely layered architecture. The layered approach for both the Compute and the Storage service boosts production readiness, maintainability and upgradability (Koukis et al 2013). The modular design allows for linear scalability, gradual extensibility and ease of operations. In Figure 2, an overview of the Synnefo architecture is presented, showing the major components on each layer.

For the deployment of PublicaMundi, the Synnefo stack was installed on the available server infrastructure, and a Web User Interface (UI) was made available to the administrators in order to maintain and manage the cloud resources for the integration environment. In Figure 3, the administration page of the Synnefo UI is shown. Through that interface, the user can have an overview of the virtual machines available, can restart, shut down, start up, delete, and create a virtual machine. For the deployment of labs.geodata.gov.gr several virtual machines were created, using the Debian 7 GNU/Linux operating system. This strategic decision was based on the maturity and stability of this operating system, and the fact that it is used in production servers on geodata.gov.gr since 2010.

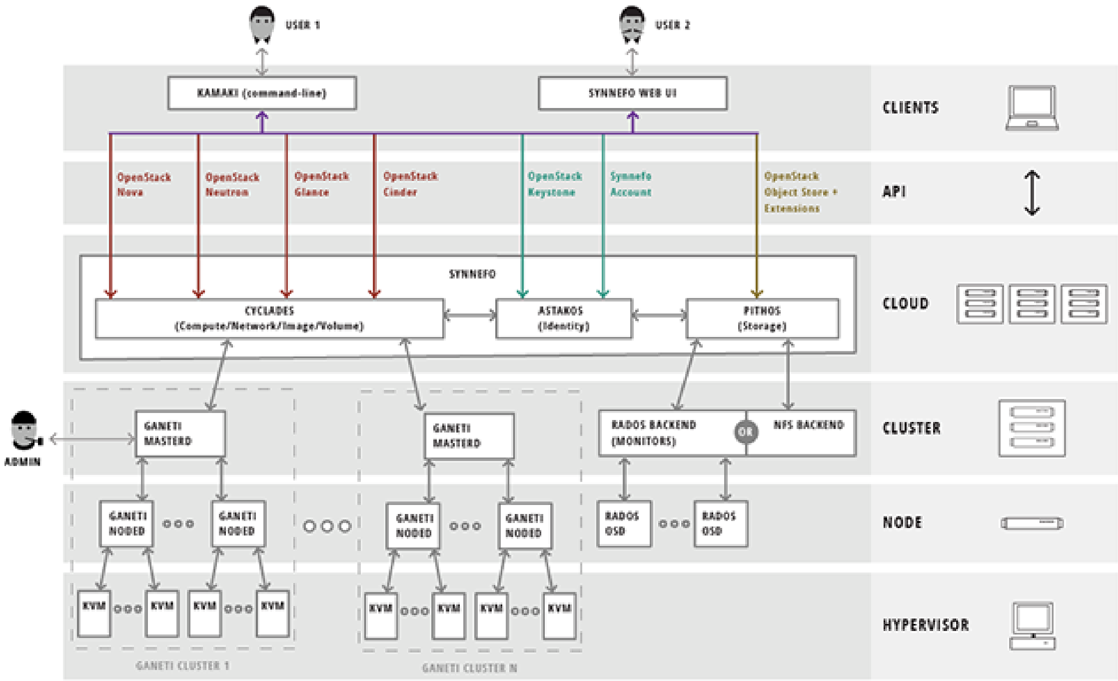


Figure 2: A detailed description of the Synnefo Architecture.

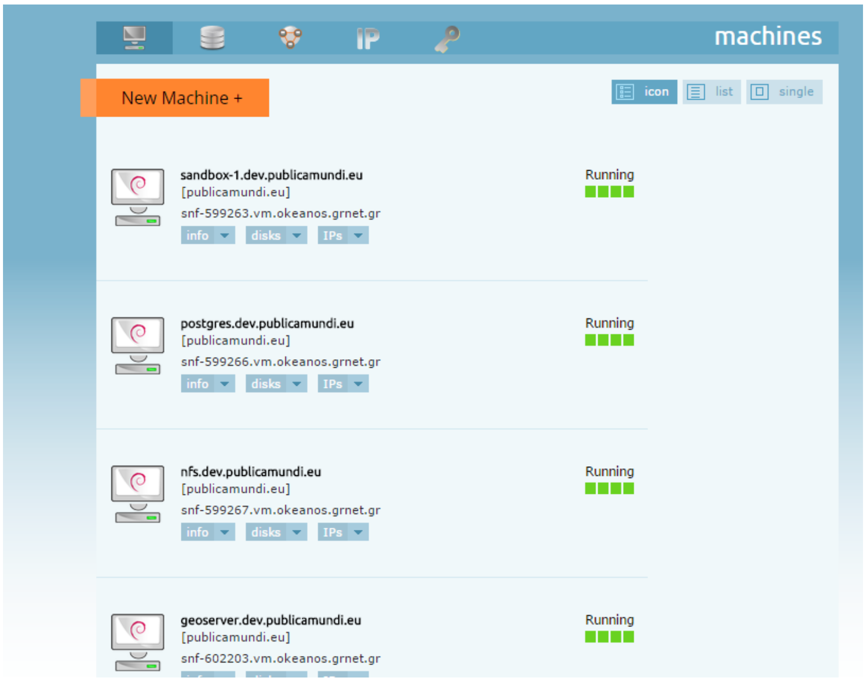


Figure 3: The main interface web page of Synnefo for administration of the cloud resources.

The software components of the system were deployed into 8 virtual clusters, with the provision of spinning up more virtual machines into each cluster if necessary. In Figure 4, the catalogue page of the integrated environment is presented.

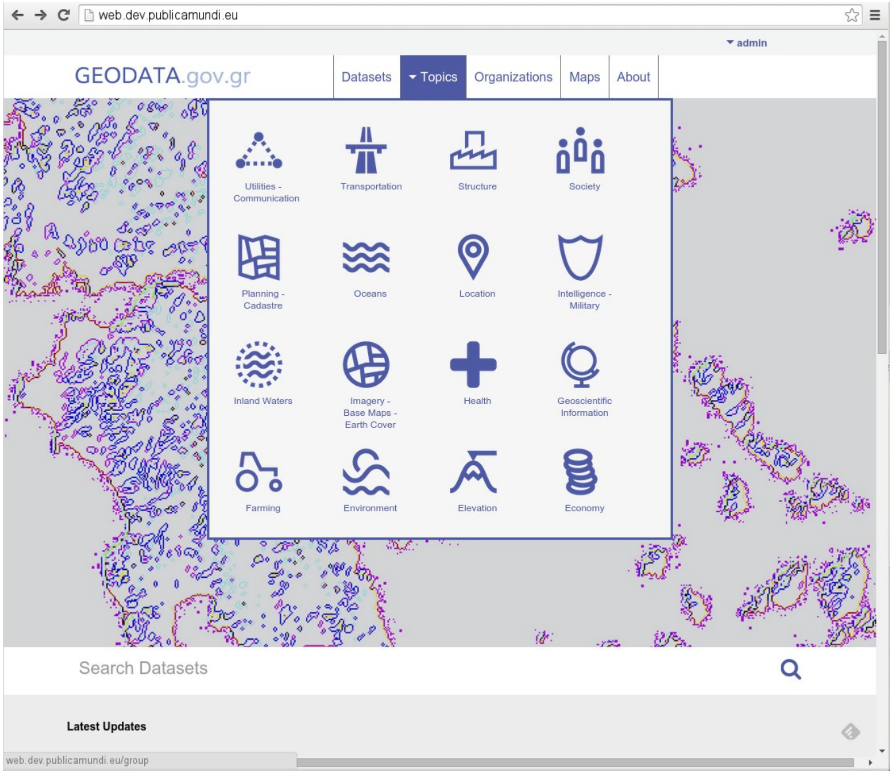


Figure 4: The catalogue front page in the integrated environment.

4 Geospatial extensions to CKAN catallogue

As described previously, CKAN is a Pylons/Python application, providing a mechanism to add functionality through extensions. This way, the core application is improved but at the same time the source code is easier to be maintained through mainstream code revisions (releases).

A main objective of the PublicaMundi project was to provide reusable and extensible software, while spatially enabling CKAN. Throughout the development process there was a focus on maintaining compatibility with upstream CKAN project, so that the new extensions would easily be applicable not only to geodata.gov.gr open data portal but also on other CKAN deployments, without breaking existing functionality.

During the development of the CKAN extensions significant progress was made in the metadata components, with pycsw (OGC CSW 2.0.2 reference implementation) being tightly integrated into CKAN. Initially, pycsw was loosely integrated under the CKAN spatial extension and used in the US data.gov project. The integration of pycsw into CKAN brought harvesting support of metadata from OGC services as well as support for INSPIRE and OGC OpenSearch Geo/Time specifications.

Moreover, CKAN was extended to support various established metadata models (including ISO-19115, ISO-19139 and INSPIRE), through OWSLib and pycsw. A new metadata extension was implemented (ckanext-schematic), providing the

capability for developing new metadata schema support in CKAN, through Pythonic Zope interfaces. This new extension also provides a way to easily create metadata editor user interfaces and extend the current CKAN publishing workflow. For geodata.gov.gr, the publishing workflow has been heavily extended, adding capabilities of automatic ingestion of geospatial vector data to a spatial database (PostGIS) and automatic publication of OGC services from within the CKAN administrative user interface (through GeoServer). Furthermore, CKAN was extended to support raster data as well (through rasdaman).
Data publishers are guided step-by-step into creating metadata in their schema of choice (Figures 5,6) or importing existing metadata. The metadata provided can then be transformed on-the-fly in any supported schema. Additionally, if an existing data catalogue or Spatial Data Infrastructure is available, the data publisher only needs to provide a simple entry point, and all available metadata are automatically harvested.
The new open data publishing workflow significantly lowers the entry barrier for data publishers, while also accommodating the data publishing needs of organizations with existing data catalogues and SDIs, maintaining full compatibility with INSPIRE and OGC web services.
For the management of data resources, a new administrator dashboard was implemented, which adds capabilities of creating new derivative resources (e.g. create a WMS resource from a GML file) (Figure 7).

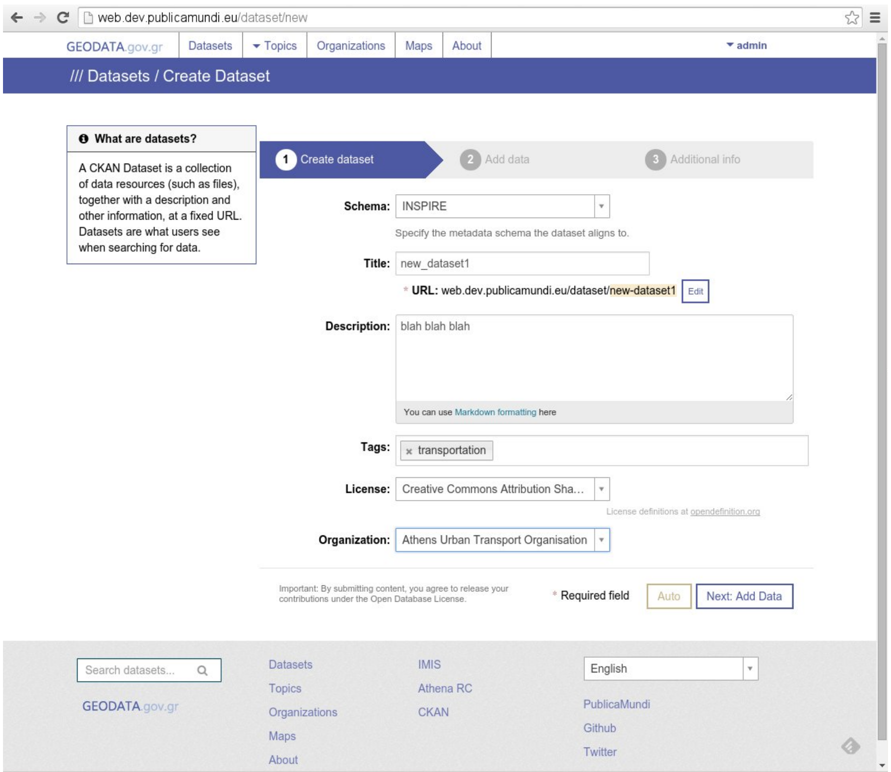


Figure 5: The publishing workflow to create a new INSPIRE dataset.

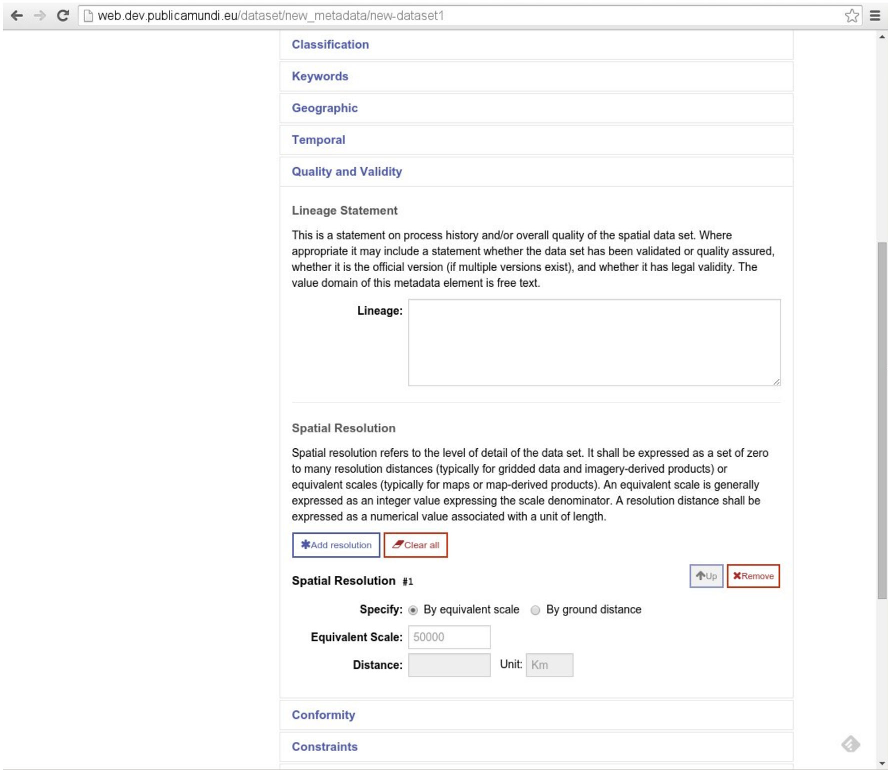


Figure 6: The PublicaMundi metadata editor (INSPIRE schema).

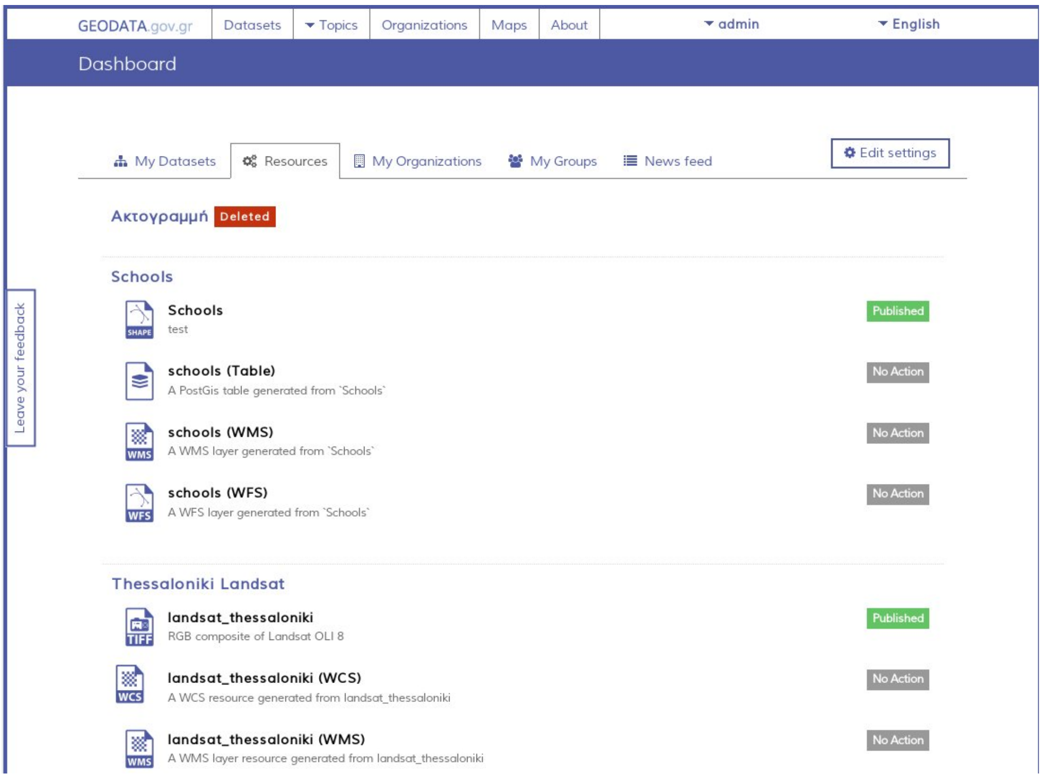


Figure 7: The PublicaMundi administration dashboard.

CKAN was extended to natively support geospatial vector data management, by integrating PostGIS, the leading open source geospatial database. Data publishers can upload geospatial data in any format and coordinate reference system. The system automatically stores the dataset and can provide it in another data format (on-demand) or through OGC compatible services. As such, data publishers can provide any data they have at hand, without additional effort into transforming their data in specific-purpose formats. Furthermore, as soon as the data is published, they are automatically available for querying and visualization with no extra effort through a Mapping API and a Data API (Figures 8 and 9).

As part of the CKAN spatial extension, the developed Mapping API was integrated in CKAN, thus providing data preview for geospatial data (Figure 8) . Also, in order to provide similar functionality as previous geodata.gov.gr, a full Javascript map client application was developed in CKAN with data search and map composition capabilities. (Figure 10).

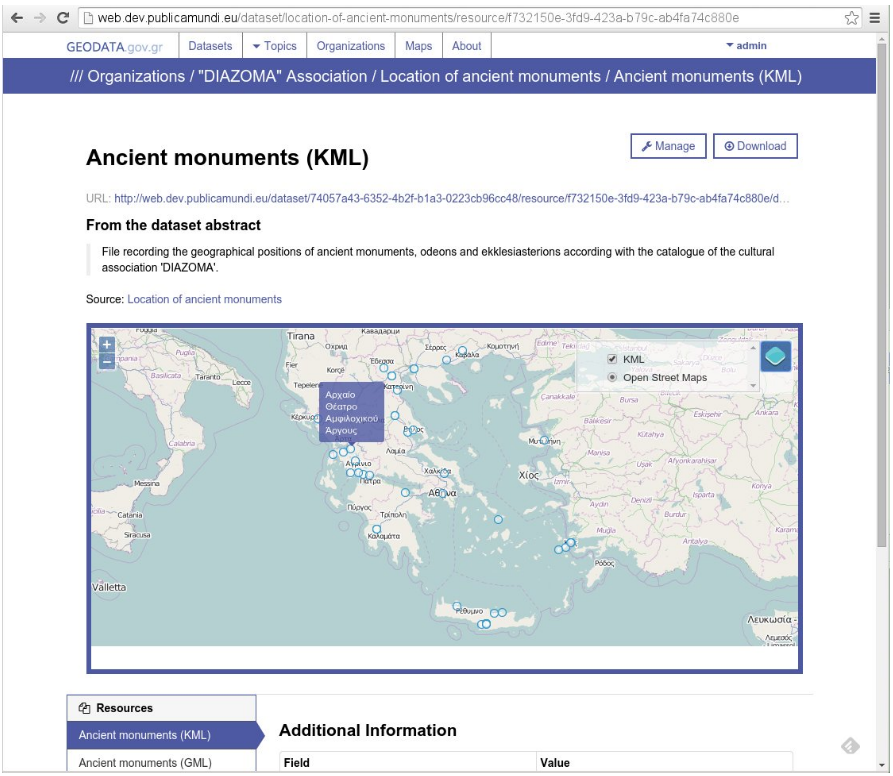


Figure 8: Geospatial data previewer based on PublicaMundi Mapping API.

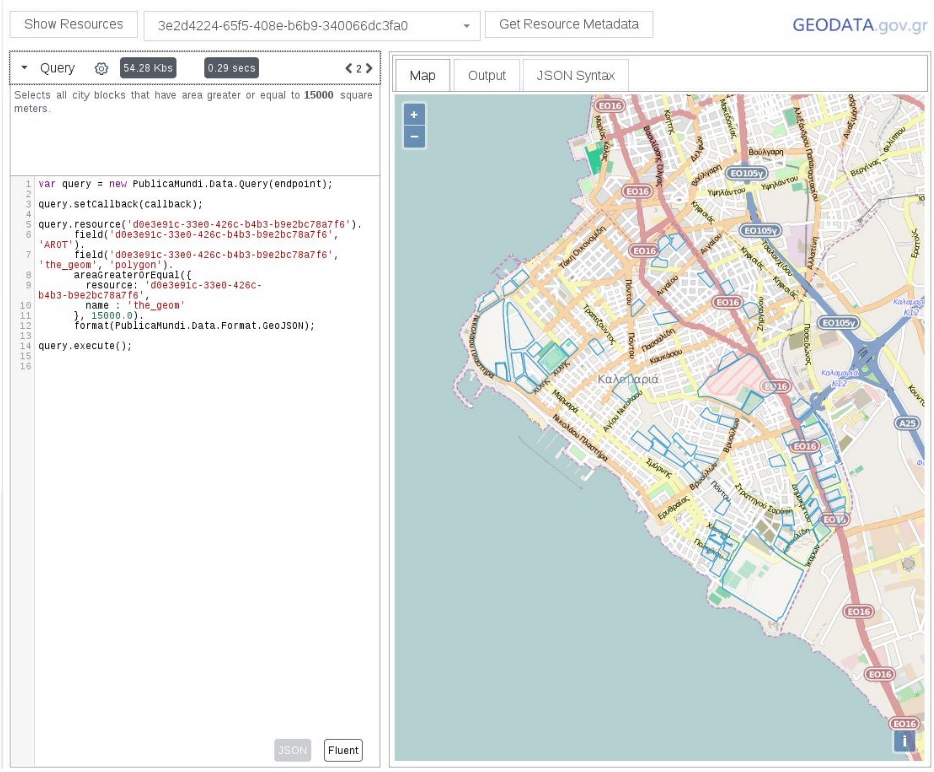


Figure 9: Demonstration of PublicaMundi's Data API.

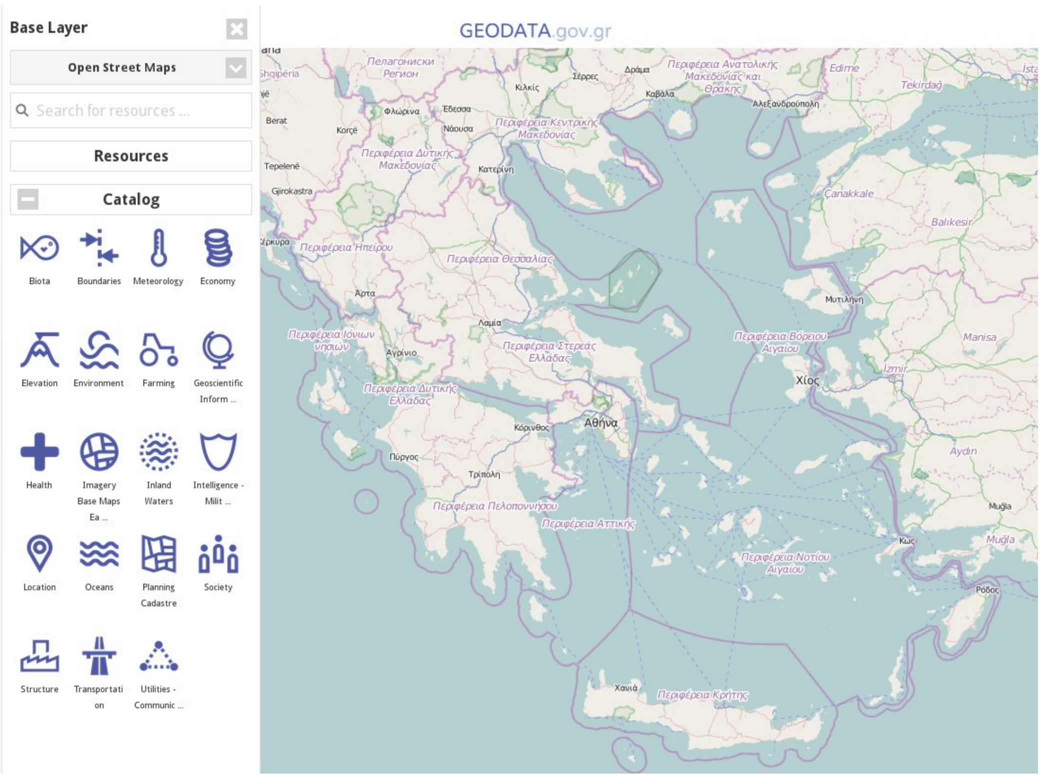


Figure 10: PublicaMundi's Mapping application.

5 Big Data support

CKAN has been succesfully integrated with the rasdaman raster data engine (through a CKAN extension named raster-storer), adding big data capabilities to the PublicaMundi's software stack, providing the ability to serve hundreds of TBs of Earth Observation data through the WCS OGC API or through PublicaMundi's Data API.
The rasdaman system is a scalable, multi-dimensional array database engine and analytics server. Rasdaman is a domain-neutral Array Database System: it extends standard relational database systems with the ability to store and retrieve multi-dimensional raster data (modeled as arrays) of unlimited size through an SQL-style query language. Array intensive services can be set up using rasdaman, such as those found in data centers offering remote sensing, sensor series, image, simulation, and statistics data. Domains taking advantage of an Array Database can be Earth, Space, and Life sciences. The query language offered over arrays enables flexibility, speed, and scalability.
On top of the Array Database System, a geospatially enabled layer is provided that adds spatiotemporal semantics over array data. Within the geo layer fall the petascope components and the related geo-import tools (currently known as rasgeo and wms-import). Geo-enablement of rasdaman provides access to geo raster standards-based services, including OGC WCS and WCPS, the OGC raster query language, along with WCS-T, and WPS. A key advantage of rasdaman is its internal management of data based on n-D tiling which allows data to be stored in the most convenient way, depending on the desired performance profile along properties, such as fast retrieval along some dimensions and fast data ingestion rates.
The PublicaMundi project is based on CKAN as a publishing platform that streamlines the process of publishing, sharing, finding and using data. Several CKAN extensions were developed to allow for ingestion of georefferenced data, be it vector or raster. For raster and earth observation data in particular, a raster storer CKAN extension was developed, in order for CKAN to be able to automatically ingest raster data and expose them through a variety of OGC services. In this context, a WCS-T standard was also implemented to allow simple ingestion of raster data into rasdaman.
The raster storer plugin receives datasets (or links to datasets) from the user, containing data in various formats (GeoTiff, JPEG2000, zip archives etc) and, using the GDAL library, reads or deduces all the necessary metadata directly from the file and creates a GML file that references the original file as the source of the raw data. This provides an important degree of flexibility. Rasdaman is loosely coupled with CKAN, allowing the use of this ingestion method in other projects that will not use CKAN as the front-end. Data is never copied or moved around, the GML file only references the original file, allowing rasdaman to read its contents and ingest it in its internal format.

6 Processing engine

An important part of the PublicaMundi system, is the ability to re-use the catalogue's open data and create new datasets with added value. To this purpose, a geospatial processing engine was implemented as part of the

system. The open source WPS implementation ZOO-Project was used at different levels of the PublicaMundi project, especially to address the goal of providing a reliable and scalable processing engine to PublicaMundi architecture, as well as a collection of generic and user-centric WPS services.
According to the OGC, the WPS standard provides a set of rules for standardizing inputs and outputs (requests and responses) of geospatial processing services. The standard defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes and clients' discovery of and binding to those processes. WPS thus allows to process geospatial data over the Internet and ZOO-Project implements it through an open architecture and a modular source code mostly written in C, Python and JavaScript.
In the presented framework, a WPS extension of the CKAN software was implemented in order to provide a bridge between the Data Catalog and the Processing Engine. The so-called ckanext-wps module brings WPS capabilities to the CKAN powered open data catalogues by introducing the possibility of adding an existing WPS server as a new CKAN resource from the CKAN administration panel. The support for WPS GetCapabilities, DescribeProcess and Execute requests support was implemented using the Python language and the CKAN logic in order to create a suitable extension. The Processing Engine can thus be directly exploited from CKAN and this allows to directly process open data and create, store, and reference new information directly from the Catalog. Every WPS service returned by a GetCapabilities request can be configured through appropriate CKAN forms and executed using any suitable geospatial resource available from the raster and vector data stores as input data. The output data finally generates a new CKAN resource and add it to the targeted data store. Further improvements are still under development and their release is already planned for the next phase of the project.

6 Future work

PublicaMundi project is still under heavy development. Future work includes multilingual support for both data and metadata storers, as well as utilities to enable publishers to crowdsource their translations of metadata and data. At the same time, work is being done to implement the upcoming OGC CSW 3.0 and WPS 2.0 standards, as well as the DCAT Application Profile developed by the European Commission. Furthermore, work is being done for data interlinking and integration of Data and Mapping API with WPS and WCPS.

References

- Baumann, P. (2014, August). rasdaman: Array Databases Boost Spatio-Temporal Analytics. In *Computing for Geospatial Research and Application (COM. Geo), 2014 Fifth International Conference on* (pp. 54-54). IEEE.
- Koukis, V., Venetsanopoulos, C., & Koziris, N. (2013). Synnefo: A complete cloud stack over ganeti.; USENIX, login 38 (5): 6–10.

Finding spatial open data via web: a SWOT analysis

Marco Piragnolo[1,2], Francesca Fissore[1], Alberto Guarnieri[1], Antonio Vettore[1], Francesco Pirotti[1,2]

[1] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, Viale dell'Università 16, Legnaro (PD) 35020, Italy

[2] Department of Land Environment Agriculture and Forestry (TESAF), University of Padova, Viale dell'Università 16, Legnaro (PD) 35020, Italy

Abstract

In the following paper we report the experience of first-time users looking for open data over a dedicated web portal which provides a catalog from third party providers in different formats. A SWOT analysis is applied to assess how accessible, intuitive and helpful were the tools and the data that are provided. We provide constructive criticism on weaknesses and discuss tentative solutions to problems encountered, which can be the starting point for further improvement of such services.

Keywords

National catalog, web access, open data, big data, SWOT analysis

1 Introduction

The data in public administrations are going in the direction of greater distribution and further "open" paradigm, in terms of licensing and usage. Open spatial and non spatial data are following the silent revolution of open source spatial and non-spatial software. Some years ago open source geographic information systems (GIS) – and related software which generally deal with spatial domain - were exploited by a small number of users, due to the steep learning curve and the lack of graphical user interfaces (GUI) which decreased the appeal from less motivated users.
Lately, "smart" oriented technologies, which are increasing exponentially, can leverage open data to enrich links between the data they provide and data from other providers, with added value for users. In the digital era new technologies for data acquisition, management and analysis provide high volume, variety and velocity of data - i.e. big data (Craglia et al., 2012; Ehlers, Woodgate, Annoni & Schade, 2013). Examples are numerous, and range from road accidents positioning for estimating road risk (Pirotti, Guarnieri & Vettore, 2011) to complex infrastructures for defining energy saving strategies (Schaffers et al., 2011); more data, faster update, and crowd sourcing is becoming a common way to enrich datasets.
The open data paradigm is often used for "promotional" reasons, but it is actually a complex principle, which follows a number of rules, licenses, and is

often a matter of debate over the different typologies thereof (Lopez-pellicer, Florczyk, Nogueras-iso, Muro-medrano & Zarazaga-soria, 2011; Shadbolt et al., 2012). Nevertheless a very important aspect is often overlooked: how accessible is the data, and how usable is it for an average user interested into integrating open data in its projects? In this paper we want to consider this problem by assessing a practical approach of using an open data portal provided by the Italian government. To provide a structured report, we apply a SWOT analysis over our experience on the utilization.

2 Method and analysis

In this work we address the following questions: (i) how can open data be exploited, e.g. using a webgis service that provides services to municipalities; (ii) what are the strengths and weaknesses of the work-flow required to access to the data.

2.1 Consultation of existing online catalog

In this study case we accessed the data base which is available in the web portal which collects open data from public administrations (PA) for the Italian Government, (http://www.dati.gov.it). The web page is well structured and provides a menu with several links allowing access to data catalogs, software, information and news related to Open Data, as well as a tool called Infografica.

2.1.1 Data catalog

The data catalog has several embedded tools: a search bar for free text search, a map-based filter, a filter with a criteria based on organizations which source the data, a filter using tags, and a filter based on data format and licensing definition. The search field allows you to query using keywords which define the domain of interest. The map filter allows to make a selection using spatial criteria. The source filter allows you to limit results based on the organizations related to public administrations that provide data. Operationally the search is direct and simple, but we note that the list is not in alphabetical order and lacks a tool for ordering results; this capability would support an easier identification of the data that are of interest, as it is now, the user has to navigate through all results. The same problem occurs for the filter by tag and by format. Another point that can be improved is the definition of elements. for example, looking for a Shapefile format you see different names belonging to the same format, i.e. “SHAPE”, “shapefile”, “shp” and “ZIP”, the latter being a file that contains inside a Shapefile. This adds complexity when obtaining data and can generate confusion in itself. Such problem is part of a complex principle related to how keywords are linked to each other and what they really mean; e.g. “shape” can intend the Shapefile or can intend the denotation of the dictionary definition of shape. There are studies over the semantics of words over the web and how they link to each other. That is also the reason that a lot of effort is applied to finding standardizations and harmonization of data representation, and also data description (e.g. the so called metadata). The first step towards this is to make it easier for machines to understand the data that we provide so it can be easier to structure and connect (Gomez-Perez, Fernández-López & Corcho-Garcia, 2004).

0/ (47)	jpg (15)	SHAPE (221)
7z (99)	JSON (3171)	shapefile (13)
altro (69)	KML (112)	shp (511)
api/sparql (16)	kmz (205)	shx (13)
application/rdf+xml (12)	n/a (7)	tif - tfw (5)
blender v.2.49b (26)	n3 (27)	tsv (51)
CSV (6458)	n-triple (7)	turtle (34)
csv-semicolon delim... (28)	ods (40)	TXT (16)
csv-tab delimited (13)	org/ (16)	wms (367)
dbf (12)	owl (8)	XLS (2949)
dxf - dwg (4)	PDF (346)	xlsx (20)
dxf (7)	php (9)	XML (3001)
geojson (58)	rar (6)	xsd (6)
gml (5)	RDF (1989)	ZIP (1094)
grid (4)	sbn (4)	zip (shp) (21)
grid esri (4)	sbx (4)	
HTML (1434)	service (89)	

Table 1: filter by format with results and in parenthesis the number thereof. The highlighted rows are the formats that present analogies with each other. In an unordered list the user may have difficulty in identifying all the resources available.

The datasets are then available for download in various formats. As for geospatial data, in a geographic information system (GIS) they can also be accessed through OGC compliant services (i.e. WMS, WFS and WCS) that enable users to connect remotely to the server of the service-provider to read the data. This approach has the advantage of a real-time link to the data with the latest version without having to download new versions. The catalog has WMS services, related to raster representation, while it is missing WFS and WCS services. The suggestion is that the government supports the development of such geospatial services, which allow for greater dissemination and data interoperability. For clarity it must also be said that a real-time access to data through services would require significantly high bandwidth. Nevertheless the future should see funding towards broad-band internet, thus this problem should not become a limiting factor. Similarly it would be interesting to improve the availability of SOS services for sensors, i.e. web services which are structured to provide values from measurements from time series, interoperable and accessible (Bröring, Jirka, Kotsev & Spinsanti, 2013). Often this type of data are provided in a non-structured tabular or semi-tabular format, via spreadsheets or published directly on the web via HTML. Such representation requires considerable effort for manipulation and formatting - the so called extract transform and load (ETL) procedures, which have a very significant impact in terms of time (thus money) (Vassiliadis, Simitsis & Skiadopoulos, 2002). Additional suggested filters are those that select using administrative limits i.e. municipalities, province or region and a temporal filter i.e. year.

2.1.2 Software

The software page lists all the executables which have been developed by public administrations and which are available to third party. There is only a single query filter with keywords.

2.1.3 Infografica tool

The Infografica page (Figure 1) provides a real-time overview of the status of open data in Italy, both graphically and through an interactive map that shows the geographical distribution of the data sources. The tools allow you to have a graphical overview, an intuitive search panel of the data store, without necessarily having to scroll through the list. It is a particularly interesting tool for navigating complex datasets. In the near future it should increase in terms of richness of contents; it already has the connotations of “big data” both in terms of physical space and also in terms of the number of results obtained by the queries, the quality (Singh & Singh, 2012) the challenges regarding its maintenance, structure, navigation and data visualization (Fox & Hendler, 2011; Choo & Park, 2013; Sagiroglu & Sinanc, 2013;). Particularly interesting is the filter on the reusability of data, which allows to focus and highlight the virtuosity of the public administrations which provide high-quality data and continuity. This tool can also be stimulating towards a continuous improvement of the datasets. The service could be ameliorated by placing a link to the dataset inside the portal itself, and not only by forwarding the user to the page of the institution/public administration.

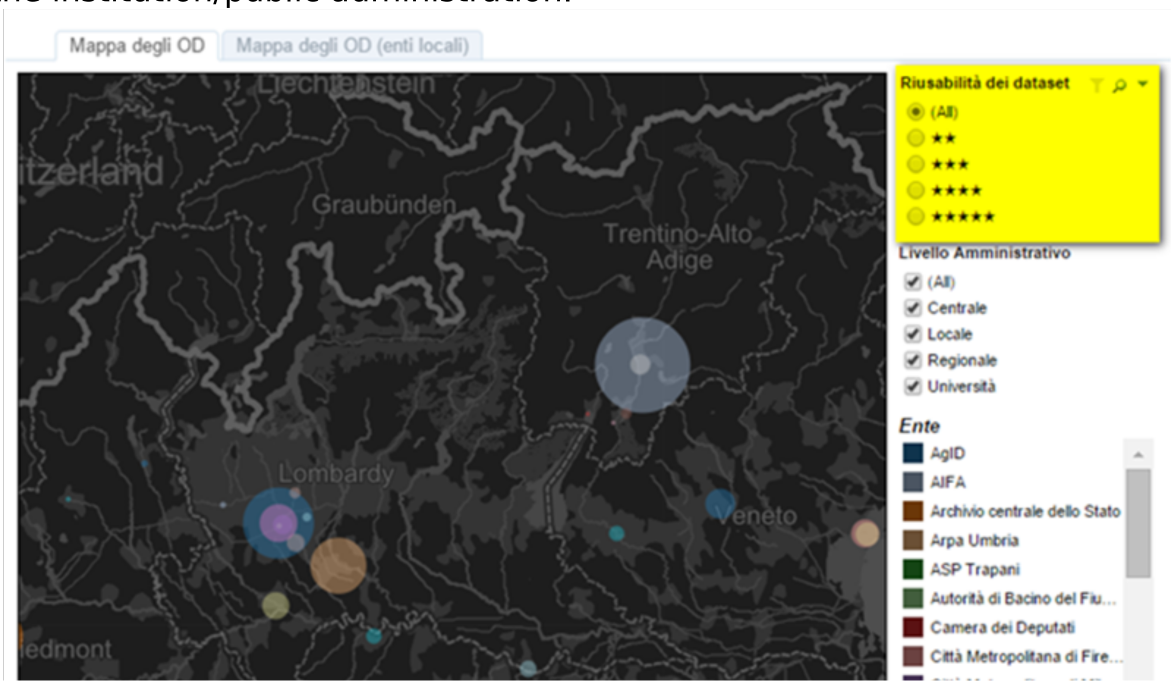


Figure 1: Infografica tool.

2.1 Data searching over an area

In this short test we focused on finding data in shapefile format and as OGC web services for the area corresponding to the province of Brescia. We tried to use a filter on the map to find data in the catalog and we note that we can only make selections with a rectangular bounding box. This poses a problem in the sense that if we select a large territory we get an overabundance of data. An interesting alternative would be to select the data set based on municipal boundaries, regional or provincial, therefore a spatial geo-processing intersection. In the field of free research we then type "Brescia" getting ten datasets: eight correctly from the Lombardy Region, one for the Liguria Region and one for the City of Vicenza; the latter two false positives (Figure 2).
The result for the region of Liguria, a false positive, the figure refers to a "digital color orthophoto of Portofino Park 1:2000 - flight on 10 June 2004". Made by a company called "Rossi of Brescia". This proves that simple key words are not enough for a clear result. The other result, the false positive related to the city of Vicenza, we get the following metadata: "an important project of digitization of rare and valuable ancient documents, a collection which is in the Bertoliana public library which was funded with funds from the sale of the shares of the West Brescia-Padova company owned by the town of Vicenza". The research of data works correctly only partially, and this leads to require further filtering to remove false positives, which are not present in the area of interest. This implies the user having to manually fulrther filter the data. As for the available format of the data there are the following possible formats: ZIP, XML , XLS , RDF , JSON , CSV , HTML.

Figure 2: right - results from searching with keyword "Brescia". Bottom left are highlighted the results. Left - formats with in parenthesis the number of results related to the format.

At a first glance we do not see Shapefile formats, but at a closer look we see that the details related to the ZIP format show that it contains the Shapefile format (Figure 3). This is a limitation as ZIP files might not be linked to their actual format which they contain, thus remaining in the false negatives results.

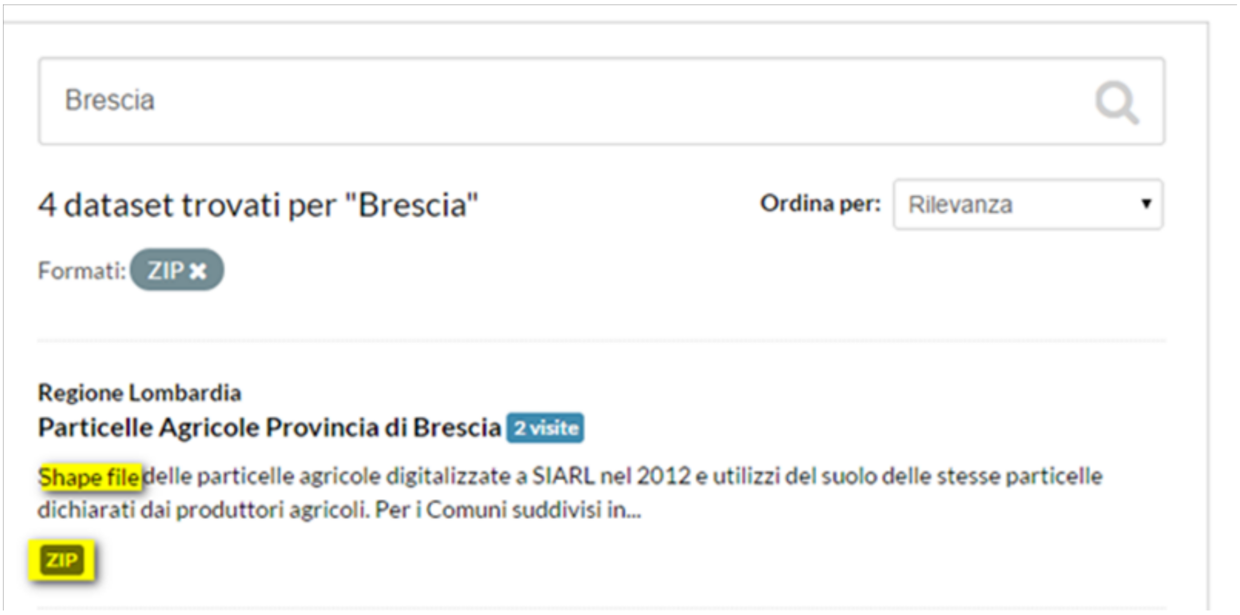


Figure 3: ZIP result containing a Shapefile.

3 Discussion and conclusions

In this section we discuss strength, weaknesses, opportunities and threats (SWOT) that, to our opinion, have come to the surface from our experience in this study. It must be stressed that our point of view is that of a user with proficiency in programming webgis platforms, that addresses as source open data from a portal.

One point is the accessibility and degree of ease for finding the data that one is looking for. For example, in our study case, we are interested in data relative to municipalities of the Lumbardy Region which are directly georeferenced. Some data are somehow geo-coded in the sense that each information (e.g. row) has a reference to the municipality. For example the number of citizens in each town is geo-coded in the sense that the town's ID allows it to be linked to a geographic element (town boundaries, area or baricenter). The available format also makes a difference. The industrial area for each town in a region is available, only in shapefile. This implies that to integrate in a spatial data infrastructure the file has to be manually downloaded, and linked.

The portal allows to monitor the distribution of open data in the Italian territory, to evaluate its quality, geographical spread as a kind of container , a kind of index to Reference to try all useful resources . To assess the portal and the catalog we found some strengths and some weaknesses. Therefore we applied a SWOT analysis.

Strengths	**Weaknesses**
Spatial filter	Spatial selection only with bounding box
Organization filter	Selection tools are not always accurate
Tag filter	Limited number of filters
Format filter	Query results cannot be ordered
License filter	Only WMS services available
Opportunities	**Threats**
Improve the selection panel / toolbar	Complex development of tools for selecting using other layers (e.g. municipality area, province area etc..)
New filtering possibilities	Temporal filter over non-continuous data might be not easy to implement
Improve management of resources	Harmonization of definition of data format requires ad hoc project
Incentive to provide more services	Tags and keywords harmonization requires investment in a specific project
Insert new tools for reading the data	The lack of WFS, WCS, SOS services will decrease the number of potential users
Further integration with the Infografica tool	Lack of quality control / quality assessment

The SWOT analysis of the Infografica tool gave the following results:

Strengths	**Weaknesses**
Allows to assess the spatial distribution of all the open data available	There is no zoom with the mouse wheel, which is a habit by users of webgis
Map has navigation capabilities	Zoom and selection panel not easily accessible
Map allows intuitive searching of the catalog	There is only a link to the pubic administration with the data, thus an indirect link
Allows to assess the spatial continuity of the spatial data available	Missing the link with resources available directly from the portal
There is a filter for quality of the data	
There is a filter for the re-usability of the data	
Opportunities	**Threats**
Integration with new resources	Potential complexity with the increase of the volume of data
Improvement of selection panel	Missing filter using temporal criteria
Improvement of selection filters	Data visualization becomes difficult
Inserting applications	

References

- Bröring, A., Jirka, S., Kotsev, A., Spinsanti, L. (2013). Making the Sensor Observation Service INSPIRE Compliant. *Proceedings of the INSPIRE Conference*, Florence, 27th of June 2013.
- Choo, J., & Park, H. (2013). Customizing computational methods for visual analytics with big data. *IEEE Computer Graphics and Applications*, 33, 22–28. doi:10.1109/MCG.2013.39.
- Craglia, M., de Bie, K., Jackson, D., Pesaresi, M., Remetey-Fülöpp, G., Wang, C., … Woodgate, P. (2012). Digital Earth 2020: towards the vision for the next decade. *International Journal of Digital Earth*, 5(1), 4–21. doi:10.1080/17538947.2011.638500.
- Ehlers, M., Woodgate, P., Annoni, A., & Schade, S. (2013). Advancing Digital Earth: beyond the next generation. *International Journal of Digital Earth*, 7(1), 3–16. doi:10.1080/17538947.2013.814449.
- Fox, P.; Hendler, J. (2011). Changing the equation on scientific data visualization. *Science*, 331, 705–708.
- Gomez-Perez, A., Fernández-López, M., & Corcho-Garcia, O. (2004). Ontological engineering. *Computing Reviews*, 45(8), 478-479.
- Lopez-pellicer, F. J., Florczyk, A. J., Nogueras-iso, J., Muro-medrano, P. R., & Zarazaga-soria, F. J. (2011). Linked open data for INSPIRE : From 3 to 5 star geospatial data 5 star Linked Data. *INSPIRE Conference 2011*. Edinburgh.
- Pirotti, F., Guarnieri, A., & Vettore, A. (2011). Collaborative Web-GIS Design: a Case Study for Road Risk Analysis and Monitoring. *Transactions in GIS*. Wiley-Blackwell Publishing Limited, United Kingdom. doi:10.1111/j.1467-9671.2011.01248.x.
- Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. In International Conference on Collaboration Technologies and Systems (CTS) (pp. 42–47). doi:10.1109/CTS.2013.6567202.
- Schaffers, H., Komninos, N., Pallot, M., Trousse, B., Nilsson, M., & Oliveira, A. (2011). Smart cities and the future internet: Towards cooperation frameworks for open innovation. *Lecture Notes in Computer Science*, 6656, 431–446. doi:10.1007/978-3-642-20898-0_31.
- Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., Hall, W., & Schraefel, M. C. (2012). Linked open government data: Lessons from data.gov.uk. *IEEE Intelligent Systems,* 27, 16–24. doi:10.1109/MIS.2012.23.
- Singh, S., & Singh, N. (2012). Big Data analytics. In *International Conference on Communication, Information & Computing Technology (ICCICT)* (pp. 1–4). doi:10.1109/ICCICT.2012.6398180.
- Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). Conceptual modeling for ETL processes. In *Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP*-DOLAP '02 (pp. 14–21). doi:10.1145/583890.583893.

A Data Scientist Exploration in the World of Heterogeneous Open Geospatial Data

Gloria Re Calegari, Irene Celino

CEFRIEL – Politecnico di Milano

Abstract

Open Geospatial Data abound today because digital technologies are more and more pervasive. The rich ecosystem of data producers and data consumers presents a multi-faceted picture of the environment, because heterogeneous data sources differ in terms of data complexity, spatio-temporal resolution and curation/maintenance costs.

In this paper, we present the challenges faced by a Data Scientist in exploring and analyzing heterogeneous Open Geospatial Data. This work is aimed at explaining the initial steps of a data exploration process, specifically aimed at discovering similarities and differences conveyed by diverse sources and resulting from their correlation analysis; we also explore the influence of spatial resolution on the dependence strength between heterogeneous urban sources, to pave the way to a meaningful information fusion.

Keywords

Open Geospatial Data, smart city, data diversity, spatial data resolution, correlation analysis

1 Introduction

The ever-increasing availability of Open Geospatial Data calls for smarter processing approaches. A Data Scientist is of course happy in front of such a wealth of knowledge about the space that surrounds us. Still the geospatial nature and the heterogeneity of such data provide interesting challenges for the data exploration.

Heterogeneous geospatial data describes the environment through the lenses of a multitude of information sources. Consequently, the collection, cleansing, curation and maintenance of specialized data sources results in a complex and expensive process, sometimes requiring manual intervention, sometimes involving error-prone automatic processing. The diverse origin of different datasets referring to the same environment represents a further challenge, because the data producers often have disparate objectives and goals that are reflected in the data itself. For a Data Scientists, therefore, the first step is to analyze available datasets in order to identify potential intrinsic dependence and inter-relationships between them. Since their heterogeneous provenance reflects specific and distinct perspectives, we need to investigate whether and how to reconcile the possibly diverging “pictures” of the urban environment those sources convey.

In this paper, we focus on the data exploration of diverse information sources that refer to the same environment; we present the issues and difficulties of such an analysis and we provide a possible best practice, based on the use of free and open source software tools, to start a process to get a deeper understanding of an environment starting from the available Open Geospatial Data. While our current work focuses on urban environments – because we believe this kind of analysis is of paramount importance in the context of the so-called Smart Cities – the Data Scientist's approach here introduced can be applied also in different spaces.

The remainder of the paper is as follows: Section 2 introduces the characteristics of Open Geospatial Data sources and the main challenges in their processing; Section 3 details the information sources about Milano used in this research; Section 4 illustrates our experiments of correlation analysis applied at different spatial resolution levels, and Section 5 concludes the paper with some perspectives on our future work.

2 Availability of Open Geospatial Data and Challenges in their Analysis

Open Geospatial Data about cities abound today. The sources of such information are constantly increasing, due to the pervasiveness of information and communication technologies in the so-called Smart City domain. In this section, without the claim of being comprehensive, we would like to give an overview of the possible urban-related datasets that can be found today and of some of the challenges in those datasets management, manipulation and analysis.

With the advent of the *open data* movement, with its call for transparency and knowledge sharing, a very large number of data sources has been made available on the Web, through a new generation of CMS systems able to give access to this wealth of information, often originating from public bodies and research activities. With special reference to urban information, local authorities have started publishing numerous datasets referring to the city environment: demographics and statistics from municipalities (e.g. distribution of population, family income, crime statistics), listing of local businesses from chambers of commerce, various levels of descriptions about the environment from an urban planning perspective (e.g. land use or land cover, cadastre information), and so on. Those data sources share the geographic aspect even if they are originated by different actors with different purposes.

Furthermore data coming from *private businesses* that used to be closed data sources are nowadays becoming available, sometimes even as open data. Examples of this kind of datasets are public utilities information, including telecommunication operators: as collateral effect of the mobile networking services, telco companies collect data about the phone activity over time and also over space (due to the positioning of transceiver towers). This kind of data can provide specific perspectives on what happens in our cities and can be a strong indicator of people presence and movement in the urban environment.

Moreover, the popularity of sensor technologies and the so-called Internet of Things (IoT) has led to the availability of massive *real-time and streaming information*, like climate sensors from environmental agencies (e.g. temperature, pressure, humidity and other ecosystem measures), smart meters and GPS traces from public utilities (e.g. energy consumption or public transportation position).

After the Web 2.0 boom, also *user generated information* about cities has become ubiquitous. Crowdsourcing initiatives like OpenStreetMap[1] have popularized the Volunteered Geographic Information paradigm of "citizens as sensors" (Goodchild, 2007) and have collected data about different kinds of points of interest in urban environments (e.g. monuments, restaurants, public services). Location-based social networks like Foursquare, Twitter, Flickr have also produced a stream of "check-ins" and geo-located information that represent the digital counterpart of human activities in the urban space.

Managing, processing and comparing those diverse Open Geospatial Data can be cumbersome for a data analyst. Besides common issues like dealing with data scale and improving data quality, we would like to highlight some challenges that emerge from the specific case of comparison between datasets referring to the same geospatial environment.

One issue arises from the *varying spatial resolution* of information sources: being produced by diverse actors for different reasons, it is quite common that datasets are heterogeneous in terms of the geospatial extent they refer to. For example, population statistics could be at municipality level, land use information from cadastre could be at building level, and smart meters measures could refer to individual points (identified by latitude and longitude coordinates). This means that those sources are not immediately comparable.

Information sources can also refer to *different time-frames*: population census is usually done every *n* years, while sensor information is potentially provided in real-time; some other data sources can be made available as historical dumps, while IoT data can have different frequency updates (every 10 minutes vs. once a day). Directly comparing those sources can lead to poor results, because connections and correlations could be traced between datasets that give different pictures about the environment.

Moreover, because of the time-frame, as well as because of the data provider, data sources can differ in terms of *reliability*: apart from pure data quality issues, in managing and processing different datasets it is important to take into account whether and to what extent the source can be trusted; for example, while data from public authorities can be considered "official", information deriving from user generated efforts or crowdsourcing campaigns can be less dependable.

For a data analyst, the first task is to prepare urban datasets for the subsequent elaboration. Besides common activities like data transformation, cleansing or normalization, we focus our attention on two specific cases that are particularly relevant in a Smart City scenario.

One issue when dealing with diverse data sources is the heterogeneity of their representation in terms of geospatial resolution. The first activity therefore is to *uniform spatial resolution*, finding a trade-off between fine-grained and coarse-grained data. This implies interpolating, aggregating or splitting information in different datasets (Gotway & Oung, 2002) so that they become comparable in subsequent elaborations.

In today's Smart Cities, information is often characterized by the temporal dimension, as in the case of IoT sensor data and real-time sources. While having both time and space dimensions enriches the city representation conveyed by those sources, time-series frequently consist of big data (Kitching, 2014), thus they require specific processing techniques. In the context of an exploratory analysis of urban data aimed at comparing different datasets, a fundamental activity is therefore *data compression*: pre-processing large-scale

[1] www.openstreetmap.org/

time-series to get a more manageable compressed representation that can be used to seek relations with other sources. Several approaches to temporal data compression exist (Mitsa, 2010); the main idea is to reduce high-dimensional temporal sequences to compact representations that summarize the original data into a pattern. The simplest example is the computation of a *temporal signature or footprint*: a 1-dimensional vector in which each element represents the average behaviour of the measured quantity over a specific time interval. Common footprints represent daily patterns, possibly distinguishing between different days of the week; the interval granularity greatly depends on the nature of data (e.g. sensors' frequency of measure).

3 Overview of the Milano Datasets employed in our work

Our case study deals with the exploratory analysis of diverse urban datasets related to the municipality of Milano in Italy.
The datasets used in the analysis are illustrated in Table 1: the open data about population demographics from Milano municipality[2]; the land use classification elaborated within the CORINE European initiative[3] with CORINE multi-level taxonomy of land cover and made available as open data by Lombardy Region[4]; the Points of Interest (POIs) of the city provided by both Milano municipality and OpenStreetMap; two months of mobile call data records provided by the Telecom Italia mobile operator for their “Big Data Challenge”[5] as Open Data.

Domain (content)	Data Source	Data Format	Spatial Resolution	Reference Period	Volume (records)
Demographics (population)	Milano Open Data	Shapefile	Census area	2011	10s
Urban Planning (CORINE land use)	Lombardy Open Data	Shapefile	Building-level resolution	2012	10 Ks
Mobile Telephony (call records)	Telecom Open Data	Tabular	City grid cells (250m x 250m)	2013 Nov-Dec	100 Ms
Points of Interest (POIs)	Milano Open Data	Shapefile	Points (lat-long)	2013	1 Ks
Points of Interest (POIs)	Open Street Map	Shapefile	Points (lat-long)	2014	1 Ks

Table 1: Characteristics of the used datasets.

Telecom dataset records every ten minutes the activity occurred in Milano in Nov-Dec 2013 (60 days), and maps it into a grid of squared cells of 250 m. Five different phone activities are stored: incoming and outcoming calls, incoming and outcoming SMSs (text messages) and Internet connection. Because of the size (100Ms data records) and complexity (including both temporal and space dimension) of the dataset, a number of pre-processing operations were required.

[2] http://dati.comune.milano.it/
[3] http://www.eea.europa.eu/publications/COR0-landcover
[4] https://www.dati.lombardia.it/
[5] http://theodi.fbk.eu/openbigdata/

To reduce Telecom dataset size and to take into account the spatial information, we decided to compress all the data of each cell into a “footprint”, i.e. a summarizing data structure which records for each time slot of ten minutes the average activity of that cell, like the one shown in Figure 1.

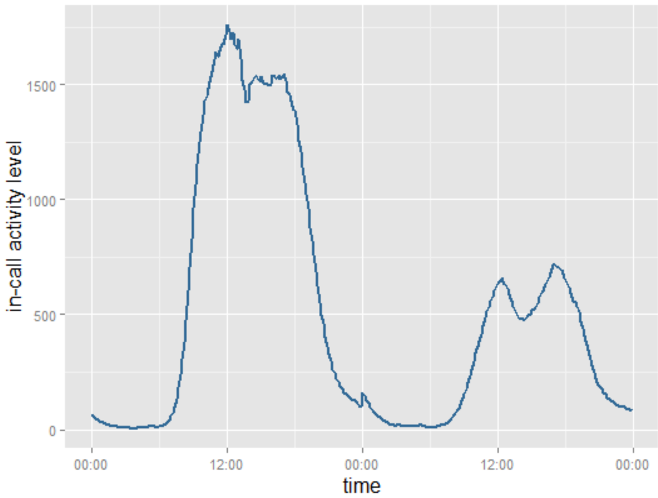


Figure 1: Telecom footprint of call-in activity. Concatenation of weekdays and holidays activities.

In particular we averaged the values recorded in the same cell and at the same time during all the day of the period under analysis, obtaining 144 values for each of the 5 phone activity types (720 values total). We successfully performed all these computationally intensive operations using R free software for statistical computing[6].

As Figure 1 reveals, we treated weekend and weekday data separately, obtaining a double footprint, one describing weekday activity and the other the holiday behaviour.

Actually, as literature confirms (Reades, Calabrese, Sevtsuk, & Ratti, 2007), weekend and weekdays data usually displays high differences in both shapes and magnitudes.

CORINE 2012 dataset provides data about the *land use* types of Milano territory and it classifies them by using CORINE multi-level taxonomy. Land use types can range from more general definition of residential, agricultural, industrial, wild areas to a more specific characterization as hospitals, roads, railways, construction sites and so on.

The dataset has a fine grained spatial granularity, actually the land use is defined at a building-level resolution.

After analyzing the distribution of the different land uses in the Milano area, we selected the level of taxonomy more suitable for our case study (by selecting specific categories or by grouping some classes together). We decided to start analyzing the two most general types of land use that could identify and feature a metropolitan area: the residential and the agricultural areas.

As regards *demographics information*, the dataset provides information about the number of inhabitants for each of the 6079 census areas of Milano municipality. A census area is a partition of the territory defined by the Italian National Institute of Statistic (ISTAT). Each census area is made up of one or

[6] http://www.r-project.org/

more blocks and each block is composed of contiguous buildings surrounded by mobility infrastructures or natural barriers like streams. The sizes of census areas are different and the median value is 12000 m^2.

Lastly, *points of interest* provided by the two sources (OpenStreetMap through the Overpass API and Milano municipality open data portal), are slightly different in terms of categories: they both have POIs about transports, schools and sport facilities, but in addition OpenStreetMap provides information about shops and amenity places of the city. On the other hand, data coming from Milano municipality are “official”, whereas the OSM dataset, being user generated data, may be less reliable or incomplete.

As Table 1 reveals, besides content heterogeneity, the datasets also differ in terms of spatial resolution and that's why a pre-processing phase was required in order to make them comparable.

Since population, Telecom and land use datasets have a too fine-grained spatial resolution (census areas, square cells of 250 m and even building level resolution) and POIs datasets consist of punctual information, we needed to find a suitable uniform spatial resolution to continue with the data exploration and analysis. We decided to choose two different spatial resolution levels for our analysis.

Firstly we mapped all datasets into the 88 district subdivision (Figure 2) defined by the local municipality (more specifically by the “PGT - Piano di Governo del Territorio”). Since these districts, named “Nuclei di Identità Locale” (*NIL*), can be considered as areas connected by mobility infrastructures/services that share common commercial activities, parks, services and meeting places, this mapping offers both a meaningful and a manageable spatial resolution. This district-level granularity is neither too coarse-grained to invalidate the datasets mapped to their level, nor too fine-grained to make a manual evaluation of analysis results impossible.

However, since we are interested in understanding how the adopted spatial resolution affects the data analysis, we also mapped all the datasets into a more fine-grained resolution, a grid of 3538 square cells of 250 m, which is the original resolution of Telecom data records (Figure 3). Even if this uniform subdivision of the space is, beyond any doubt, less meaningful and harder to interpret, we would like to find out if results are in some way similar and comparable between the two resolutions.

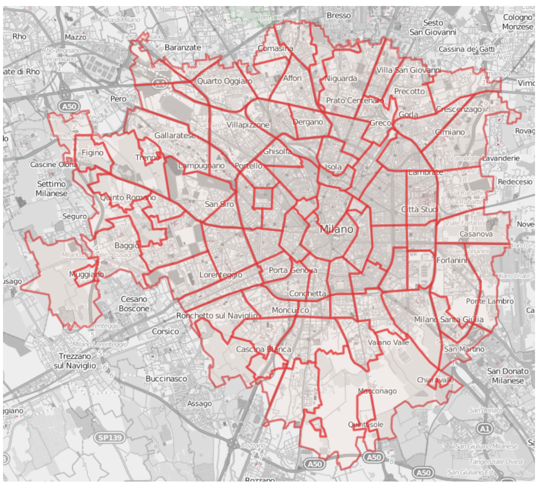


Figure 2: 88 Milano districts – NIL.

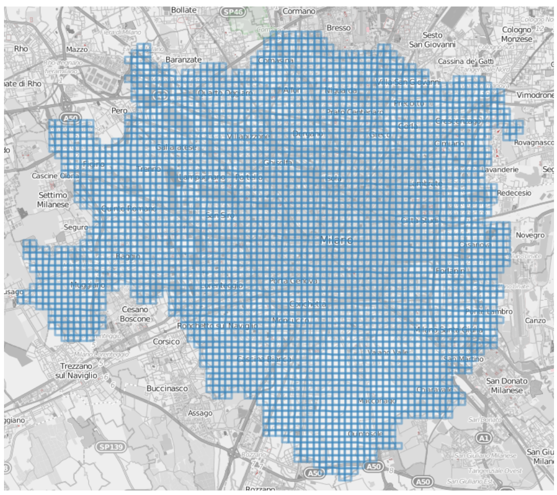


Figure 3: Grid of 3538 squared cells.

Since both the cell grid and the NIL resolution levels are more coarse-grained than the original resolutions of CORINE, demographic and Telecom data (building-level, census area and the grid itself respectively), we mapped the data in the new spatial granularities. This operation was quite easy, since we performed these interpolations using QGIS[7], a free and open source Geographic Information System. We intersected the layers describing the target resolution levels (the cells grid and the NIL subdivision) with the original data shapefiles (census area, building-level resolution and grid). Then we interpolated the original data values consequently: for example if a census area resulted to overlay 60% on a cell and 40% on another adjacent cell, its population value was split and attributed 60% to the first cell and 40% to the second one. It stands to reason that this approach can lead to an approximation, but we can consider it appropriate for our case study.

A different kind of mapping was required for POIs data, as they are data points described by latitude and longitude pairs. In this case, we simply computed POIs density in each district and in each cell.

Once all the datasets are represented at the same spatial resolution levels, another best practice is to perform *data normalization*, which becomes more relevant when we are dealing with districts, as they have different surface areas.

As regard Telecom data, at cell-level we firstly adopted a normalization-in-time methodology on the cells data records, which consists of dividing each value by the average activity of all cells at that given time. When we moved to NIL level we enriched the normalization process by calculating a sort of phone activity density, by dividing the normalized-in-time information by the NIL surfaces in order to take into account the different districts areas, so to have also a normalized-in-space data.

As regards population and POIs, we normalized the data by the area of each district/cell, obtaining the population/POI's density for each NIL and cell.

As we dealt with percentages, normalization process was not required for CORINE data.

To sum up, all the datasets have been adapted to both resolution levels and the experiments illustrated in Section 4 were repeated for both granularities.

4 Data Exploration Experiments with Milano Datasets: Correlation Analysis

Open Geospatial Data provide information about the city in different forms and with different levels of complexity. The simplest example consists in the variation of some quantity over the urban space: if the city environment is for example divided in N neighborhoods, a dataset can be a 1-dimensional structure (a vector of size N), in which each element represents the magnitude of some quantity in each neighborhood (e.g., density of population, average temperature, income indicators).

Whenever several datasets have such a form, or can be summarized by a single variable (interpolating or averaging some quantity), a best practice is to analyze each of those dataset to understand data distribution, to identify outliers and so on. Then, in order to identify possible correspondences between different datasets, we can apply statistical analysis and compute comparison metrics. This kind of correlation analysis is described below.

[7] http://www2.qgis.org/it/site/

f we consider datasets' pairs, i.e. we compare a variable with another variable, we apply *bivariate analysis* to the two datasets to determine if there is any relationship between them. To measure whether and how those two variables simultaneously change together, different correlation indexes can be computed. The most common measure of dependence between two variables is the *Pearson's product-moment correlation coefficient* (Pearson, 1895), defined as the covariance of the two variables divided by the product of their standard deviations:

$$r = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{(N-1)s_x s_y}$$

where $\bar{x}$ and $\bar{y}$ are the sample means and s_x and s_y are the standard deviations of the two variables.

Pearson's *r* is between -1 and 1 and indicates the strength of the linear relationship between the two variables.

We applied bivariate analysis on all the datasets presented in Section 3 and we measured the relationships between them by computing the Pearson's *r* index.

We performed the analysis at both NIL and cell level, with the aim of discovering if some analogies exist between results obtained from datasets with different spatial resolution.

Since we needed 1-dimensional vectors, we pre-processed the two *n*-dimensional datasets, CORINE and Telecom, as follows. Regarding CORINE land use, we decided to separately use data about residential and agricultural areas; in this way, we got two vectors, one representing residential use and the other one representing agricultural use of Milano districts and cells in terms of percentages.

On the other hand, for Telecom data we decided to apply *Principal Component Analysis* (Jolliffe, 2002), a statistical procedure that converts a set of possibly correlated observation into a set of values of linear uncorrelated variables (the principal components), through an orthogonal transformation. PCA is used for dimensionality reduction, as just a few principal components can be able to explain the overall variance of the problem. From our analysis, the first principal component of the telecom dataset explained about the 90% of our dataset's variance, we considered it a good approximation for representing the whole dataset as a 1-dimensional vector.

Figure 4 lists the results of all the pairwise comparisons performed at NIL level. The plots on the diagonal show data distributions which turn to be left-skewed for all the datasets.

Analysing Pearson's *r* coefficients, we found that the highest positive correlations exist between phone activity and residential land use (0.83) and between phone activity and POIs (0.84 and 0.85). Those results suggest that a correlation between those datasets can actually exist. The respective plots in the lower-diagonal parts of the scatter-plot-matrix also show that data fits quasi-linear models.

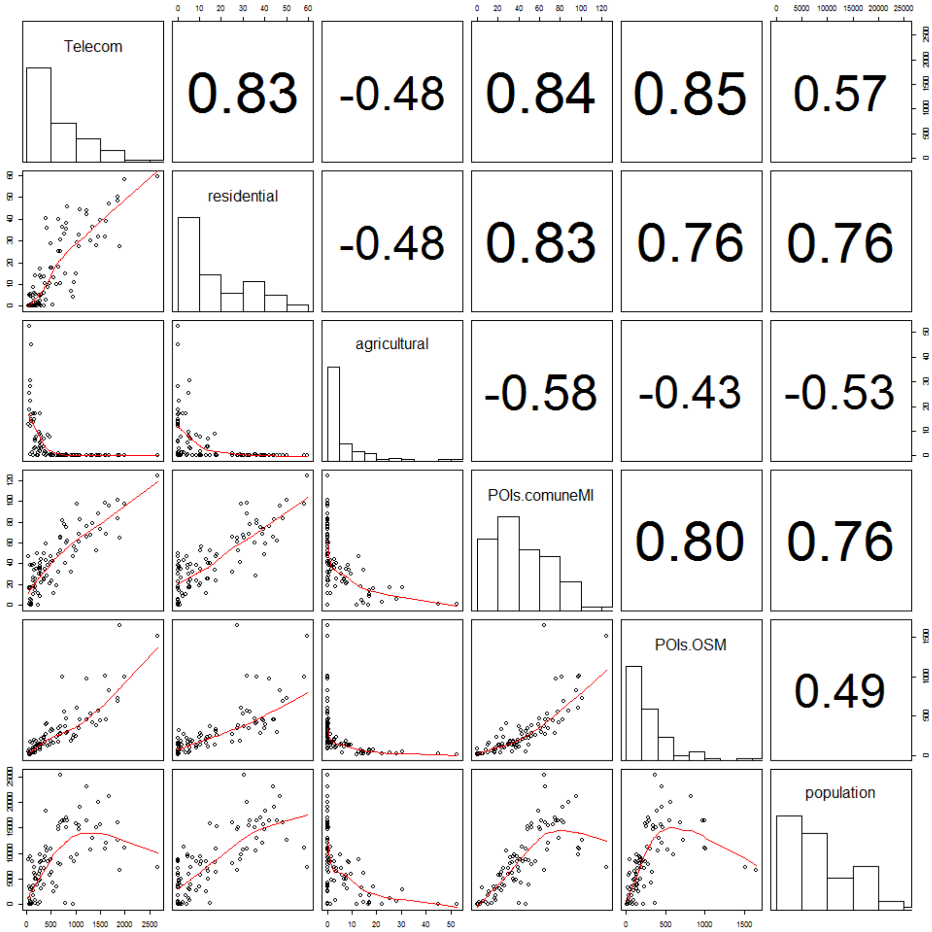


Figure 4: Scatter plot matrix with correlation between datasets at NIL space granularity level: Pearson's r correlation coefficients are above the diagonal, dataset distribution is on the diagonal and the plots of correlation are below the diagonal.

On the other hand, as can be expected, we observed negative correlation between the agricultural land use and all the other datasets (Figure 4). We interpret this result as a sign that human actions (like those recorded by mobile activity or by presence of POIs) are inversely related to agricultural areas.
Then we compared these indexes with the results obtained using the cell spatial resolution level. The Pearson's *r* (Figure 5) at cell-level are all lower than the ones at NIL-level. The higher Pearson's *r* values are again between Telecom data and residential areas (0.66) and between Telecom and POIs from OpenStreetMap (0.68).
We can notice that the correlation values at cell-level involving the population are the ones closest to the NIL-level results and the correlation between population and residential areas is even the same (0.76). These results suggest us that the choice of the suitable resolution level can have a significant impact on the correlation results.
The 250 m squared cell maybe is so small that there cannot be enough information to characterize it properly or the data available are not enough punctual and precise.

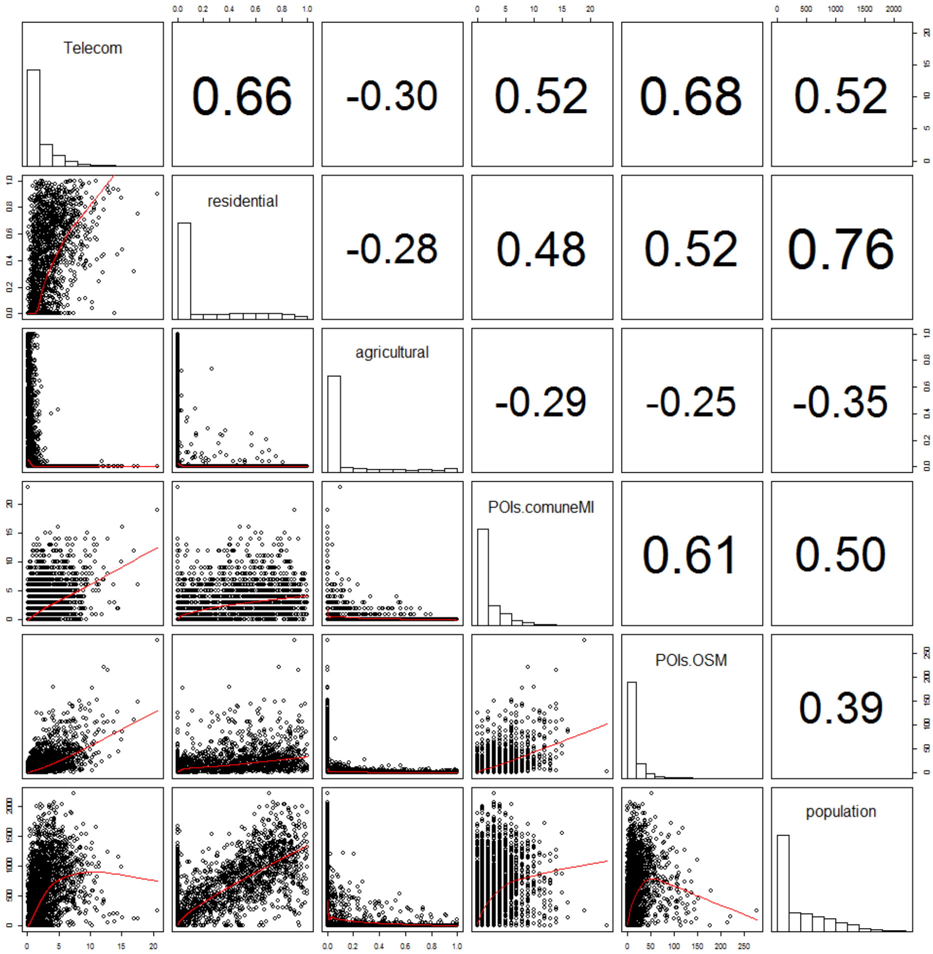


Figure 5: Scatter plot matrix with correlation between datasets at cell spatial resolution level: Pearson's correlation coefficients are above the diagonal, dataset distribution is on the diagonal and the plots of correlation are above.

Since we found that both at NIL and at cell level there was a good correlation between the average phone activity and the population density (0.57 and 0.52 respectively), we deepened this analysis by investigating the change in correlation values at different times of the day.

In particular we wondered if the correlation could change during the day accordingly to the everyday human behaviour pattern (get up, go to work, come back home in the evening).

For this reason we extracted the incoming and outcoming call activities at six different hours (at 4 am, 8 am, 12 am, 4 pm, 8 pm and 12 pm) and we computed a 1-dimensional vector for the average mobile activity at each specific hour.

We did these experiments using the datasets at both resolution levels and taking separately weekday and weekend activities. Figure 6 plots the results obtained.

At first sight, we can see that weekday and weekend profiles are different: while the former has peaks of correlation in the morning (8 am) and in the evening (after 8 pm), the latter reaches the maximum values at mealtimes (at 12 am and 8 pm).

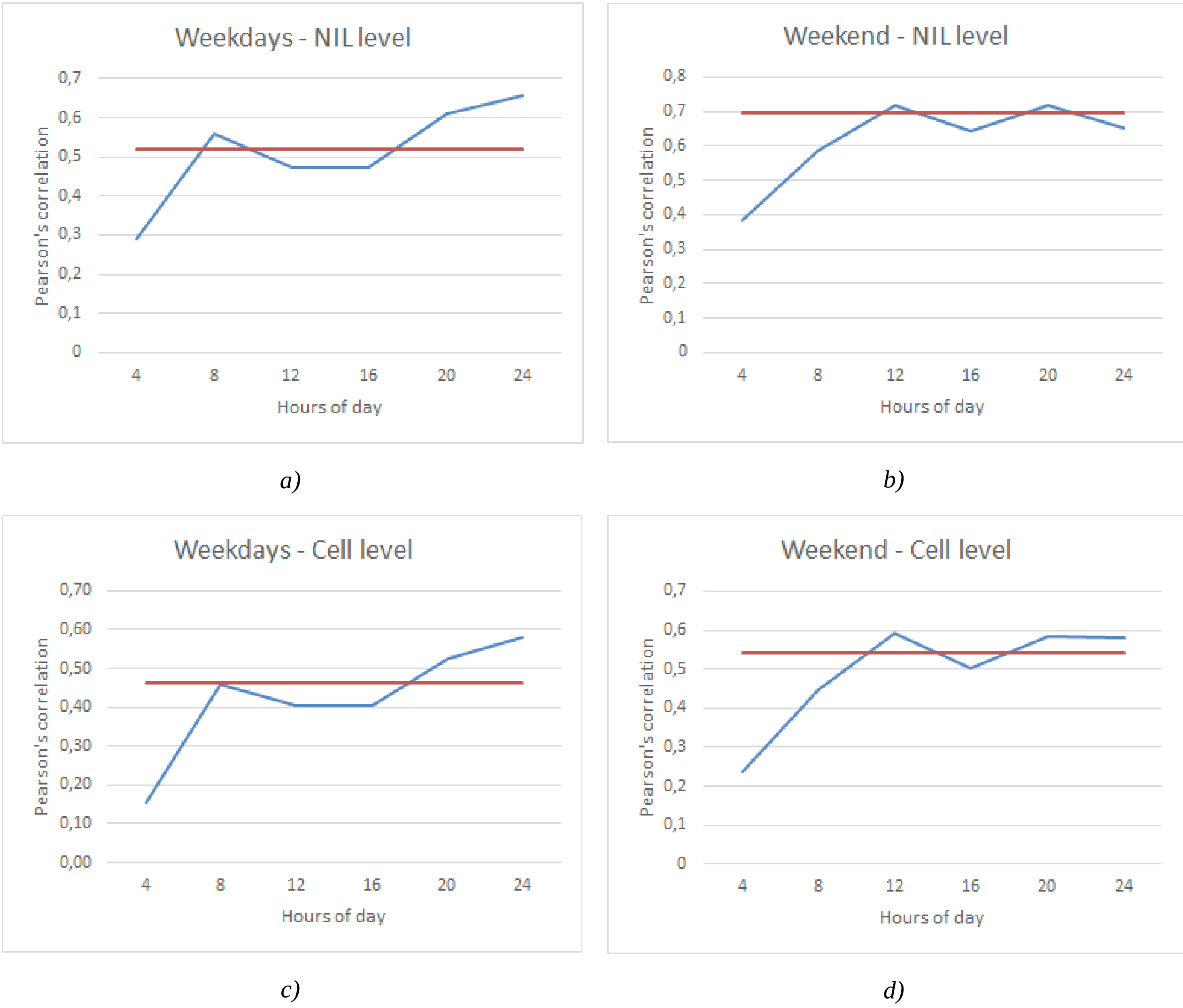


Figure 6: Plot of Pearson's r coefficient obtained by comparing the population density with the weekend and weekdays in-out calls at different hours of the day (4am-8am-12am-4pm-8pm-12pm) – blue line – and with the average in-out calls activity – red line. Plots (a) and (b) refer to data at NIL-level whereas plots (c) and (d) at cell-level.

The weekday profile proves our hypothesis that phone activity is related to the everyday human behaviour pattern: the correlation between phone activity and population density is lower than the average during the day when people are at work and become higher when people come back home in the evening.
In addition, in the weekend the average correlation is higher than in weekdays. This trend confirms that the amount of phone activity is correlated with the actual presence of people at home.
The shape of the profiles at the two different granularity level are quite the same and, as already seen above, correlation at cell level is overall lower than the one at NIL-level.

5 Conclusion

The proliferation of Open Geospatial Data related to the urban space calls for smarter solutions to process and mine such data. In this paper, we presented our best practices for data exploration process and in particular we illustrated our data analytics experiments on a set of heterogeneous datasets related to the city of Milano. Our purpose is to identify dependence between diverse sources, to understand if distinct datasets provide similar representations of the city. We experimentally evaluated the correlations between datasets at different levels by employing different spatial resolution (coarse-grained district level vs. fine-grained cell level).

Our empirical results show that correlations between different sources exist, even if their strength depends on spatial resolution. As a consequence, investigations on urban data must take this evidence into consideration and tailor experiments at the most suitable resolution level to obtain meaningful results.

It is worth noting that, given the available information, we decided to perform such experiments even if the sources do not refer to the same time-frames (2011-2014). While this means that our results risk to be less significant, we believe that the general trends showed by our analysis could be confirmed by datasets referring exactly to the same period.

This further refinement analysis is on our agenda, as soon as we get urban datasets homogeneous in terms of their temporal dimension.

Starting from the promising results in terms of correlation between heterogeneous sources, we aim to extend our investigation toward a predictive approach, wondering if it would be possible to learn the characteristic of a dataset from the information given by multiple and diverse data sources. In particular we are wondering if it would be possible to use one or more “cheap” datasets – like often are open data – as proxy for more “expensive” data sources. In other words, would it be possible to (semi)automatically generate or revise an outdated dataset, which otherwise would require a costly human work, on the basis of the content of other up-to-date information sources?

To reach this goal we need to move from our initial exploration analysis to statistical learning experiments (Minasny, 2009) and to supervised machine learning (Kotsiantis, 2007) techniques, making use of both regression and classification algorithms from the simplest linear models to the more complex and powerful Neural Network and Random Forest. Given the intrinsic spatial nature of our datasets, spatial autocorrelation and dependencies between data could also exist, thus the Moran’s I index is a very important indicator. In this case the adoption of Spatial Autoregressive models could provide benefits to the analyses.

References

- Goodchild, M.F. (2007). Citizens as Sensors: the World of Volunteered Geography.*Geojournal*. doi:10.1007/s10708-007-9111-y
- Gotway, C. A., & Oung, L. J. (2002). Combining Incompatible Spatial Data. *Journal of The American Statistical Association*. Doi:10.1198/016214502760047140
- Jolliffe, I. T. (2002). P*rincipal component analysis*. New York: Springer.
- Kitching, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*, 1-14.
- Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of

Classification Techniques. *Informatica*, 31, 249-268.
- Minasny, B. (2009). Trevor Hastie, Robert Tishirani and Jerome Friedman, The Elements of Statistical Learning (Second Edition) (2009) Springer Series in Statistics, ISBN 0172-7397, 745 pp. *Geoderma*. doi:10.1016/j.geoderma.2009.08.001
- Mitsa, T. (2010). Temporal data mining. *Chapman amp. Hall/CRC.*
- Pearson, K. (1895). Note on Regression and Inheritance in the Case of Two Parents. *Proceedings of The Royal Society of London*. doi:10.1098/rspl.1895.0041
- Reades, J., Calabrese, F., Sevtsuk, A., & Ratti, C. (2007). Cellular Census: Explorations in Urban Data Collection. *IEEE Pervasive Computing*. doi:10.1109/MPRV.2007.53

Open Data integration in context of Smart-Communities: a company's perspective

Roberto Metelli[1], Consuelo Marini[1], Francesco Pirotti[2]

[1] Linea Com S.r.l. - Linea Group Holding S.p.A Viale Trento e Trieste, 38 - Cremona (CR), Italy

[2] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, Viale dell'Università 16, Legnaro (PD), Italy

Abstract

Comprehension of what is and how to use Open Data has many levels. In the following article we illustrate a work in progress "entry-level" situation where a publicly owned company wants to create added value using Open Data in the context of a Smart-Community vision of the areas it is involved in.

Keywords

Energy efficiency, Multiutilities, Smart City, Smart Community, Open Data

1 Introduction

The future in management of public elements is tightly coupled with data acquired from sensors which are becoming more common. This is due to lower costs of the sensors themselves, thanks to advances in technology, and to increased availability of distributed sensors, ranging from smart-phones (crowd-sourcing) to smart lighting, smart grid, sensors for any measurable quantity which is of interest. Together this paradigm is called "smart" - applied to cities and communities. Its focus is to integrate (link) data as much as is needed and possible, linking sources (sensors) with data bases, grids and structures in general which allow management and analysis of the data, adding value and leveraging better usage. The way that data is managed and integrated with data from other sources introduces the importance of Open Data (OD) in a bigger picture of integration of projects related to smart cities/communities.

Linea Com Srl, is a company which is part of the LGH SpA Group which has been operating in the ICT field for more than 15 years. It supports a number of municipalities which are partners of group of five companies which constitute LGH (Cogeme Rovato – Astem Lodi – Asm Pavia – Aem Cremona – SCS/SCRP Crema). Together these companies cover a large part of the territory of Regione Lombardia, municipalities which are spatially close and / or connected to the territory. Linea Com's mission is to support public administration (PA) in more than 880 municipalities in the provinces of Bergamo, Brescia, Cremona, Lodi, Mantova e Pavia covering more than 3.5 million residents. Since 2007 a strong

push has been given to converting to Open Source principles, in particular in the group dealing with spatial data services. It successfully coordinated the acquisition and structuring of the topographic geodatabase for more than 200 municipalities, providing a free webgis service to the public were users can access all layers from the topographic database and other spatial layers, view them and download them fully or partially in the main open (and not so open) formats e.g. Shapefile, DXF, DGN, GeoJSON, GeoUML) as well as access them as OGC WMS services. Internally this has been done also for managing the underground infrastructure (GAS) for more than 50 municipalities. External OD services (AGEA WMS services, Provincia di Brescia WMS, eGeos WMS etc...) have been added to leverage information availability, thus providing a central hub for spatial data.
The above introduction of Linea Com as part of a larger group closely related to the territory and dealing with multiple aspects of municipalities, it is the ideal promoter of the use of data at its best potential, related also towards "smarter" cities and communities. Some of the latest examples are the following: (i) the past funding for improvement of cartography converting it to INSPIRE-compatible formats, the topographic geodatabase – now all data are in a Postgis structure and available as WMS service; (ii) the Virgo European Project which wants to create a registry of infrastructures which populate our territory below-on-above the ground; in this case OGC services will be used as well as proprietary software, with integrated solutions; (iii) Attiv@bili is another potential future project (if funded) were OD will be analyzed using BigData analytics; (iv) the webgis service provided by Linea Com has a public open profile providing several tools for visualization but also for free downloading of data in Open formats (GML, GeoJSON – see figure 1). Other examples of future data integration which Linea Com wants to achieve is for the service of smarter integration in refuse management, renewable resources, water and sewage planning and management. Coordination and collaboration between stakeholders in the LGH group allows to promote synergies and solutions to improve well-being and societal challenges such as security, services and energy resources.

2 Stakeholders and promoters related to Open Data

The innovation towards "smarter" communities must have a solid base, which is made of two pillars: network of infrastructures and sharing data/knowledge via Open Data (Figure 2). Related technologies must reach out spatially as much as possible, to be able to promote a new culture related to sharing data and knowledge. Not only technologies must reach far and deep to households of final users, but another important aspect is that of developing sensibility towards the matter. Education in this sense must be promoted at initial levels of education, as well as towards "elderly" users which need to be updated to new technologies to avoid being isolated in a fast-moving society. To achieve what is presented, a first step is to create a network of sensors to read data, and a second step is to provide intuitive interaction between the data and the users at different levels of capability depending on their final goals. Connectivity with 4G technology coupled with new street-lighting can be a very interesting example. Light-poles in cities can become a hub for monitoring the area for security (cameras) and provide "infotainment" – basically an interaction with access to digital information, and also provide data to improve mobility thus lowering the carbon footprint. Another fundamental aspect is the spreading of open-source technologies in public administration offices. This has often been seen as an "obvious" and "easy" task, but it is not so straight-forward. Changing a work-flow of several offices and numerous users has always been seen as a risk, especially some years ago when open-source software was not so compatible (even now trying to open this odf with Word would not keep the layout). Open-source users usually have a higher level of proficiency with computers, and provide solutions that seem simple to them, but that might sound like jabber jargon to others. This is changing now that many open source projects have become much more user friendly and interoperability solutions are becoming more common. This will surely push for an easier transition from expensive licensed software to open-source software and especially formats, which will be easily shared, and thus following the principle of re-utilization.

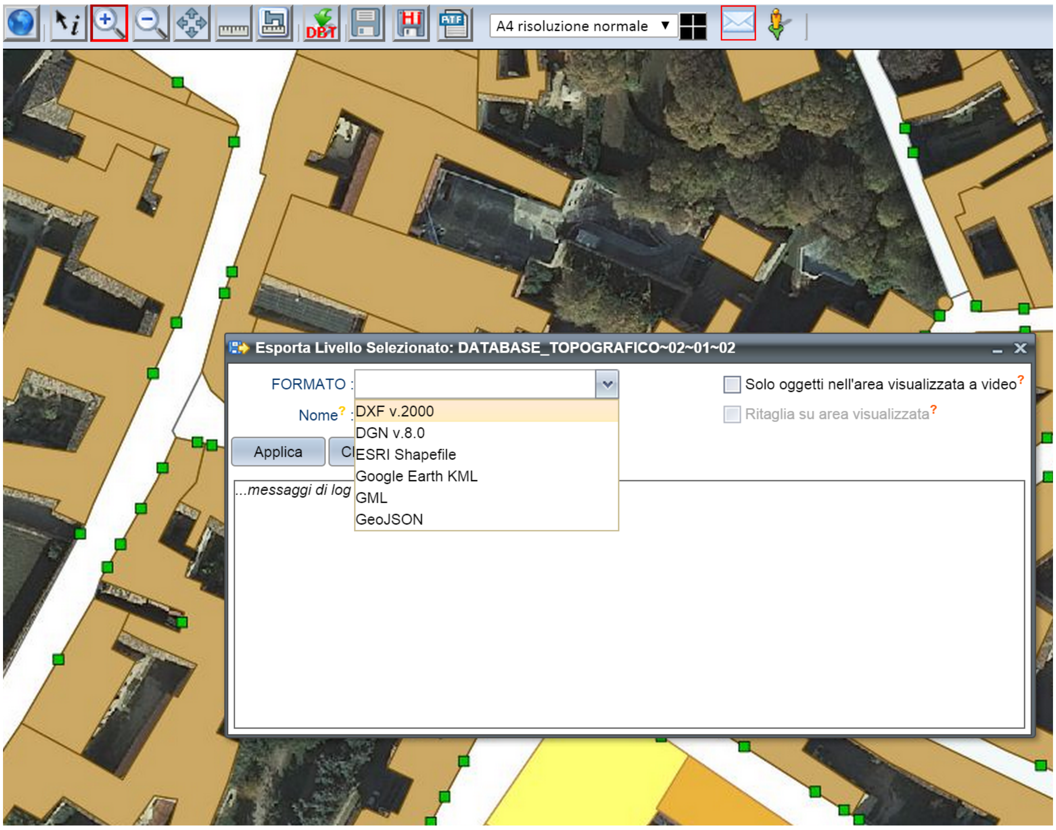


Figure 1: Example of topographic geodatabase with possibility to download data in different spatially enabled formats.

Spatial and non-spatial Open Data is fundamental with what has been said above. Of course some type of data cannot be completely set in the open domain, as privacy and security are important issues, nevertheless there is an enormous amount of data which is underused, like a gold mine which has not been tapped yet. Putting such data in the public domain with Open Data licenses would thrust a new era of sharing and collaboration, providing new utilization for the data and improving collaborative networks between users in the same public administration, between public administration, and externally to public administration. Linea Com wants to create value from its data by enhancing its accessibility and thus its fruition. One way is the education towards using and providing Open Data towards all players involved. This can be done in many domains, such as the public bodies themselves, administration, and single citizens which are encouraged to use and be aware of data and appreciate the multiple uses that derive from it.

2.1 Open Data principle in a Smart City context

It is often the case that when implementing a project dedicated to integration of elements focused towards a Smart City context, the data which is available is not structured, standardized or either in a digitally usable format. This lower the overall quality and requires a significant amount of processing to modify the data and render it in a more usable format. This effort towards standardization is carried out at many levels and in many cases. Open Data can be a starting point for forcing stakeholders to relate to their data in a way that "cares" also for uses other than theirs. Thinking that "my data will also be used by others, using other software" will, probably, create a positive mechanism where users will try to structure their data and save it in interoperable formats. This will facilitate sharing and usability and thus add value to the data itself.

2.1.1 Spatial data

The Italian Agency for Digital Italy (Agenzia per l'Italia Digitale - AGID) has stated that "spatial data constitute a knowledge base for all policies related to land-use management". This might seem obvious for users of spatial data and city planners, but it is often not so for policy makers which do not have a clear idea of the process behind planning phases for transportation, security, risk assessment, environmental hazards etc. It is therefore a "best practice" when planning "smarter" cities to give utmost importance to spatial data. Also in this case, like mentioned in the previous section, Open Data can leverage all kinds of positive mechanisms which allow a private citizen to get information and give information for its peers.

This is were Linea Com would like to make a difference: providing services pushing on structuring and opening data, spatial and not. The benefits from this type of effort are more than documented in literature and in best-practice scenarios. It is now the moment to jump from research and development to actual implementation in districts and municipalities, educating ourselves and others to sharing and opening, knowing that "you get what you give" is also true for data.

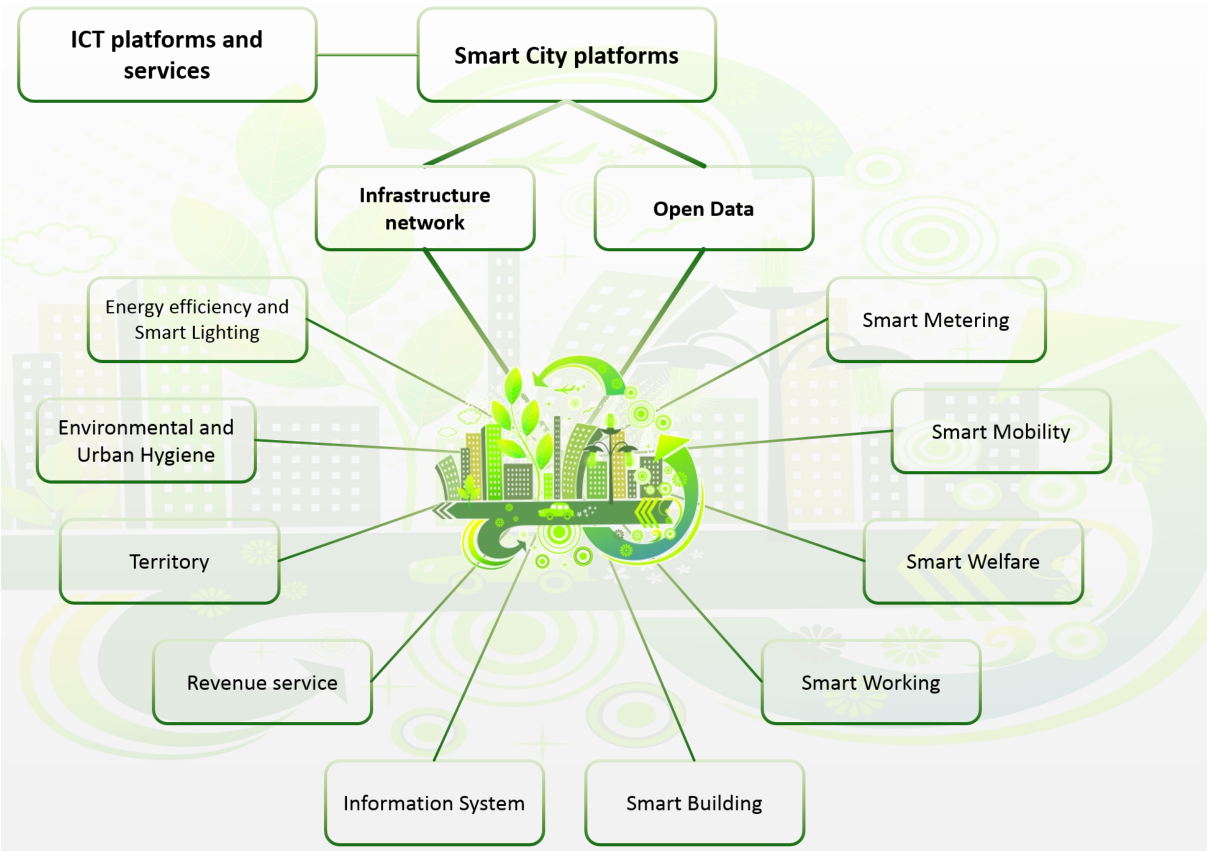


Figure 2: Linea Com's concept of integration of Smart Community with the two pillars, Open Data and Infrastructures.

References

✔ Dati territoriali. (2015). *Agenzia per l'Italia Digitale.* Retrieved February 2, 2015, from http://www.agid.gov.it/agenda-digitale/open-data/dati-territoriali

IMAGO.gIS: an open solution to improve the fruition of public administration data

Oscar Campolmi[1], Gianluca Giori[1], Michele Beretta[1]

[1] Imagogeo srl, via Cavour 2, 22074 Lomazzo, Como, Italy (www.imagogeo.it)

Abstract

An huge amount of open data like Topographical Database, cadastral data, environmental restrictions are produced by public administrations to follow national and regional directives but often they aren't directly reacheable for citizens and municipal technicians. Using GFOSS (QGIS, PostgreSQL/Postgis, QGIS server and Web Client) our project IMAGO.gIS aims to join all this data and make them usable for technicians and private users.

Keywords

Cadastre, public administration, GeoDB, WebGIS, QGIS, QGIS Web Client

1 Introduction

Municipalities of Lombardy Region produce and update every year a large amount of Geospatial Open Data about their territories to attend some regional directives (Regione Lombardia, 2005). Instead this data represents a considerable cost for the city budget they often aren't easily avaible for private workers, citizens and municipal employee too.
The aim of our solution is the complete, simply and quick fruibility of this data. Our developing philosophy embraces the Open Source way that allows to exploit them, produced by public administration, by public administration themselves, community and private worker.
This philosophy allows to prevent additional expenses by the public administration that already spent a lot of money to realize and update this geospatial data. Moreover avoids expenses like property software licenses.

2 Local administration geospatial data

Geospatial Open Data produced or available from local administrations are divided into three levels: the regional topograpichal database also called *DataBase Topografico Regionale* (DBTR from now), the ruling plan for territory and related environmental restrictions also called *Piano di Governo del Territorio* (PGT) and cadastral map and data.

2.1 DBTR

DBTR is defined by various regional directives (Regione Lombardia, 2008) accordant with the "*Protocollo d'intesa stato-regioni ed enti comunali per la realizzazione dei sistemi informativi cartografici di interesse generale*" (Intesa

GIS, 2006). Lombardy Region defines the technical specifications for the DBTR and his styling rules; following this specifications the local administrations realize and update the related territory DBTR.
DBTR stands for the old *Carte Tecniche Regionali* (Lombardia, 1979) that represented all the natural and antrophic feature over the entire territory.
The database features are divided into *CLASSI* (e.g. "Unità volumetrica", "Area di circolazione veicolare", etc..) which are divided themselves into *TEMI* and *STRATI*. For each of them is defined an univocal style.
Lombardy Region DBTR can be found at Lombardy's Geoportal (http://www.cartografia.regione.lombardia.it/) as Geospatial Open Data (Regione Lombardia, 2008) following the INSPIRE directive (European Parliament and Council, 2007).
Due their complexity this data aren't easily manageable by public administration and general users, consequently they aren't employed despite the production expenses.

2.2 PGT and environmental restrictions

The called *Piano di Governo del Territorio* (PGT) and the enviromental restrictions (idrogeological, wells' buffer, archaeological, etc...) determine the intended use, restrictions and appropiates technincal specifications for the entire municipal territory (e.g.: "Nucleo di antica formazione", "Tessuto produttivo", "Tessuto turistico", etc...). Every municipality creates compulsorily it within a geospatial version and relative custom style.
PGT and environmental restrictions are sometimes downloadable in shapefile™ format by the PgtWeb portal (https://www.multiplan.servizirl.it/pgtweb/pub/pgtweb.jsp), although data are partial and withot style yet.
Like DBTR data, they are produced but aren't often used for lack of software or knowledge. For their complexity is difficult to use them in the digital or papery format without the correct right tools, with risk to have the wrong information from them.

2.3 Cadastre

Cadastre is managed by the Ministry of Finance – Agency for the territory. They produce a cartographic version representing the results of measurement operations of the territory and, therefore, able to provide the information to allow both the geo-referenced real property (parcels and buildings) and technical parameters relevant for the definition of their potential production of income (eg. land area of the parcels, the parcels manufactured location compared to urban). It defines each property and land registered and connected to a physical owner.
Hard copy is available for citizens by the responsible local authorities after a request (often for a fee). Anyway also the administrations often aren't able to manage digital information and they work only with hard copy within the usual problems of overlap them with the rest of cartography.

3 IMAGO.gIS

The aim of our developed products is to give easy and ready to use (configurated and stylized data) tools for the municipal technicians and public administration in general, so they can powerfully exploit and develop the territorial informations. At the same time we give to the administration a web tool to disclosure this data to citizens and private workers.
The solution proposed is a client-server structure which consists of a a GIS project based on a geographic database that allows the complete management of spatial data (DBTR, PGT, environmental restrictions and cadastre) properly configured in accordance with regional and individual municipalities standards.
For the use within the municipal offices, the layers are inserted into a remote spatial database (PostgreSQL with spatial extension PostGIS), which is connected to several client-side terminals of the municipal offices. On single terminal resides the QGIS project that retrieves and displays data from the remote DB (Figure 1). This avoids unnecessary duplication of an huge amount of data, allowing the fruition to a potentially infinite number of users. This configuration also allows a significantly faster update of the data and especially ensures a data uniformity on all terminals.

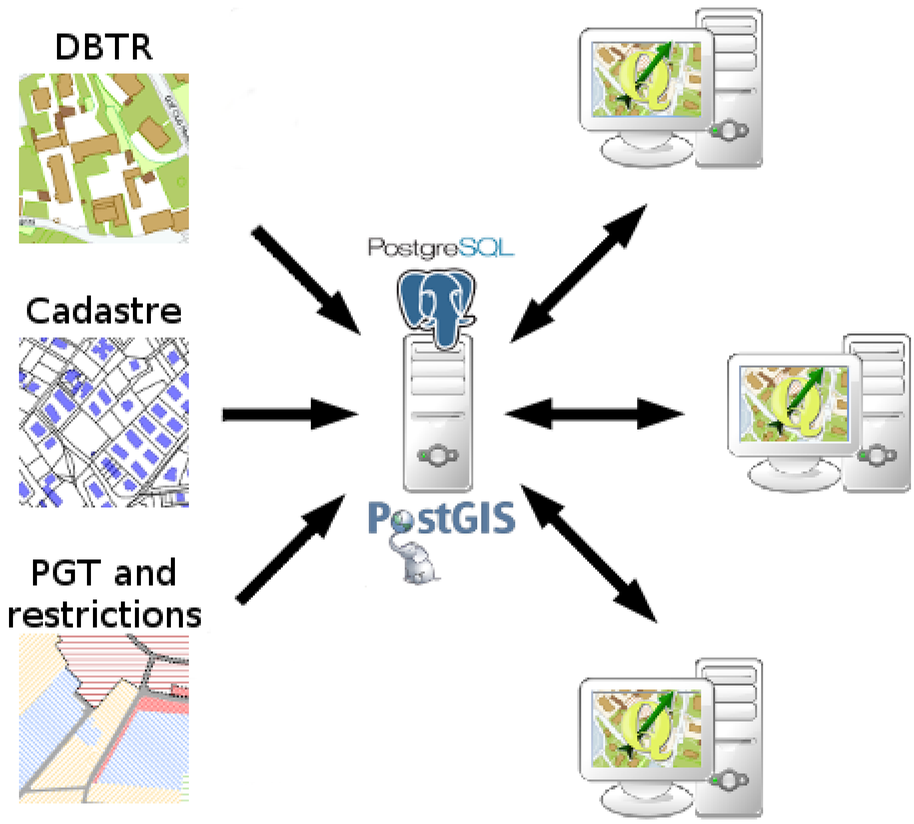


Figure 1 - Structure of our desktop solution for local administration's offices

Through the creation of specific plugins made with Python script, with graphical user interfaces developed with PyQt4 language, functions are implemented to simplify and speed up certain operations of municipal technicians, related to daily work and public relationship.

3.1 Developed plugins

We developed plugin using a "one click - one function" philosophy for quick and easy basic GIS operations and for the management of data concerning the cadastre, intertwined with other layers information, especially PGT and environmental restrictions.

3.1.1 DBTR Functions

This series of plug-in allows a number of basic GIS operations, thougth for less experienced users and focused on operations and analysis typical of a technical office.
Among these, the quick feature identify allows the immediate selection and zoom of the map elements, in addition to the attributes display. Moreover the user could create some geolocated notes, attaching them files of any type and sharing them with all other clients. This function is possible thanks to the client-server structure mentioned above and allows a greater efficiency of communication between the various administration offices and increasing the territorial system information content. Creation of buffer zones around one or more features for planning purposes with the possibility to show wich other features of other layers intersect the latter (eg, if the municipality have to widen a road, they could see directly the impact that this will have and wich lands are interested). Fast search for streets and house numbers. Download in CAD format (DXF) of portions of the map.

3.1.2 CAT Functions

This functions allow some operation over the cadastre layers: buildings layer and parcels layer.
One function allows to search building and parcels by cadastral number and zoom it. The second feature is one of the most important and useful that we had developed for municipal technicians: it permits to generate and print a certificate called "*Certificato di destinazione urbanistica*" (CDU) that specific which urban areas belong a given parcel (buildable, restrictions, etc..). For one or more selected parcels the python script founds the corresponding PGT and environmental restrictions; then another script prints an editable document within the urban rules and limits of that PGT and environmental restrictions; the document is obviusly personalized for each municipality (with logo, name of the technician, template, etc...). This function is useful because without it the municipal technicians need to overlap the papery maps of cadastre with PGT, environmental restrictions, etc. (20-25 different layers) to search the correspondings areas, and then write down the urban rules. This operation may takes some hours and suffer of subjectivity errors and the risk of forgetting some territorial informations. Using this plug-in they need only few minutes to insert parcels numbers or select them directly on map and to check the certificate validity. Also are automatically inserted the standard urban rules (otherwise often not present) and a map extract into the document.

Another fuction allows to identify the buildable parcels in a selected area and shows some informations like extents and owners. For example to check the total municipality buildable territory for planning and taxation purpose.
The last feature searchs all the parcels under a specificated environmental restriction and shows some informations like extents and owners too. For example the administration could verify all the citizens affected by idrogeological risk and improve the emergency plan.

3.1.2 CAT+ Functions

This group of functions allows some advanced operations over the cadastre layers. Advanced search by number and/or owners of buildings and parcels; the resultant interface shows all the information about the parcels and buildings, even the singles cadastral plan and the planimetric composition which shows all cadastral subordinates. This first group of functions allows greater control of the properties over the territory for technical purposes (planning, public works, etc.) and especially for taxing audits.
In order to have an efficient tax control, the administration tax section is already able to identify any anomalies in the citizens declarations overlapping DBT and cadastral buildings. Furthermore, additional functions allow to deepen the problem by identifying all those buildings located in cadastral map but which do not have a correspondece in the cadastral database (and maybe illegal).

3.2 WebGIS

All the layers of desktop GIS are published in a WebGIS, so every citizen or interested private worker could consult it by home or own office. We use the same remote PostgreSQL/Postgis database to store all the geospatial data, Apache as web-server and QGIS Server and QGIS Web Client to publish data. That client makes use of QGIS specific WMS extensions (e.g. highlighting, printing, metadata, etc.), reads the configuration from the WMS GetCapabilities command and builds the layer tree accordingly.
This structure allows us to configure one QGIS project usable both from desktop by technicians and by generic users through the web. In addition a single configured web client is used to publish all WebGIS of all the municipalities (Figure 2).

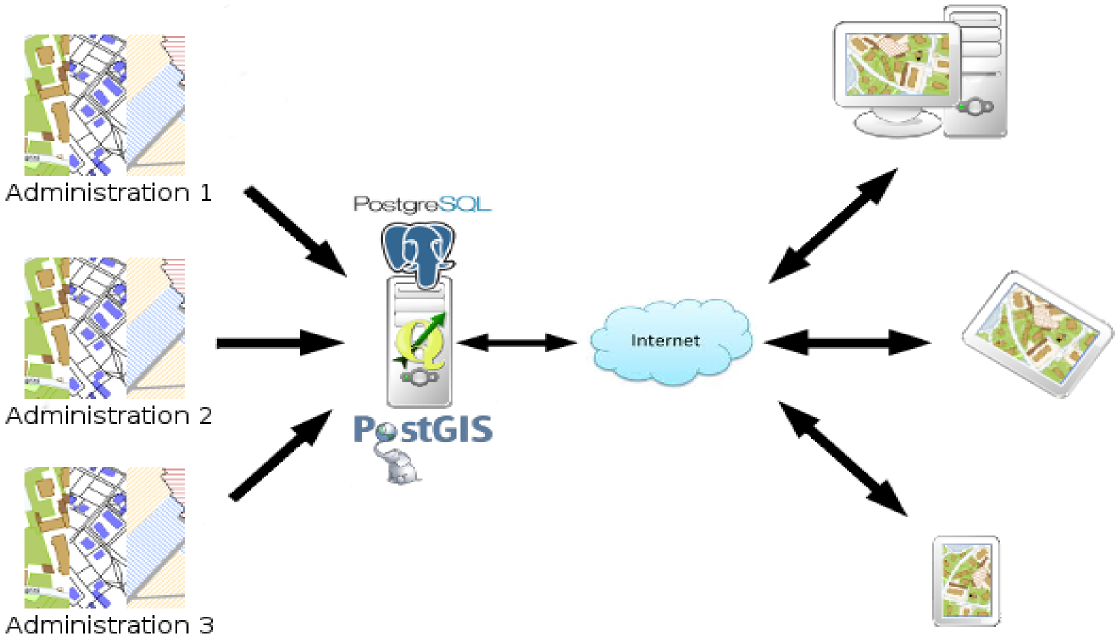


Figure 2 - Structure of our WebGIS service for multiples local administrations.

Advanced function can be used only by experts technicians and can't be published on a WebGIS, however QGIS Web Client support some useful functions like measuring areas, print parts of a map, etc.
We have also developed and personalized some additional functions that allows to search and show cadastral features (using the WSGI search scripts that QGIS Web Client gives) and a Geolocation feature (using OpenLayers library) that technicians can use for their in situ work (Figure 3).
With this tool, the public administrations offer a service of great interest for the public and private workers. Furthermore, in accordance with the law that require the publication on the institutional website of the PGT data, the WebGIS allows a dynamic and completely customizable reading of that information. This aspect is completely different from today standards because public administrations very often publish PGT data almost exclusively with large PDF files hardly readable due to overlapping layers and styles, presence of different files for different information, long time of download and difficult consultation (pan and zoom).

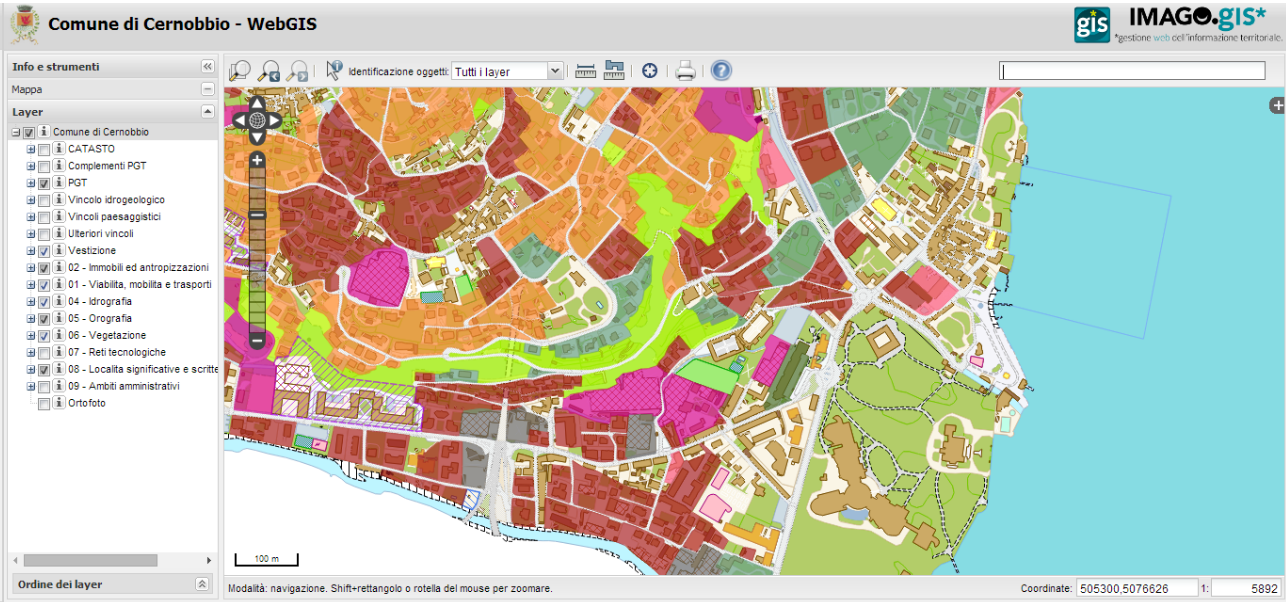


Figure 3 - A WebGIS example, published with QGIS Server and Web Client.

4 Conclusions

The advantages of our solution IMAGO.gIS are several and include many aspects of the public administration and the relationship with its citizens.

We allow the exploitation of data whose production has had an important cost otherwise this data aren't used due to lack of skills and software.

Through the use of Open Source Software, licensing expenses are zeroed and the investment of the municipality is focused on services: complete configuration, customization and staff training. The efficiency of institutions is enhanced with our custom land management features, in the interoperability between offices and in the relationship with users.

Our solution finally allows to create a bridge between the administration and citizenship due to the publication of the data through a WebGIS.

Publishing them through the web, potentially Open Data became really Open and the citizens seem to appreciate this service as we can see through the increasing visits.

References

- ✔ European Parliament and Council (2007). Directive 2007/2/EC of the european parliament and of the council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Retrived from http://inspire.ec.europa.eu.
- ✔ Intesa GIS (2006), Intesa stato regioni enti – locali sistemi informativi territoriali comitato tecnico di coordinamento, Specifiche per la realizzazione dei data base topografici di interesse generale. Retrived from http://www.centrointerregionale-gis.it.

- Regione Lombardia (1979). Legge regionale n. 29 del 04-06-1979, Norme per la realizzazione di un sistema di informazioni territoriali e della cartografia regionale. Retrived from http://www.regione.lombardia.it.
- Regione Lombardia (2005). Legge regionale n. 12 del 11 marzo 2005, Legge per il Governo del Territorio. Retrived from http://www.regione.lombardia.it.
- Regione Lombardia (2008). D.G.R. n. 8/6650 del 20 febbraio 2008, Specifiche Tecniche aerofotogrammetriche per la realizzazione del Data base topografico alle scale 1:1.000 e 1:2.000. Retrived from http://www.regione.lombardia.it.
- Regione Lombardia (2008). D.G.R. n. 8/6650 del 20 febbraio 2008, Specifiche di contenuto e schema fisico di consegna del Data base topografico. Retrived from http://www.regione.lombardia.it.
- Regione Lombardia (2008). D.G.R. n. 8/6650 del 20 febbraio 2008, Specifiche tecniche di rappresentazione dei Db topografici alle varie scale. Retrived from http://www.regione.lombardia.it.

Roadmap to open geodata in Romania

Codrina M. Ilie[1], Nadia H. Panchaud[2]

[1] UTCB, Groundwater Engineering Research Center

[2] ETH Zurich, Institute of Cartography and Geoinformation

Abstract

In Europe as well as worldwide, the open data initiative is on an ascending trend. Even though globally, the support of civil society, private sector, legislators is considerable, the built frameworks must be adapted within each country context considering a multitude of parameters, that vary from economic and social awareness to technical and legal implementations.

In Romania, the research is conducted through the collaborative research project GEOIDEA.RO. The project baseline is to analyse the Romanian environment and draw guidelines for a sustainable open geodata initiative, given the Romanian particularities.

The paper concentrates on presenting the results of the complementary research directions of the project and further focuses on open geodata implementation in Romania, as a country member of the European Union that signed the Open Government Partnership and that has a growing community of open geodata users, supporters and promoters.

Keywords

Open geodata, Romania, GEOIDEA.RO, geoportal, cartography, roadmap

1 Introduction

Political leaders, civil society voices, economists, social analysts and increasingly, business innovators, demand, support and even promote open data with respect to their own qualifications and tasks. Nevertheless, the paradigm shift is not easy to accomplish and constant and thorough analysis is needed to re-confirm the advantages of the open data movement.

International initiatives, such as the Open Government Partnership (OGP), and regional initiatives, such as the European Directive of the Public Sector Information (PSI), have strongly pushed forward the release of public data over the internet under a free license. However, given the diversity of each country's economical, technical, civic, cultural environment, these structures must be adapted within each country context.

In Romania, the analysis of particularities for open data implementation is conducted through the collaborative research project GEOIDEA.RO. The main objective of the project is to improve the scientific basis for open geodata model adoption in Romania.

In order to accomplish the specified task, the team of researchers has developed different, but complementary research directions: on one hand,

reviewing and analysing the international strategies and legislation regarding open data, analysing the impact of adopting open geodata policies in Romania and on the other hand developing a pilot web geoportal to host relevant samples of geodata considering research issues such as data harmonisation, user friendly navigation and data quality.
As seen thorough our research, advancing the open geodata movement in Romania requires a combined approach, not only identifying the interested parties and developing context for discussions, but also setting up knowledge transfer and hands-on activities, e.g. hackathons. The paper further presents a roadmap to open geodata, with insertions of Romanian particularities. Using dedicated surveys for public agencies, civil society, and private companies, we identified obstacles and drivers for open data in Romania. Analysing the economic and legal frameworks offered the possibility of highlighting the advantages of open geodata in the Romanian community. Furthermore, investing significant efforts in building an open source pilot web geoportal that integrates relevant samples of geodata with respect to data harmonisation, data quality, user friendly navigation, and geodata visualization has completed the research directions for drawing up a roadmap in initiating, developing and maintaining an open geodata initiative in a European Union member country.

2 Inception of the Open geoData movement in Romania

The open geodata movement in Romania predated the open data one. It started in 2007 when geo-spatial.org was launched. geo-spatial.org is a web portal that contains articles, tutorials and data from geospatial domains of expertise such as: digital cartography, historical cartography, neogeography, virtual terrain modeling, remote sensing, geographical informational systems and global positioning systems. It promotes and support the usage of open source software for geospatial and militates for open access to geographical data. All materials published on geo-spatial.org are under Creative Commons Attribution-ShareAlike 3.0 license[1]. Furthermore, geo-spatial.org initiated and developed a collaborative project named eHarta that provided, with the help and within the geo-spatial.org community an open collection of thousands of old georeferenced maps, published and documented (Crăciunescu et. all, 2011). In 2011, the eHarta project won the “Better Data Award” at “Open Data Challenge”. The prize was awarded at the Digital Agenda Assembly, being held in Brussels on the 16th and 17th of June 2011.
Therefore, at the time of the dynamic emerge of the open data movement at European and international level, the Romanian geospatial community understood the great significance of what was proposed.

It is important to highlight from the beginning that a distinction between data and geodata does not exist in the Romanian legislation, as a practical implementation of a principle of the neutrality of the law (except for the INSPIRE Directive). In fact, due to the same principle, the term data is most of the times not even included in the definition of documents (as regulated by the PSI legislation) or of information (as is the case of legislation about the access

1 https://creativecommons.org/licenses/by-sa/3.0/ Last accessed May 2015

to information) (Manolea, 2013). This lack of distinction between the terms has been kept beyond the legislative context, as a considerable amount of data made available is or can easily be georeferenced.

In 2001, Romania adoped the Law for Access to Information 544/2001 that even though it paved the way towards an open communication bewteen public administrations and the public, did not take into consideration the re-use of the public data. Furthermore, when in 2007 Romania had to implement the Public Sector Information Re-use European Directive, the legislative authority decided, as in many other EU countries, to draft a completly different law, 109/2007. This proved to be a much less effective tool, as it lacked any punitive measures for insitutions that did not answer to public requests (Manolea, 2013).

In 2011, Romania makes yet a new step towards open data by adhering to the Open Governement Partnership and then, presenting in 2012 its first National Action Plan (NAP). The responsability of its implementation was placed under the National Anticorruption Strategy (NAS), which made OGP more visible, but there were no financial resources allocated for its implementation. Even more, this adjacency created more confusion when dealing with responsibilities as the NAS is run under the Ministry of Justice and OGP runs under the General Secretary of the Government. As a consequence, the list of assumed actions that were fulfilled was less than satisfactory (Bucur, 2012). A positive turn in the OGP implementation came with the creation of the Department for Online Services and Design (DSOD), under the Prime Ministry Chancellery. DSOD received the clear mission of assisting and monitoring the national e-government programs and the OGP implementation. The first public action was organizing a development day (also called hackathon) at the headquarters of the Romanian Government, in Victoria Palace on the 23rd and 25th of February. This marked the beginning of the journey towards the launching and supporting the official open data portal of Romania.

3 Building the Open geoData initiative

Taking into consideration the actions and activities that pushed forward, supported and promoted the use of open data in Romania, it can be acknowledged that it was a combined approach. The community proved its interest and support through launching community geoportals such as geo-spatial.org. On the other side, DSOD initiated the dialog with public administration representatives. In June 2013, a first result of the collaboration between GEOIDEA.RO team, civil society and the DSOD took place by organising the first big open data event in Romania, Moving forward Romania. Open Discussions for an Open Romania[2]. The event was a stepping stone in the Romanian Open Data movement, with participants from all sectors and invited speakers from the Open Knowledge and Open Government Partnership. The open data event was a side event organised within the framework of the FOSS4G Central Eastern Europe 2013[3].
In September 2013, an intense round of meetings between representative from

2 http://2013.foss4g-cee.org/program/opendata Last accessed May 2015
3 http://2013.foss4g-cee.org/ Last accessed May 2015

ministries, the DSOD and representative of the civil society were held. There were a total of 17 meetings sprawled over a period of one month with the main reason to raise awareness within the public administration regarding the concept of open data, the related legislation and how this movement could benefit the institution and the society. All meetings were advertised and open to the public. Anyone interested could freely participate. The tangible result of each meeting was a list of datasets that could be released as open data on the Romanian official open data portal. As a result of these efforts, in November 2013 data.gov.ro was launched.

On the Romanian open data roadmap, the implications of the civil society and, in particular, of the GEOIDEA.RO research team was a valuable contribution. The GEOIDEA.RO team has joint all relevant meetings with national agencies, such as the National Agency for Cadastre and Land Registration, the National Agency for Agriculture Payments, the National Institute of Patrimony and so on, bringing its contribution in geodata related discussions. As a natural continuation of the collaboration between the Romanian Government and GEOIDEA.RO team, an agreement was signed in December 2013 with the scope of shaping the framework for collaboration. In the following year, GEOIDEA.RO's research team and DSOD co-organized and participated at a significant number of open data events in Romania, either being targeted towards the civil society, the open data consumers, or towards public administrations, the data providers as a first instance.

Even thorough, all events contribute to the advancement and prolifertion of the open data movement, a specific kind present an obvious particularity of practicality. These events are called hackathons or development day and have as a finally goal the creation of applications that use open data to produce diverse services and products. Over the course of the GEOIDEA.RO development, a number of such events were organized, with the active implication of the GEOIDEA.RO team. A pertinent example was the Open GeoData hackathon organised together with the Moving forward Romania. Open Discussions for an Open Romania, where 15 people engaged into 4 open data projects, regarding administrative transparency in Romania, spatial distribution of national TV stations, spatial distribution of official prices for basic products and statistical information on every county, city and commune in Romania.

Figure 1 presents the development of the GEOIDEA.RO project, correlated with the roadmap of the open data movement in Romania and with significant open data events at European or International level.

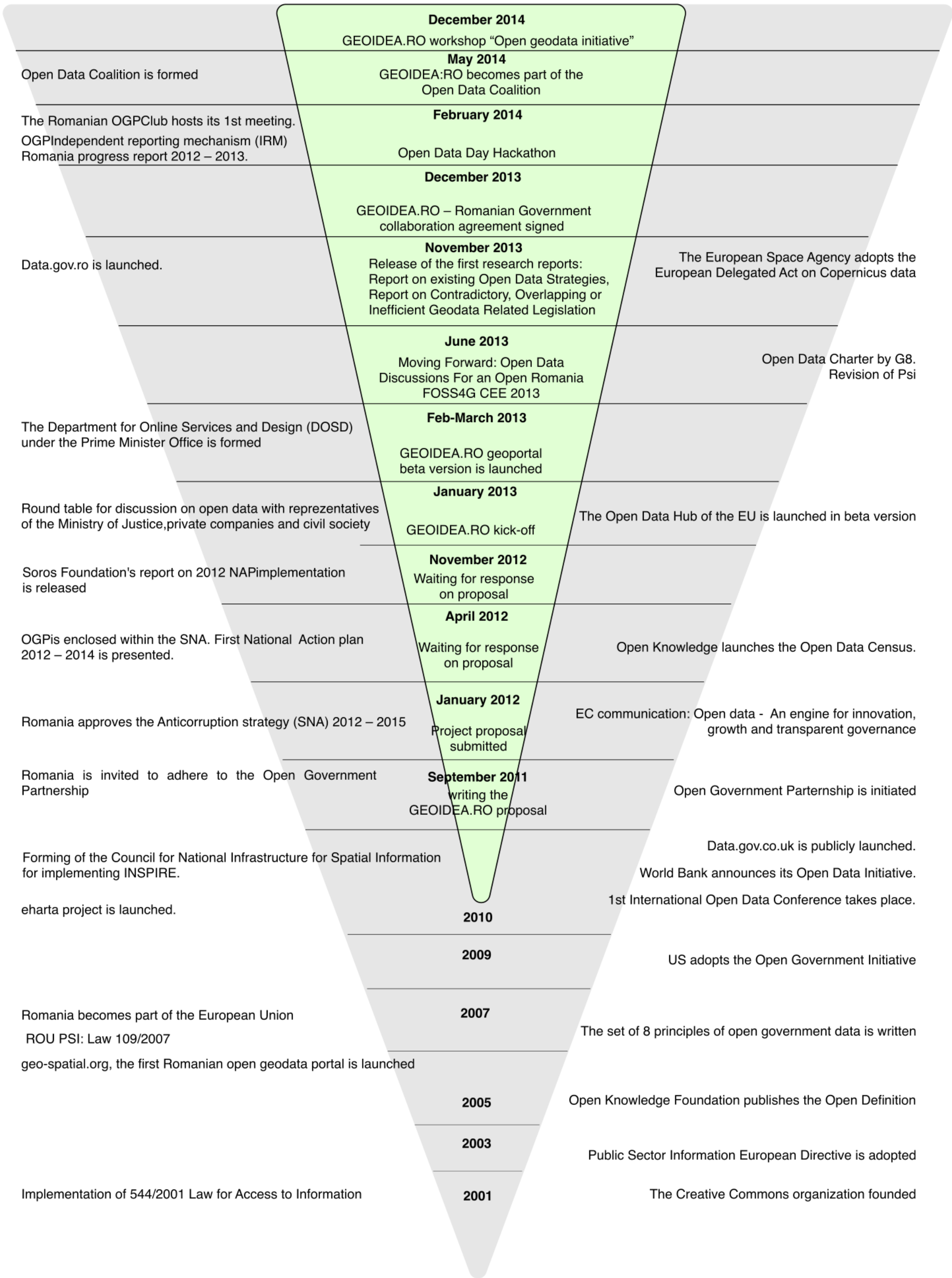


Figure 1. GEOIDEA.RO and the Romanian roadmap to Open geoData

Apart from the efforts that the GEOIDEA.RO research team has done in building the open data momentum in Romania, significant efforts have been invested in building a geoportal that would provide samples of open geodatasets for the Romanian territory. The beta version of the geoportal was launched in the second month of the project and has been improved and populated with datasets ever since.

3.1. Developing an Open Geoportal for Romania

3.1,1 Implementation

An important task of the GEOIDEA.RO project was the development of an open source geoportal for the publishing, viewing and downloading (geo)data released under an open license by public institutions. The requirements for the geoportal were twofold: first, it had to offer the basic functionalities that one would expect from a geoportal, including direct download of data, and second, to provide an intuitive and user-friendly interface.

By virtue of the expertise and experience of one of the project's partner, the Institute of Cartography and Geoinformation at ETH Zurich, the GEOIDEA.RO team could built upon a previous developed architecture. However, several of the components and workflows were not yet fully based on open source technologies, and thus a transposition to open source technologies was required

In current state, the geoportal offers the following functionalities: map-based view of the datasets; information button to access the attributes value; metadata information, including copyright and license; search location by name; selection of area to download via drawing and coordinates input; dataset download as a whole; combination of different datasets into a user map; spatial navigation through panning, zooming, drawing of a rectangle and map re-centering. Presently, it holds 152 data layers grouped within 24 datasets for viewing and downloading.

3.1.2 Architecture

While developing and improving the geoportal over the last two years, several challenges were encountered. It must be mentioned that the transposition to fully open source workflows and components was successful and there was no impact on the capabilities of the geoportal, mostly due to the fact that open technologies in the area of web services, databases and GIS software are now significantly advanced. However, some aspects required more effort and expertise in fine-tuning and integration within the existing architecture than it would have been with an out-of-the box proprietary solution. The actual architecture is based on a classical three-tier architecture, leveraging the advantage of server side technologies to offer a relatively light client (see Figure 1).

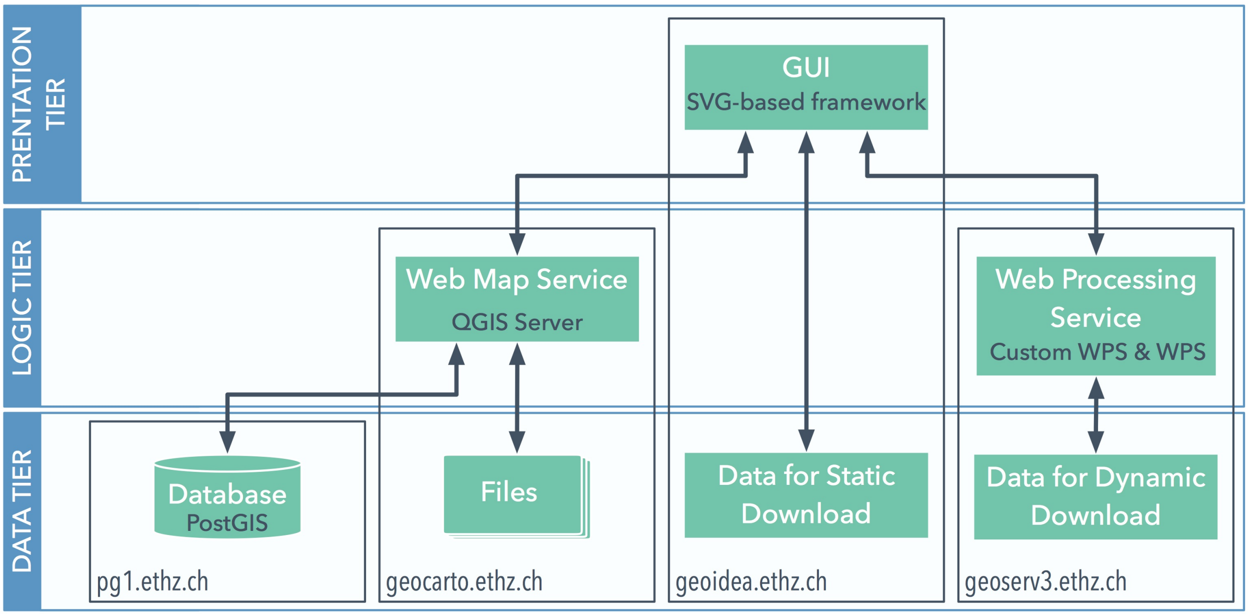


Figure 1: Architecture of the geoportal

3.1.3 Open geodata

Regarding the available open data and the publishing procedures, it has been proven that open data most often comes in many forms and in many states of organization and quality, thus convincing the team to develop internal rules and conventions about their processing and manipulation. . For instance, a naming convention has been developed for the datasets and a subset of the most important metadata fields of the ISO 19115 standards as a metadata minimal reference for the datasets on the geoportal has been extracted. Furthermore, a thematic organization of the data within the geoportal was necessary. The following organization was based partially on the project historic development and on the different data types. The data are organized within four *Map Categories* (*Raster Maps, Raster Data, Vector Data, Data.Gov.ro*), which are then organized in *Map Products* (see Figure 2). A *Map Product* corresponds to one set of data with a common thematic, the same origin and the same geographical extent.

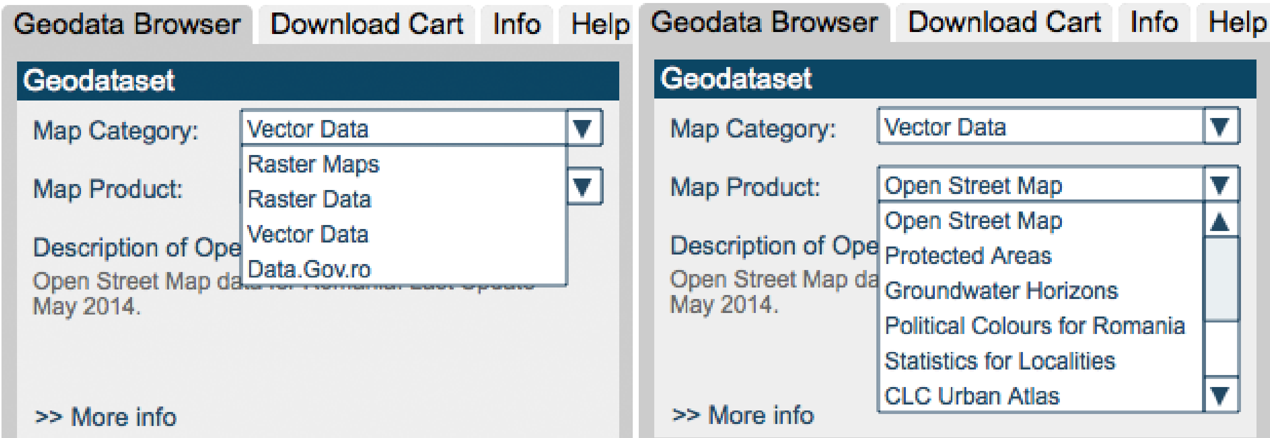


Figure 2: Map Category and Map Product Selection

Regarding the processing workflow, two situations were encountered. One, the dataset was released by a Romanian institution and thus in the Romanian official projection system (i.e. EPSG:31700 *“Dealul Piscului 1970/ Stereo 70”*), and second, it was produced by an international organization which provide the data in a global projection system (e.g. EPSG:4326 “*WGS 84*”), required thus clipping and reprojecting. This step proved to be time-consuming due to the diversity of the data: not one dataset could be processed exactly like a previous one. As a consequence, the preparation and publishing of the geodata on the geoportal is one of the most time-consuming tasks of the workflow.

3.1.4. Graphic User Interface

For the Graphic User Interface (GUI), SVG-based framework was used(developed by carto.net[4], expanded for the GeoVITe project[5]) that was adapted to the project's needs. More specifically, improving the selection area for the download and its integration to the rest of the geoportal from a user's point of view was attempted (see Figure 3). Additionally, it was take into consideration the users feedback concerning the download function. In the first version of the geoportal, this function was only available via the selection of a specific and limited geographical extent, due to the historic limitation of the download speed, processing the download of bigger datasets in a considerable amount of time. However, this issue has lost its meaning and several users requested for the possibility to download a dataset form the entire geographical extent at once. Thus, this functionality was implemented in a straightforward manner that does not load the sever with additional requests: via static links to the data on the sever. The implementation can be seen in the GUI in Figure 4.

4 http://carto.net/
5 https://geodata.ethz.ch/geovite/

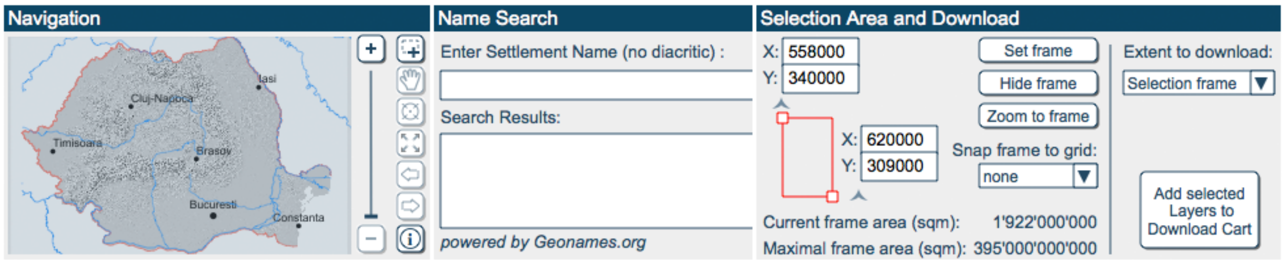


Figure 3: Different functionalities of the GUI: overview map and navigation tools, name search, download tools

Every time the user adds layers to the download cart, the parameters of the request are saved, allowing the user to browse further through the geoportal for additional data. Then, the user can go to the download cart (see Figure 4) and start downloading the data that were requested.

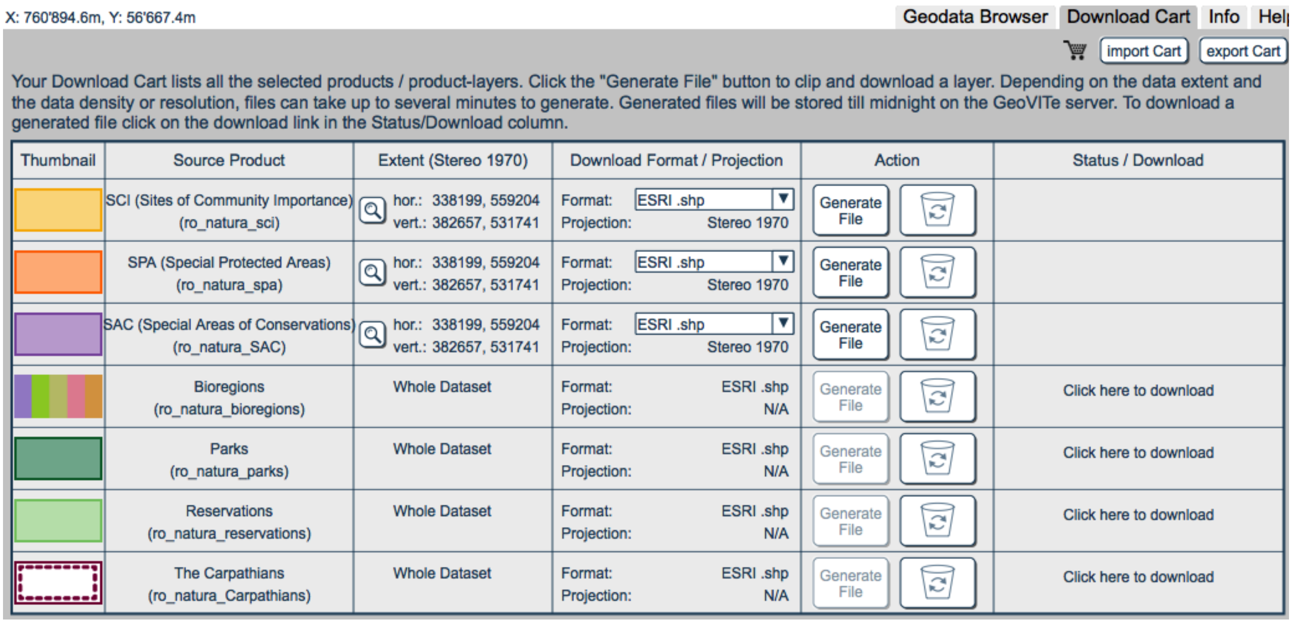


Figure 4: Download Cart: the user requested 7 layers, 3 for a specific extent (top) and 4 for their whole extent (bottom).

3.1.5 Next steps

As the main aspects of the geoportal are functional, the focus has sifted towards small improvements that are still to be done (fine-tuning of the architecture, minor GUI improvements), and on the automation of the processing workflow for the publication of geodata on the geoportal. Some important steps require a direct human control (e.g. data symbolization, projection system issues, etc), but others can be automated. For instance, the generation of the metadata templates from the parameter file, which would allow creating the metadata page for each map product with the basic information already in WMS settings in an automated manner, leaving the task of manually filling up information regarding licensing or attribute values.
Furthermore, a simplification of the procedure to link the WMS and download

services to the geoportal is needed to reduce the potential of error along the workflow. This procedure presently requires a good knowledge of the system and data content, as well as an eye for details.
These automation and simplification aspects aims at supporting the non-expert for the publication of geodata on the geoportal, especially for the tasks that are repetitive and trivial, but can be sources of errors.

4. Conclusions and discussions

It has been proven that the roadmap to open (geo)data in Romania has not been a paved way, but it certainly has been a collaborative adventure. The combined approach, civil society's actions, GEOIDEA.RO activities and the active pressure of the DSOD has proven fruitful. Today, data.gov.ro hosts a number of 249 datasets. Although it may not count for much in comparison with other national initiatives, given the nature of some available datasets, such as the postal codes, railway stations, schedule and many more, the improvement is significant. Nonetheless, Romania has still a considerable road to travel towards building a sustainable, powerful, significant open data paradigm in public administrations.
The interest, support and scrutiny of the open data consumers in Romania is, as well, a constant and powerful driver of the open data initiative. The multitude of events based on open data, such as hackathons, conferences, the OGP Club meetings increase the interest of all sectors' interest, be it private, academia or even public. Applications, such as politicalcolours.ro, developed using open data are yet another example of the involvement of people from the community.

When concerning the legal framework, it has been proven that in Romania's context essential was not the European Directive for Public Sector Information Re-use, but the context offered by adhering to the Open Government Partnership. Due to the political instability of the first year of OGP, the agreed commitments were far from fulfilled. Nonetheless, 2013 brought a significant enhancement in the open data movement in Romania.

GEOIDEA.RO was the first research project in Romania to address the delicate matter of open geodata. In three years of activity, it has successfully collaborated with interested civil organizations, the Governmental officials in promoting, sustaining, developing and researching methods to develop the open (geo)data initiative in Romania. The efforts invested in processing, georeferencing and symbolizing following cartographic rules the open datasets from the data.gov.ro portal introduced higher potential of what was becoming available.

References

- Bucur, A. (2012). OGP Report. Soros Foundation Romania.
- Crăciunescu, V., Constantinescu, Ș., Ovejanu, I., Rus, I. (2011). Project eHarta: a collaborative initiative to digitally preserve and freely share old cartographic documents in Romania. E-Perimetron 6 (4), 261-269.
- Manolea, B. (2013). Report on contradictory, overlapping or inefficient geodata related legislation. GEOIDEA.RO. Retrived May 2015, from http://geoidea.ro/pdf/geoidea_id04_final.pdf

The research was supported by the Romanian Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI) through the Joint Research Project under the Romanian-Swiss Research Programme, Project Number: IZERZ0-142129 through the collaborative research project Geodata Openness Initiative for Development and Economic Advancement in Romania (GEOIDEA.RO).

Open Community Data & Official Public Data in flood risk management: a comparison based on InaSAFE

Riccardo Pasi[1,2], Cristian Consonni[1,3], Maurizio Napolitano[1]

1 Fondazione Bruno Kessler (FBK) – ICT – Digital Commons Lab (DCL)
2 Università IUAV di Venezia – Dipartimento di Progettazione e Pianificazione in Ambienti Complessi (DPPAC)
3 Università degli Studi di Trento – Dipartimento di Ingegneria e Scienza dell'Informazione (DISI)

1 Introduction

Flood and – in general – natural hazards cannot be prevented; however, measures can be taken to mitigate their impacts and prevent them from becoming disasters. Disaster management has been defined as «(the) continuous process that aims at avoiding or reducing the impact of natural hazards» (Poser, Dransch, 2010). Poser and Dransch (2010) have also outlined the importance of using up-to-date and accurate information in all phases of disaster management, as the need of integrating information from many different sources including in-situ sensors, aerial and satellite images, administrative, statistics and socioeconomic census data. New Internet technologies have facilitated fast and easy data collection from the public, giving rise to the idea of using Volunteered Geographic Information (VGI) in disaster risk management.

The paper discusses the opportunities and challenges of using VGI for disaster management, with particular focus on information for the prevention phase. This case study is based on flood risk assessment in two recently flooded cities in Veneto, Italy. We used InaSAFE, a free hazard and risk modeling application integrated in QGIS as a plug-in. InaSAFE offers the capacity to compare hazard and exposure official data with community crowdsourced data. In the case study we compare the results obtained by InaSAFE when using as input the data describing buildings (as exposure layer) drew from OpenStreetMap and from official public data. The goal of this work is answering the following question: Can OSM be used to collect exposure data for DRM? The paper ends analyzing different data sources opportunities and limits.

2 Flood and Disaster Risk Management

Between 1998 and 2009, Europe suffered over 213 major damaging floods (including the catastrophic floods along the Danube and Elbe rivers in 2002), which caused some 1126 deaths, the displacement of about half a million people, at least €52 billion in insured economic losses and severe environmental consequences (EEA, 2010).

The strategic relevance of this issue has prompted the European Parliament to issue the Directive 2007/60/EC on flood risk assessment and management. Floods are natural phenomena, but it is possible reducing their likelihood and limit their impacts: this is the reason why flood risk management plans have been introduced in the European Community legislation, with their focus on prevention, protection and preparedness.

Flood risk management is a subcategory of disaster risk management (DRM), defined as «the systematic process of using administrative decisions, organization, operational skills and capacities to implement policies, strategies and coping capacities of the society and communities to lessen the impacts of natural hazards» (UNISDR, 2009). DRM includes five main stages, these being:

- RECOVERY: the restoration of basic social function;
- RECONSTRUCTION: the full resumption of socio-economic activities;
- MITIGATION/PREVENTION: the permanent reduction of the risk, minimizing both vulnerability and hazard presence;
- RELIEF: the set of activities implemented after the impact of a disaster in order to assess the needs, reduce suffering and limit the direct disaster consequences;
- PREPAREDNESS: the hazard reduction through measures that ensure the organized mobilization of personnel, funds, equipment and supplies within a safe environment.

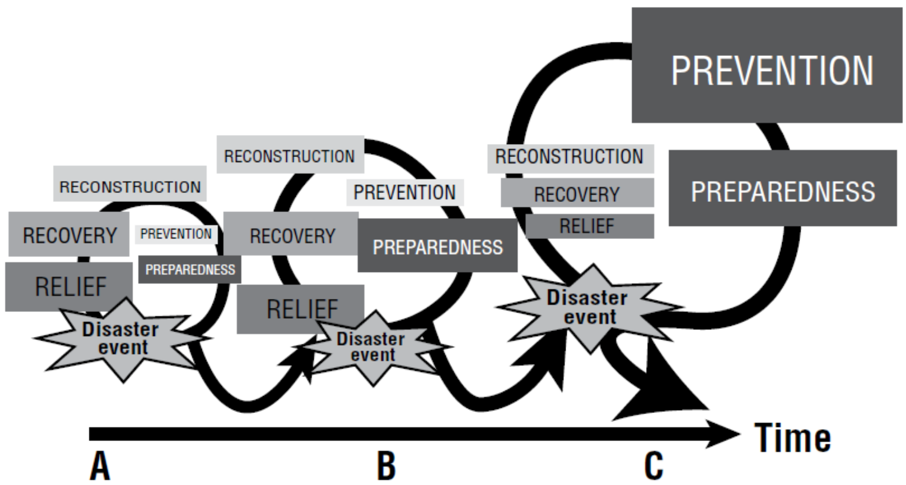


Figure 1: Disaster Risk Management evolution (UNCTAD, 20012)

In the past few decades, the focus has slowly shifted from disaster recovery and response to risk management and mitigation, and ways to reduce the

vulnerability of communities by strengthening their capacity to develop coping strategies (Birkmann, 2006). Figure 1 shows how disaster risk management has been portrayed differently over time. The size of the boxes indicates the importance given to each of the phases; the size of the circles indicates the time between two successive disastrous events.
Geospatial data and technologies (GD&T) are now an integral part of disaster risk management because both hazards and vulnerable societies are changing in space and time.

In the post-disaster phases GD&T plays a major role in rapid damage assessment thanks to optical, thermal and microwave satellite data, but also through crowdsourcing initiatives for collaborative mapping in emergency situations. Equally, in the pre-disaster phases GD&T contributes supporting hazard and risk assessments through remote sensing data, Digital Elevation Models (DEMs) and Interferometric Synthetic Aperture Radar (InSAR) data coupled with census and statistical datasets but also with participatory GIS and web-GIS to spread and collect information among inhabitants.

3 Volunteered Geographic Information for DRM

3.1 The VGI phenomenon

Geographic information provided voluntarily by individuals occupies an important and emerging role among GD&T. This phenomenon has been termed by Goodchild (2007) as “Volunteered Geographic Information” (VGI): WikiMapia, OpenStreetMap and Google Map Maker are examples of platforms used by volunteers to gather geographic information.

The VGI phenomenon has been possible thanks to some enabling technologies, among which Web 2.0, GPS and geocoding services, high quality graphics hardware and high capacity internet connection diffusion. It has become possible for citizens to determine their position accurately, and in turn to gain the ability to make maps from acquired data and eventually to develop cartographic design skills previously possessed only by trained cartographers (Goodchild, Glennon, 2010). The term neogeography has been coined by Turner (2006) to describe the breaking down of the traditional distinctions between expert and non-expert in the geographic information context, since all of the traditional forms of expertise can now be acquired through the use of technology. It is the same process that Butler (2006) called democratization of GIS and that Sui D. Z. (2008) called wikification of GIS, stressing the web-based mass collaboration component, which relies on free individual agents to come together and cooperate to improve an operation or solve a problem.

OpenStreetMap (OSM, https://www.openstreetmap.org), in particular, is an international project aiming to create a free and open map of the world, built entirely by volunteers surveying with GPS, digitizing aerial photography, and collecting existing public sources of geodata (HOT, 2015). The project is supported by the OSM Foundation, an international not-for-profit organization committed to encouraging the development and distribution of free geospatial data. In OSM, online volunteers and remote mappers contribute to create a detailed, precise and up-to-date map, whose underlying data are freely available, in the same sense of freedom that is characteristic of F/LOSS communities (Coleman, 2013), to modify and use under an open license[1] (MapGive, 2014).

3.2 VGI for humanitarian response (post-disaster phases)

Recent disasters have shown the great potential and usefulness of VGI (Goodchild, Glennon, 2010). The response to the 2010 Haiti earthquake demonstrated that non-professionals have significant potential to contribute to post-disaster information gathering (Soden, Palen, 2014). The response also showed that there is a great willingness by volunteers to contribute to such efforts (UNCTAD, 2012). For these same reasons, the U.S. Department of State’s Humanitarian Information Unit recently launched MapGive, an initiative that makes it easy for new volunteers to learn to map and get involved in online tasks (MapGive, 2014).

In general, the literature tends to highlight the role of VGI in the post-disaster phases (Goodchild, Glennon, 2010; HOT, 2015; Poser, Dransch, 2010; UNCTAD, 2012), typically in the immediate aftermath of the event for damage estimation, response planning or spatial data gathering (especially in places where base map data is often scarce, out of date, or rapidly changing). In some cases, in fact, OSM has provided the cheapest source of geographic information, and sometimes the only one available. The Humanitarian OpenStreetMap Team (HOT) has been created to address this need, it is a group – incorporated as a U.S. based 501c(3) non-profit organization – that coordinates the creation, production and distribution of free mapping resources to support humanitarian relief efforts through OSM. HOT acts as a bridge between the traditional humanitarian responders (World Bank – GFDRR[2]) and the OSM Community. The team works both remotely and physically in countries and the majority of the activities of HOT occur remotely: when an event occurs a search for existing data and available satellite imagery is performed with the

1 All the data that we used from OpenStreetMap are © OpenStreetMap contributors and are available under licensed under the *Open Data Commons Open Database License* (OdbL) which is available at http://opendatacommons.org/licenses/odbl/. Further informations can be found at the following page: http://www.openstreetmap.org/copyright (URLs accessed on March, 9th 2015).

2 Global Facility for Disaster Risk Reduction and Recovery (GFDRR).

goal of making this data freely available through the HOT Task Manager, a collaborative mapping platform designed to coordinate the mapping of an area by several remote volunteers. On these basis, the community of OSM is called to improve the map available for the given area. Sometimes, after the initial response, other activities are performed on the ground with the aim of training local people to participate in OSM. In this way the collection and reuse of data is enhanced. (HOT, 2015)

3.3 VGI for sustainable development (pre-disaster phases)

This capability to glean information remote people working on image data but also using ground and local knowledge enables the map to become a locally owned resources, thus proving a great advantage in all the phases DRM (Soden, Palen, 2014). VGI offers a great opportunity to enhance awareness because of the potentially large number of volunteers to act as 'sensors' observing important disaster management parameters in their local environment (Poser, Dransch, 2010).

Besides post-distaster management, this opportunity becomes even more important considering its application in the other DRM cycle phases, especially in prevention and in preparedness. We argue that crowdsourced data can be used for gathering data beforehand in order to succeed in a sustainable development.

A prime example of this has been the work of HOT in Indonesia to gather information about buildings. The focus has been data preparedness and disaster risk reduction in order to minimize the growing rate of exposure and the rising vulnerability. The work of HOT has been part of a broader framework of activities carried out by the World Bank - GFDRR and AIFDR[3] that developed an open source risk modeling software: InaSAFE (Indonesia Scenario Assessment for Emergencies).

InaSAFE is a plugin for QGIS, a free and open source desktop geographic information system (Quantum GIS Development Team, 2014). InaSAFE itself is free and open source, published under the GNU General Public License (GPLv3) that produces «realistic natural hazard impact scenarios for better planning, preparedness and response activities» (InaSAFE Project, 2014). It combines one exposure data layer (e.g. location of buildings) with one hazard scenario (e.g. the footprint of a flood) and returns a spatial impact layer along with a statistical summary and action questions useful to prepare a contingency plan: (a) What are the areas likely to be affected? (b) Which hospital / schools / roads / … will be closed?

3 Australia-Indonesia Facility for Disaster Reduction (AIFDR), a partnership for regional disaster reduction involving Australian (AusAID) and Indonesian (BNPB) governments designed to strengthen Indonesia's ability to reduce the impact of disasters.

Realistic hazard scenarios require scientific, sound and up-to-date data hazard information as well as up-to-date, scale appropriate exposure data and InaSAFE is designed to provide a simple but rigorous way to combine data from scientists, local governments and communities. Information on hazardous areas and people or assets location are usually provided by government departments: in this case, it is possible to import directly raster or vector layers. If the spatial data does not yet exist, InaSAFE external tools – such as the OSM downloader – allow to import data from this ‘external’ source quickly and easily. Its connection with OSM allows more detailed information to be collected, while being part of QGIS allows easy spatial analysis (InaSAFE Project, 2014).

4 Case study: InaSAFE application to compare exposure data

InaSAFE offers the capacity to compare hazard and exposure official data with community crowdsourced data, collected through the experience and knowledge of people. In this case study we will compare the results obtained by InaSAFE when using as input the data describing buildings, which will constitute the exposure data, obtained from OpenStreetMap and from official public data. The goal of this work is answering the following question: can OSM be used to collect exposure data for DRM?

InaSAFE allows to insert hazard data related to any type of natural disaster. We will focus our analysis on floods data given that this type of risk is relevant in Italy (CNR-IRPI, 2015), but this work could be extended to other types of risk.

4.1 Data collection and processing

InaSAFE combines exposure and hazard data layers. This case study employs as hazard data the extension of floods recorded in the recent past in Italy. Two shapefiles – both belonging to the Region of Veneto – has been analyzed:

- Data describing the floods occurred on 1 November 2010 in Vicenza (VI). This dataset was downloaded from the official website of the Municipality[4] and is depicted in fig. 3.
- Data describing the floods occurred on 26 September 2007 in Mestre (VE). This dataset has been produced by the emergency structure created for that specific event[5].This dataset is depicted in fig. 2;

These shapefiles do not contain the information about the height that water has reached, therefore it has not been possible to define thresholds and impact functions to determine when a building has to be considered damaged, but we were able only to tag buildings with a “flooded” or “not flooded” label.

4 http://www.comune.vicenza.it/file/87645-Alluvione_Novembre_2010.shp.zip (URL accessed on March 9th, 2015

5 “*Commissario Delegato per l’emergenza concernente gli eccezionali eventi meteorologici del 26 settembre 2007 che hanno colpito parte del territorio della Regione Veneto*” (O.P.C.M. 3621/2007). The website from where we downloaded the datasets was closed on December 31st, 2012, when the Italian Government declared the emergency to be over.

Figure 2 – 3: Hazard data layer no. 1 (L), flooded areas in Venezia on 26/09/2007, and no. 2 (R), flooded areas in Vicenza on 01/11/2010
[Datasets visualized in QGIS over a layer of satellite imagery from Bing – Bing Aerial layer: © 2015 Microsoft Corporation © 2010 NAVTEQ © Harris Corp Earthstart geographics LLC]

Regarding exposure, InaSAFE can handle just one exposure data layer at a time and many receptors could have been considered such as people, buildings and infrastructures. We decided to focus this case study on building footprints, employing two different spatial data sources:

- Official data, produced by the Region of Veneto and freely downloadable in vector format at 1:5.000 scale[6] (Carta Tecnica Regionale Numerica – CTRN);
- Open community data: extractions of OSM data on municipal basis have been downloaded (in PBF format[7]) from a web service provided by the Italian OSM community[8]. They have been imported in QGIS through the QuickOSM plugin.

The data processing followed different steps for the two data sources: we found that CTRN data required more attention and a more time-consuming processed to be utilized for our analysis.

CTRN data are organized in cartographic tiles, forcing us to search and select tiles in advance; furthermore, many shapefiles exist for each tile: we had to identify where building information were stored, download and merge them. We then removed from the resulting shapefile the entities not related to buildings: courtyards, steps, graveyards, gardens, etc.; finally we clipped the building data to the flood extent and we added a field (named "type"), which is required

6 http://idt.regione.veneto.it/app/metacatalog/ (Data have been downloaded on December, 2015)

7 The Protocolbuffer Binary Format (PBF) is primarily intended as an alternative to the XML format. It is used to support random access at the OSM 'fileblock' granularity, preserving the order of OSM entities, and tags on OSM entities.

8 http://osm-toolserver-italia.wmflabs.org/estratti/it/comuni/ (Data for the area of Venice have been downloaded on December 19th, 2015 and data for the area of Vicenza have been downloaded on December 15h, 2015)

by InaSAFE, copying in it building typologies from an already existent field (named “descrz”).
OSM data, instead, were downloaded in a unique file on municipal basis: no search, selection and merge operations were needed, but we had to select which geometry entities to import (“multipolygons”). Then we filtered out the entities not related to buildings as we have done for CTRN. Finally we clipped the building data to the flood extent and we added a field (named “type”), encoding in it building typologies: in this case, in order not to lose information about building typologies, we used multiple fields (“name”, “type”, “amenity”, “building”, “landuse”, “leisure”) as inputs for the encoding process. Some reclassification and harmonization of tags was required.

4.2 Data comparison and results

In this section we presents the results of the impact calculation we produces using InaSAFE. The system outputs a layer representing flooded buildings together with statistics of flooded buildings detailed for various building typologies. For each hazard dataset, we have repeated the same process using different exposure layers.

Figure 4: InaSAFE output layer no. 1: flooded building in Venezia using official (L) and OSM data (R)

Figure 5: InaSAFE output layer no. 2: flooded building in Vicenza using official (L) and OSM data (R)

As shown by the maps reproduced in fig. 4 and fig. 5, at small to medium scales, there are no visual remarkable differences between the output layers produced by InaSAFE: official and community data are comparable. This result can be explained considering the choice made by the administration of Region of Veneto to open part of its geographic information databases in 2011. CTRN data has been imported in OSM[9] and these data served as complete and solid basis for the community that than focused on improving the map. More significant differences can be detected at a larger scale. Here, the difference can be explained with the lack of updating of official data: the official datsets dates back to 1994 for Venezia and to 1999 for Vicenza. This is a considerable difference since OSM data is constantly updated. albeit unevenly and unsystematically. An example is shown in fig. 6, depicting the area of VEGA, the Venetian Science & Technology Park, built in the decade 1993–2004. While the official map still shows the pre-existing industrial buildings, OSM mappers have drawn the new layout of the area.

Figure 6: QGIS screenshot on VEGA area: official data layer (L); OSM data layer (in the middle) [© OpenStreetMap contributors]; Bing Aerial photograph (R) [© Microsoft Corporation]

Further differences between official and community data can be highlighted considering non-spatial statistics on building typologies.
Table 1 and 2 show that community data achieve a greater information granularity: OSM data contains a building types classification richer than what is available in the CTRN data. This greater detail can contribute to refine current disaster risk management results and methods, both in prevention and in emergency phases.

9 The import occurred in the last months of 2011, following a clear method that can be consulted at: http://wiki.openstreetmap.org/wiki/Veneto/Guide_e_documentazione/Import:_dalla_CTRN_Veneto_a_OSM (URL accessed on March, 9th 2015)

VENEZIA (2007)	BUILDING TYPE	OSM data			OFFICIAL data		
		TOTAL BUILDINGS	**BUILDINGS FLOODED**	**%**	**TOTAL BUILDINGS**	**BUILDINGS FLOODED**	**%**
	Bell Tower	16	0	0	15	0	0
	Chimney	8	2	25	14	5	36
	Church	122	31	25	109	29	27
	Civil building (undef.)	26448	8553	32	24515	8430	34
	Collapsed	33	1	3	115	18	16
	Commercial	18	1	6			
	Courthouse	1	0	0			
	Cultural heritage	89	1	1			
	Farm auxiliary	160	23	14	119	7	6
	Fire station	2	0	0			
	Hospital	3	0	0	8	0	0
	Hotel	11	2	18			
	Hut	12309	3872	31	12347	4041	33
	Inflate balloon tent				4	1	25
	Industrial	1893	643	34	1983	760	38
	Mall	7	0	0			
	Monument				3	2	67
	Museum	3	0	0			
	OTHER	30	3	10	89	10	11
	Parking (building)	3	0	0			
	Police station	1	0	0			
	Public building	21	1	5			
	Residential	49	17	35			
	Restaurant	3	0	0			
	Roof	1052	136	13	700	172	25
	School	93	33	35	87	32	37
	Silo	190	59	31	442	196	44
	Sports facility	19	3	16	57	8	14
	Squat	1	1	100			
	Storage tank	131	100	76			
	Supermarket	6	0	0			
	Theatre	4	1	25			
	Townhall	2	0	0			
	Train station	5	0	0	6	0	0
	Underconstruction	29	9	31	37	18	49
	University	7	0	0			
	TOT	**42.769**	**13.492**	**32**	**40.650**	**13.729**	**34**

Table 1: Absolute and percentage values of flooded buildings split by building typology in Venezia (2007) using OSM and official data.

	BUILDING TYPE	OSM data			OFFICIAL data		
		TOTAL BUILDINGS	BUILDINGS FLOODED	%	TOTAL BUILDINGS	BUILDINGS FLOODED	%
VICENZA (2010)	Bell Tower	38	10	26	52	12	23
	Church	362	57	16	409	65	16
	Civil building (undef.)	36063	3435	10	36933	3577	10
	Collapsed	3	0	0	117	14	12
	Commercial	51	4	8			
	Courthouse	3	0	0			
	Cultural heritage	58	10	17			
	Depot	33	0	0	38	0	0
	Ederle Camp	1	0	0			
	Farm auxiliary	746	92	12	775	92	12
	Fire station	2	0	0			
	Garage				2	0	0
	Hospital	152	17	11	165	17	10
	Hotel	3	2	67			
	Hut	4582	390	9	4753	407	9
	Industrial	2054	137	7	2375	157	7
	Inflate balloon tent				5	2	40
	Monument				9	3	33
	OTHER	12	0	0	1079	111	10
	Police station	26	0	0			
	Public building	10	1	10			
	Residential	25	1	4			
	Restaurant	63	7	11			
	Roof	10270	810	8	10547	852	8
	School	428	91	21	561	122	22
	Silo	212	10	5	429	15	3
	Sports facility	5	4	80	46	15	33
	Supermarket	2	0	0			
	Theatre	4	3	75			
	Toll booth				1	0	0
	Tower				19	0	0
	Townhall	19	5	26			
	Train station	14	0	0	18	0	0
	Train accessory building				7	0	0
	Under construction	23	1	4	78	2	3
	University	2	0	0			
	TOT	**55266**	**5087**	**9**	**58418**	**5463**	**9**

Table 2: Absolute and percentage values of flooded buildings split by building typology in Vicenza (2010) using OSM and official data.

During the identification of exposed assets from a flood risk point of view, the most sensitive building typologies available in the CTRN dataset are hospitals, schools and industrial buildings. Conversely, the civil building category is too broad and undefined to be considered useful: are these building inhabited? Which functions are contained?

OSM data help to answer these questions identifying residential buildings, public facilities (town halls, court houses, police stations, universities), commercial activities (hotels, restaurants, malls, supermarkets), cultural heritage sites (museums, theatres, historic buildings). Some OSM categories – as fire stations and storage tanks – may even be used to assess diffusion of

pollutants after a flood.
Flooded building percentages are homogeneous in both urban contexts despite the different data sources used (around 30% in Venezia, 9% in Vicenza). Our analysis with InaSAFE detects a difference in the total number of buildings available in OSM and CTRN data: in Venezia OSM detects more buildings than the CTRN, while in Vicenza the situation is the opposite. This difference may be explained with the fact that the two cities belong to different socio-economic contexts and they may have experienced different transformation dynamics, but this analysis would be outside the scope of this paper; we can give a more limited explanation noting again that OSM data are updated up to December 2014, while please CTRN date back several years (1994 for Venezia and 1999 for Vicenza) and for this reason we would have expected a greater number of elements in OSM data in both cases, while for Vicenza this is not the case. A more detailed analysis of CTRN data in the area has given some indications for this apparent shortage of buildings in OSM data, namely: (a) the presence of some errors in a cartographic tile of the CTRN and (b) during the import of CTRN data in OpenStreetMap building shapes have been simplified in OSM (fig.7).

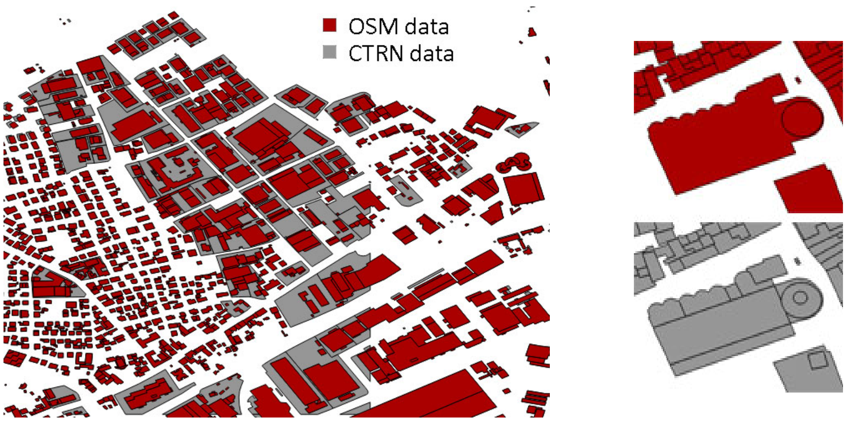


Figure 7: Large scale comparison of CTRN and OSM data in Vicenza showing fine-grained differences: errors in a cartographic tile of the CTRN (L) and examples of OSM simplified building geometries compared to CTRN data (R).

At the same time, it should be noted that OSM data are not complete. For example, while OSM data show that in Vicenza there are only 10 public buildings, 3 hotels and 2 supermarkets, we can assume that the actual numbers are higher.

4 Conclusions

This paper discussed the opportunities and challenges of using VGI for disaster risk management, focusing on its application for prevention and risk assessment rather than in response and post-disaster phases. The case study, based on flood risk assessment in two urban contexts of the Region of Veneto, Italy, has been used to compare different exposure data layers, targeted on buildings drew from regional official maps (CTRN) and OpenStreetMap (OSM).

Our aim was to investigate if OSM – at the state of the art now – could be used to collect exposure data in DRM. To the opening question, our analysis show that the answer is affirmative. The risk assessment profiles obtained with the two data sources are comparable and, in addition, OSM data granularity and richness is higher. OSM data are also easier to retrieve and process than CTRN data, thus simplifying the professionals task in huge or short term scopes.
Furthermore, we did not exploit the existing OSM data to their full potential: the quality of our results on buildings could be improved using other available information such as (a) information from OSM nodes encoding points of interest, and (b) information from OSM 'landuse' tags. This information could be used to refine the specification and detection of building typologies.
An example is shown in the figure below (fig.8). Note that only the buildings depicted in red have a defined OSM attribute associated to them. In this case we chose only some attributes ('amenity' and 'shop' tags) from the entire punctual OSM database. Nonetheless it is clear how information granularity could increase by mean of this simple geoprocessing step; here – for instance – exposure maps can be enriched with a new feature class: commercial destination at ground floor.

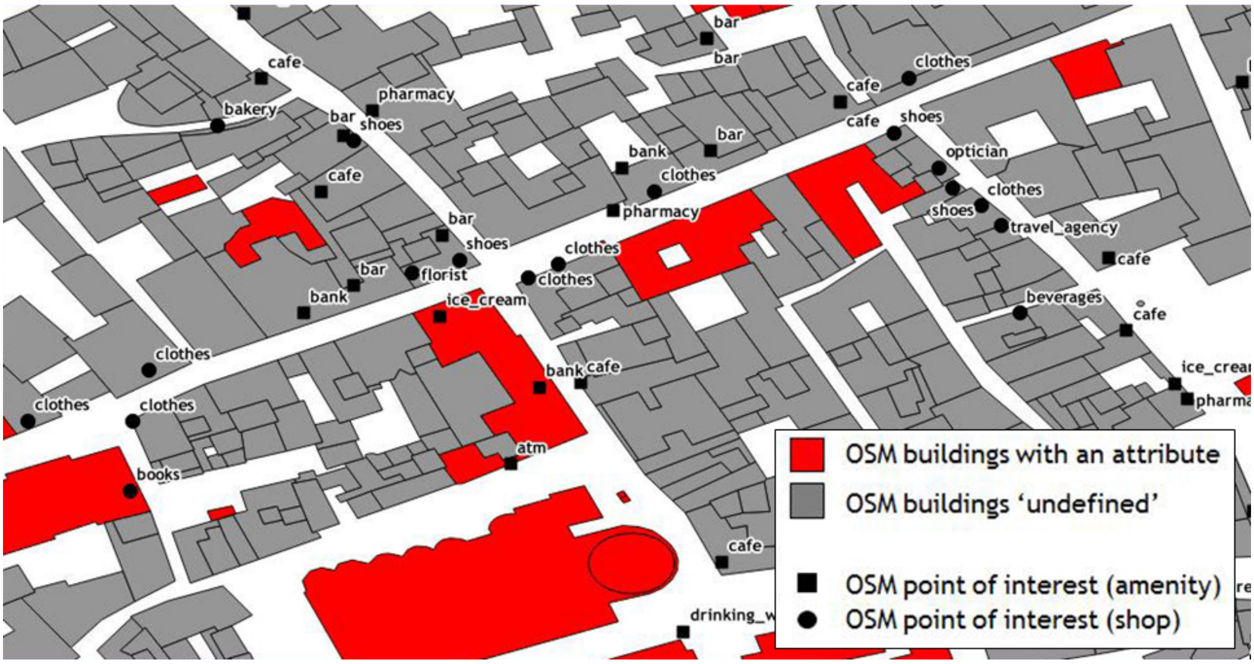


Figure 8: preliminary example on points of interest and buildings intersection

Another advantage related to using OSM data would be the possibility of improving its data: the institutions should foster the use of OSM by citizens and they should organize meetings with OSM mappers could be organized in order to identify some strategic elements for flood risk assessment in order to plan mapping campaigns, a practice already in use in the OSM community with the name of “mapping parties”. This practice has already shown its potential in international initiatives such as MapGive and the work the Humanitarian OpenStreetMap team.

Some of the elements that could be added to OSM and used to refine the building vulnerability assessment in flood events are:

- the existence of underground floors (which anyone can easily identify through the presence of basement windows, descendent stairs, parking/garage ramps, etc.). In this regard, we note that the existing OSM tag “layer=-1” it does not report the existence of underground elements by itself, but it is used to order overlapping objects such as intersecting roads;
- the land use of ground floors (retail/office, public service, residence, warehouse, etc.);
- the approximate height – in relation to the road – at which the most vulnerable urban land uses are located, due to the buildings shape and structure (pilotis, mezzanine floors, basements, etc.).

Through the practice of editing the map citizens would also gain a greater awareness about flood risk in their city and neighborhoods: spreading this culture of risk among citizens would fulfill a great, and usually neglected, need in flood risk management, whose benefits have already been shown in humanitarian post-disaster initiatives. In addition, even if this paper is focused on flood risk, InaSAFE can analyze hazard data related to any type of natural disaster and further work could extend this analysis to other types of risk.

To conclude, this case study has shown how the choice of the Region of Veneto to open its geodatabase has allowed the import of CTRN data in OSM. The community of OSM has taken advantage from this import, relying on map with higher spatial homogeneity, but in return they are updating continuously this data, that otherwise remains static and outdated thus eventually useless. In turn, regional and local authorities could take advantage from this update; more, they may even promote the use of OSM prompting their technical and operational staff (technicians, police, Civil Protection volunteers, etc.) to update OSM constantly and timely. We believe that this process would benefit all the communities interested in having accurate, up-to-data and reusable geospatial data.

References

Butler, D. (2006). Virtual globes: The web-wide world. *Nature*, 439, 776–778.

Coleman, E. Gabriella (2013). *Coding freedom: The ethics and aesthetics of hacking*. Princeton, NJ: Princeton University Press. Retrieved April 20, 2015, from http://gabriellacoleman.org/Coleman-Coding-Freedom.pdf.

CNR-IRPI – Consiglio Nazionale delle Ricerche, Istituto di Ricerca per la Protezione. (2015). *Rapporto periodico sul rischio posto alla popolazione italiana da frane e inondazioni – Anno 2014*. Retrieved April 20, 2015, from http://polaris.irpi.cnr.it/report/last-report/.

European Parliament. (2007). *Directive 2007/60/EC on the assessment and management of flood risks*. Official Journal of the European Union. Retrieved April 20, 2015, from http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007L0060&from=EN.

EEA – European Environment Agency. (2010). *Mapping the impacts of natural hazards and technological accidents in Europe: an overview of the last decade* (EEA Technical report no 13/2010). Copenhagen: Author. Retrieved April 20, 2015, from http://www.eea.europa.eu/publications/mapping-the-impacts-of-natural.

Goodchild, M. F. (2007). Citizens as sensors: the world of Volunteered Geography. *GeoJournal*, 69 (4), 211-221.

Goodchild, M. F., Glennon, J. A. (2010). Crowdsourcing geographic information for disaster response: a research frontier. *International Journal of Digital Earth*, 3:3, 231-241.

HOT – Humanitarian OpenStreetMap Team (2015). Web site retrieved April 20, 2015, from: http://hot.openstreetmap.org.

InaSAFE Project. (2014). *InaSAFE documentation* (release 2.1.0). Retrieved December 10, 2014, from http://inasafe.org/pdf/InaSAFE-2.1-Documentation-en.pdf.

MapGive (2014). Web site retrieved April 20, 2015, from: http://mapgive.state.gov.

Poser, K., Dransch, D. (2010). Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica*, 64, 1, 89-98.

Quantum GIS Development Team (2015). Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. Retrieved April 20, 2015, from: http://qgis.osgeo.org.

Soden, R., Palen, L. (2014). *From Crowdsourced Mapping to Community Mapping: The Post-Earthquake Work of OpenStreetMap Haiti*. Retrieved April 20, 2015, from University of Colorado, Boulder, Department of Computer Science. Web site: https://www.cs.colorado.edu/~palen/palen_papers/HaitiCOOP_Final.pdf.

Sui, D. Z. (2008). The wikification of GIS and its consequences: Or Angelina Jolie's new tattoo and the future of GIS. *Computers, Environment and Urban Systems*, 32, 1-5.

Turner, A. (2006). *Introduction to neogeography*. Sebastopol, CA: O'Reilly. Retrieved April 20, 2015, from http://highearthorbit.com/neogeography/book.pdf.

UNCTAD – United Nations Conference on Trade and Development. (2012). *Geospatial Science and Technology for Development* with *a focus on urban* *development, land administration and disaster risk management* (UNCTAD Current Studies on Science, Technology and Innovation No. 6). United Nations, New York and Geneva. Retrieved April 20, 2015, from http://unctad.org/en/PublicationsLibrary/dtlstict2012d3_en.pdf.

UNISDR – United Nations International Strategy for Disaster Reduction. (2009). *Terminology of disaster risk reduction*. United Nations, Geneva, Switzerland. Retrieved April 20, 2015, from http://www.unisdr.org/files/7817_UNISDRTerminologyEnglish.pdf.

What is the Value of Open Data

Hans Gregers Petersen[1]

[1] Septima P/S, Denmark

Abstract

In 2012 the Danish Government announced an ambitious plan, the "Basic Data Program", to free most of the basic data in Denmark. We have found that several business and software development processes have changed with the release of data, and it has brought an extreme growth in the use of data. Both have affected our business possitively. The wider use of data has also meant that public servants have had to adapt; to face fear of errors and ever more demanding customers.

Keywords

Basic Data, Open Data, Business Model

The Danish Release of Basic Data

In 2012 the Danish Government announced an ambitious plan, the "Basic Data Program", to free most of the basic data in Denmark. All basic data that is not confidential should be freely available to citizens, companies, NGOs etc.

The Danish government had a wholehearted and rather broad approach to which data was covered:

"*...various core information about individuals, businesses, real properties, buildings, addresses, and more.*" (The Danish Government, Local Government Denmark [DGLGD], 2012)

and

"*...As a general rule, all basic data is to be made freely available to all public authorities, private businesses and individuals.*" (DGLGD, 2012)

The idea behind the release of the basic data was to gain growth and efficiency in both private and public sector, and to ensure open and free access to information. These objectives were to be reached by significantly lowering the costs behind using data that had already been paid for once – from being mostly very expensive the data was now to become free.

This new strategy also involved paying just a little more upfront, and then expecting this small investment returned in the form of a more efficient public sector:

"...map data, cadastral maps, Central Business Register data, and company data will be financed by the government and released to the public and the private sectors..." (DGLGD, 2012)

As well as a focus on coordinating the many registers and dataset holders in national and local government to gain efficiency both financially (by not paying for double registration) and IT-wise (by not doing multiple registration of the same information in different registers:

"...to ensure efficient, effective and coordinated development and use of basic data, a cross-institutional basic-data committee is to be established..." (DGLGD, 2012)

Can You Sell a Free Beer?

Partly as a consequence of the release of Basic Data this abstracts author started a consulting company together with six former colleagues in January 2013. There was a strong belief that money could be made from the value of our knowledge and through helping companies and the public sector use these newly freed data. This was also supported by the success of freeing the Danish address dataset which have been shown to be a very possitive case (Danish Business Authority, 2010)

What we have found is that it has not only changed what we can do for both private and public sector clients, it has also changed **how** we do it. As data is now free, we do not have to wait for (public sector) clients to approach us with ideas (to complete on **their** data) – we can now approach (both public and private sector) clients with products and proposals of our own. An apparently tiny thing as being able to product develop on our own, has turned the business model upside down in some instances. Although you cannot sell a free beer, you can sell the knowledge of how to open the free beer, or a ready-to-use bottle opener, and possibly some consulting on how you can get to enjoy the free beer the most.

Is Free Data Worth Anything?

One could think that free stuff is usually not worth anything. The release of basic data in Denmark has shown that this is not always the case. The use of the data has seen an extreme growth since their release – and the wider use is spread over many types of users (Steffensen, 2013). Denmark has commisioned a consolidation of the distribution of basic data in a central data distribution platform, to facilitate this process even more (Bjørn-Møldrup, 2013).

Even the Danish society has been affected by these new possibilities - for instance news papers and NGOs now have far easier access to a broad spectrum of map data that can be used to analyse and visualise information.

Open Information Means Insight and Involvement

In the Danish public sector there is a pride around the quality of data. There is also an acknowledgement of the fact that every dataset contains errors. The freeing of data therefore – naturally – also means that these errors will be seen by others. It was therefore a natural concern amongst some public servants that errors in data might give grounds for incorrect conclusions when data was publicly available. We have seen, though, that even though the free data is once in a while used wrongly, the feared panic (e.g. over flooding risks or the like) has not arrived.

What has indeed been seen is that the opening of data has not been a one way distribution. The users have demanded more information to be freed, more services to be created, and even data formatted in ways that are not compatible with good governmental practice – sometimes these demands have been within reason, sometimes not.

References

- ✔ Bjørn-Møldrup, Mads; Danish Geodata Agency. (2013). *Grunddata og Datafordeleren*. Copenhagen, Denmark: Kortdage Conference, Geoforum. Retrieved From: http://www.kortdage.dk/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fKortdage+2013%2fPr%26aelig%3bsentationer%2fFrie+data%2fKortdage+Datafordeler+-+Mads+M%26oslash%3bldrup.pdf
- ✔ The Danish Business Authority (2010). *Værdien af danske adressedata - Samfundsgevinster som følge af frikøb af adressedata mv. I 2002. Copenhagen, Denmark:* ERHVERVS- OG BYGGESTYRELSEN. Retrieved From: http://www.adresse-info.dk/Portals/2/Benefit/Vaerdien_af_danske_adressedata_2010-07-07l.pdf
- ✔ The Danish Government, Local Government Denmark. (2012). *Good Basic Data for Everyone – A Driver for Growth and Efficiency. The eGOVERNMENT strategy 2011-2015.* Copenhagen, Denmark: Rosendahls – Schultz Grafisk. Retrieved From: http://uk.fm.dk/publications/2012/good-basic-data-for-everyone/
- ✔ Steffensen, Bente; Danish Geodata Agency (2013). *Frie data - hvordan er det gået? Copenhagen, Denmark: Kortdage Conference, Geoforum. Retrieved From: http://www.kortdage.dk/Admin/Public/DWSDownload.aspx?File=%2fFiles%2fFiler%2fKortdage+2013%2fPr%26aelig%3bsentationer%2fFrie+data%2fKortdage2013_BenteSteffensen_udennoter.pdf*

Privacy in positioning - a review

Scott Cadzow

Cadzow Communications Consulting Ltd

Abstract

The role of geo-location in a wide range of applications and the use of remote servers to provide both localised and personalised service mean that there is some reason for caution from users with regard to the data that is shared across the network. The intent of this paper is to show that the privacy risks can be countered successfully and illustrates how this has been addressed using examples from a number of EU funded projects including i-Locate, i-Tour, i-Scope, UNCAP and SUNSHINE.

Keywords

Privacy, Human Factors, Dignity, Protection, Security, Confidentiality

1 Introduction

Security is quite simply badly understood in technical domains and where security goes the need to protect privacy inevitably follows and in like manner is also misunderstood. In the traditional security arenas the guiding principles have been built around the “need to know” and have stood us in good stead for many years. What legislation and society have added to the “need to know” is the idea of having a “right to know” and this complicates a little the actual provisioning of security functionality. Looking at this in more detail we can obviously identify broad classes of user who have the need or the right to know and conditions in which these rights change.

The geo-space (real world) that we live in has many restricted spaces, often signposted, but just as often not. When we move our real world to the virtual world we need to be able to replicate the existing restrictions and this has implications on the means we adopt to implement security and privacy controls.

Figure 1: A real world access control sign.

Our primary access control in the real world is achieved through signs (for example see figure 1) and education that give us rules for behaviour. In the real world though we are relatively flexible in implementing rules and again when using rules for controlling access to data in the virtualised GeoSpace we also have to understand that sometimes they have to be broken.

Figure 2: Conditioning in access control.

Our conditioning tells us that men shouldn't use the women's restroom but should that always be the case (see figure 2)? Should an ambulance or fire engine be allowed to drive down a one way street the wrong way?
In fact those who play computer games such as Civilisation can visualise this geo-mapping problem: When you gain territory you see more of the map. In real world geo-maps what you see will be determined by your access rights.

2 Location and smart city

In 2012, 50% or more of the population in every EU country lives in non-rural areas (urban or intermediate regions, Eurostat, 30 March 2012). The continuing growth in city (non-rural) populations poses a challenge for services and infrastructures which need to grow at an equal rate in order to support rising demand. These increased populations will create new economic opportunities and social benefits for the cities in which they live. To ensure that our cities can support the population demands, they must become smarter through improved utilisation of the resources available. Whilst this paper describes how geospatial technologies can be used to support the development of Smart Cities it also addresses the rights for location privacy in this environment.
Many applications in smart cities are location centric: Find the nearest restaurant; Show me the way to ...; What happened here in history?; and so on. The user's role in most of these applications requires revealing current and, often, planned locations, and implicitly revealing behaviour too.
In addition to the user the city reveals behavioural information tied to location: A restaurant is here; A sports facility is here; A centre of religion is here (church?). When building geo-apps location rules have to be applied too: If road-x is one way north to south and entry requires travelling north then entry shall be disallowed. If the geo-rule is broken who should know and who should act? Obviously driving on a one-way street in the wrong direction is illegal but can a smart city integrate with a smart vehicle to inhibit such behaviour?
There are wider ethical and societal issues with location, behaviour and privacy. If you are in a cinema but forgot to switch off your phone should your phone recognise your location and the societal rules that apply there to enable silent mode? Should your phone deny incoming calls with the message "the person you have dialled is in the cinema and cannot take your call"? Obviously there is a potential breach of privacy here and it concerns how you consent to the use of knowledge of your location and any inferred behaviour. It may be reasonable

to post an entry on Facebook or similar social media site that you're at the cinema and loving the film (notwithstanding any local rules this may break). It may be reasonable for a parent worried about the whereabouts of a child to interrogate the child's phone for its location. But what if the parent is an abusive one and the child is hiding? What if the requestor is an employer and the employee was meant to be at work and not the cinema?
Routing is one of the more obvious location based apps and when using online (connected) apps there is a surprising amount of personal data even if the service is claimed to be anonymous. Routing apps need to know where you are, where you'e going, how you intend to get there, how many of you are travelling (singly, in groups, with children), your preferences (using public transport, quiet routes, avoiding motorways). The overriding concern is how to give a personalised service without overindulging in gathering of what is termed Personal Identifiable Information (PII). In this case it is not the explicitly obtained information that is of concern but the possibility to infer information from the data offered to the service. Whilst most consent statements identify data that is explicitly gathered very few extend to the inference of data even though this is the result of processing the explicitly obtained data.

3. Security and complexity

When security experts are asked about protection they often, too often, think in terms of cryptography and the advantages that can bring in proof that security solutions work. The cycle of conventional good security work is shown in Figure 3 where the aim is to identify risks or threats to the system, and sequentially prevent them, make preparations if the threat is enacted and respond to that scenario. Quite properly this cycle has a pessimistic view of the world – no matter what we do in identifying and preventing bad things we will be affected at some point. These are all activities, i.e. things we need to do to assure ourselves that something is secure.

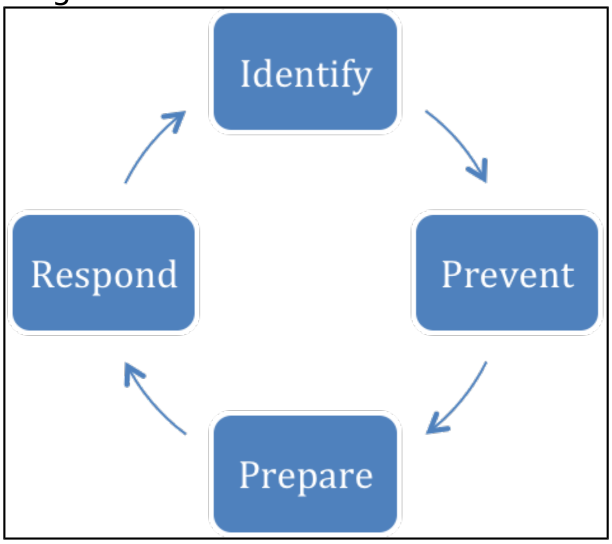


Figure 3: The security cycle.

Where security experts take a mostly negative view of the world the general public is remarkably positive about how they interact with the world both real

and virtual. However the positive view can very rapidly become negative if and when someone suffers identity theft, a virus infection, or even a simple component failure. Part of the problem is complexity: Modern ICT systems are massively complicated and the point where the human being sits is often at the end of a very long chain. When we consider the complexity of protecting the user we should very simply and crudely consider the number of variables we have to have control of and taking the factorial of them. So a system with say only 4 variables has a complexity measure of 24, adding just one more variable takes this measure to 120, adding 2 takes it to 720 and so on. We can think of many ways to consider complexity but the simple rule is that the more variables there are the more complex is our control of the system and the greater the likelihood of missing something (identifying ways of controlling and eliminating 24 problems is obviously less error prone than controlling and eliminating 120).
Security experts in general also try and consider ways to simplify a system to one that allows them to solve only 4 problems:

1. Confidentiality
    a. Ensuring that data transmitted by Alice to Bob can only be seen by Bob
    b. Conventionally provided by encryption across open networks
2. Authority
    a. Ensuring that if Alice is trying to do something that she actually is allowed to
    b. Enabled by a large number of techniques including Access Control Lists (list of identities allowed to access the system), and various access control schemes based on identity, role, consent, trust, location. Many are managed by policy control engines.
3. Integrity
    a. Internal consistency or lack of corruption in electronic data
    b. Ensuring that changes to the system can be identified
    c. Often provided for data using cryptographic hashes
4. Identity
    a. Ensuring that Alice is really Alice and not Bob pretending to be her
    b. Often achieved by Identity Management infrastructures and strong authentication

There is an impact of applying these measures – they change the system and that has to be factored in too. In other words how do you protect the system from the changes imposed by the protection?

4. Security and privacy

In addition many security measures are privacy negative. What this means is that in order to provide strong proof of identity you may need to declare additional private information, e.g. there are more than 2 Alices in the world, so how do we distinguish good Alice from bad Alice? Or from the other good Alice? Similarly, to determine authority this may be dependent on your physical location, e.g. are you allowed to make an international phone call from your current location?

It is also worth noting that these words as applied by security experts are slightly adrift of normal, particularly literature based, interpretations. For example Killinger (Killinger 2010) states that "Integrity is a personal choice, an uncompromising and predictably consistent commitment to honour moral, ethical, spiritual and artistic values and principles." For a network or machine system to demonstrate integrity of this nature is quite different from simply expecting it to be “whole” when measured against a known state of being whole.
The approach taken in i-SCOPE is to view a city as a layered geographic entity. The base layer is the topographical model of the physical environment. Subsequent layers can be added, limited only by the availability of knowledge. To a large extent cartographers have been using this process for centuries. However, digital geospatial technology allows us to massively expand the volume and number of datasets that can be layered and, with the immense processing power now available, build new capabilities from the datasets that we have now or in the future. In extending the ideas of i-SCOPE the i-locate project has moved indoors and to the Internet of Things (IoT) space with a view to identifying where objects are and how to get to them. An object in IoT can be both natively mobile or can be relocated with assistance (i.e. somebody can move an item from location A to location B).
i-SCOPE integrates open source technologies and existing services and applications developed by the project’s partners. This creates a comprehensive toolkit promoting interoperability through the use of OGC and other open standards for data exchange and services. In turn this allows the independent development and deployment of functionality provided by different web-services. Pre-existing technologies that were not available as individual services will be incorporated in the project’s service layer to ensure compliance with the overall schema.

5. The privacy and security dimension

Smart Cities are environments in which people live, learn, work and play. How they do that and how they interact with the Smart City allows the Smart City to learn about them. This implies that a Smart City can build up very detail models of the behaviour of citizens both as generalised communities and as individuals. The success of the Smart City then depends on the development of mutually supportive (symbiotic?) relationships between the city and the citizen. A citizen looks to the city to provide a huge range of services, and has a reasonable expectation that in order to receive those services some personal data has to be shared with the city. For example in receiving education for children the education authorities in the city need to know the age, sex, religion and academic attainment (maybe even academic desires) of every child in the city. In giving this data to the city the citizen also has a reasonable expectation that the city will not divulge this knowledge to anyone who asks. The protection of such explicitly sought and delivered data can be protected by policies of consent and fair use, and by ring-fencing the storage. However a Smart City gathers significant volumes of data that may be linked to individuals when data mining tools and similar are applied to it. This PII revealing data is not obvious and may not be picked up when the Smart City’s data designers

conduct their Privacy Impact Assessments. Furthermore ring-fencing data is not a reasonable practice either as it leads to massive duplication, and this is almost always going to lead to inconsistencies.
Therefore, a Smart City has to have a data policy that determines how data is shared between different use cases.
The protection model in i-SCOPE and i-locate and in each of SUNSHINE and UNCAP (the projects supporting and referred to in this paper), in common with other evolving web services, is policy based validation of security and privacy. In this model each interacting service exchanges its policy with regard to privacy and security attributes (e.g. identity of users, integrity and confidentiality of data) and when each party agrees them they create a privacy and security contract that binds their sharing of data. Any future interaction is then validated as complying to the contract allowing run time privacy and security protection to be given across the system.

6. Conclusion

3-D Urban Information Models provide an extremely powerful mechanism to deliver the vision of Smart Cities. The iSCOPE project is creating an infrastructure that will demonstrate how these models can integrate spatial data from a wide variety of sources to create tools for analysis, visualisation and simulation. As far as possible the project is integrating services and applications already developed by partners in the project. However, some further work is needed to ensure that gaps are filled and the objectives of both iSCOPE and Smart Cities are fully realised.
It is extremely important to reflect on the implications of the project for safety, security and privacy of the citizens of Smart Cities. Data can be synthesised and analysed in ways that cannot be anticipated by the data creators and developers of the system. Therefore it is imperative that a Smart City has a comprehensive data policy to determine how the privacy of citizens can be safeguarded.

7. Acknowledgements

The author acknowledges support from a number of sources in preparing this paper, where in large part the European Community through the FP7 and H2020 initiatives has been instrumental in providing this support. The author is however solely responsible for the content of this paper and asserts that it does not represent the opinion of the Community and that the Community is not responsible for any use that might be made of information contained therein.

i-SCOPE

The project has received funding from the European Community, and it has been co-funded by the CIP-ICT Policy Support Programme as part of the Competitiveness and innovation Framework Programme by the European Community (http://ec.europa.eu/ict_psp), contract number 297284.

SUNSHINE

This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community under grant number 325161

i-locate

The project has received funding from the European Community under contract number 62104029

UNCAP

This project has received funding from the European Union's Horizon 2020 research and innovation programme - Societal Challenge 1 (DG CONNECT/H) under grant agreement No 643555

References

- Eurostat. *urban or intermediate regions,* from http://ec.europa.eu/eurostat/web/main/home.
- i-locate from www.i-locate.eu.
- i-SCOPE from http://www.iscopeproject.net/.
- Killinger, B. (2010). *Integrity: Doing the Right Thing for the Right Reason*. McGill-Queen's University Press. p. 12. ISBN 9780773582804.
- SUNSHINE from http://www.sunshineproject.eu/.
- UNCAP from www.uncap.eu.

Applying Archaeological and Forensic Science methods and experience to outdoor and indoor mapping

Alexander Cadzow[1], Scott Cadzow[2]

[1] University of Bournemouth

[2] Cadzow Communications Consulting Ltd

Abstract

The science of archaeology, and forensic analysis methods, are intrinsically geo-centric but have not been viewed as “geo” or as mapping applications even though they require the recording of the locations of items and features. The purpose of this paper is to show the role that both forensic method and archaeology can play in the development of geomapping for recording of data at crime scenes, at archaeological sites and similar. By synergising geo-knowledge with existing forensic methods we can build a richer archaeological dataset for future analysis.

Keywords

Design, standardisation, forensic method, archaeological, indoor mapping, outdoor mapping

1 Introduction

The application of archaeological and forensic science methods and experience to outdoor and indoor mapping would rely on taking existing methods and methodologies and simply reapplying them to another area of work. Surveys carried out for archaeology research have yielded important information about key sites. For example, at Stonehenge researchers used a single multi-receiver electromagnetic induction (EMI) survey conducted over a 22 hectare area within the Stonehenge landscape the results of which shed light on the genesis of the landscape, and allowed for a better definition of potential paleo-environmental and archaeological locations. (Smedt P et al, 2014). This allowed greater understanding of the historic landscape for future research and conservation planning. This example can be expanded as the rationale to show that the techniques and methodology of archaeological survey have applications beyond that relatively narrowly defined domain. The discussed applications are not intended to improve positioning knowledge over the traditional means of technology which offer greater accuracy and positioning, rather they offer the means to give additional data regarding geo-locations and in adding that knowledge in the form of data to be used as a set of guidance markers in further position and navigation applications.

The techniques that are most commonly used for surveys in archaeology and forensics are ground penetrating radar (GPR), global positioning satellites (GPS), magnetometers, and earth resistivity. They allow features and sites to be “seen” and recorded when in the ground or on the surface.

GPR and GPS can be used in urban environments effectively to record features and see beneath the ground. Researchers and investigators already use GPR for urban surveys in cemeteries for example to work out where there is new or available space for new graves. GPR can give researchers and investigators the ability to “see” through modern urban construction materials such as tarmac and concrete in order to detect and to utilise features such as pipes and lines. While in a rural environment the methods GPR, GPS, Earth Resistivity and Magnetometer can all be used with few restrictions as we move into urban areas and urban green space magnetometers become progressively less useful as uncontrolled metal objects and materials (e.g. cars) interfere with the instrument.

2 Review of Current Practises

2.1 Archaeology

2.1.1 GPS

GPS surveys have been used in archaeology almost as long as commercial sets have been available. The GPS data collected in the UK are to Ordnance Survey standard. This allows the dimensions of the features to be recorded. More points equal higher resolution of detail to be collected about the features. Points can be collected using an automated roaming setup which allows a large area to be covered quickly but may lead to less detail of features being collected or missed all together. It can be used indoors as well as outdoors though depending on the location collecting GPS data does depend on satellite signal strength and the number of satellites that can be picked up. Though with the coming of the Galileo satellite network the issue of signal strength and satellite coverage should become less of a concern. The powerful ability of GPS surveys to record large sites was shown in the Chacas Valley Project in Peru. They focused on gathering both archaeological data and topographical information as necessary. The kinematic GPS method they applied gave the needed accuracy in representing the geometric properties of the objects and to preserve the geometric properties between the structures and to document the relations with the ground morphological aspects of the site. The data collected allowed the space relation between geographic and urbanized landscapes to be realized. (Capra A et al, 2002) Furthermore, data generated from GPS surveys can be immediately transformed into thematic maps useful for integration into a geographical information system (GIS) application which is part of the smart city architecture. Part of the development of smart city data sets wills most likely rely on digitizing traditional paper records and bring them into the GIS application and merging them with the any new data that is collected or created. The site of the ancient city of Throikos in Southern Greece through excavations had led to the creation of many detailed hand drawn maps of the different insulae. But no high quality overview map existed, only a generalized version. With GPS they carried a topographic survey of the area then digitizing the old maps they brought the data together into a GIS and proceed to match and merged the cartographic data together. (Lieffering et al, 2011) This allowed a better overview of the ancient city to be observed. Any smart “type” project will have to rely on old and new data sets being brought together allow people to have access to the full amount of information that is available. Instead of having to use both new smart application and a legacy application to have the full amount of information available. The discussed method would be the easiest to bring into indoor and outdoor mapping as it is already being used as a navigation method and this extant applications offers a way to turn what is seems to be purely a way of navigating into a way to collect geographic data.

2.1.2 GPR

The concept of radar is familiar to most people. In this instance, the radar signal – an electromagnetic pulse – is directed into the ground. Subsurface objects and stratigraphy (layering) will cause reflections that are picked up by a receiver. The travel time of the reflected signal indicates the depth. Data may be plotted as profiles, or as plan view maps isolating specific depths. This makes GPR a non–destructive method of subterranean mapping.

GPR can be a powerful tool in favourable conditions (uniform sandy soils are ideal). It is unique both in its ability to detect some spatially small objects at relatively great depths and in its ability to distinguish the depth of anomaly sources. The principal disadvantage of GPR is that it is severely limited by less-than-ideal conditions. The high electrical conductivity of fine-grained sediments (clays and silts) causes conductive losses of signal strength; rocky or heterogeneous sediments scatter the GPR signal and data collection is relatively slow.

The reflected echoes that allow precision of features to be picked up varies by quality of the equipment. When GPR was first used it was planned to be the gold standard but poor results and mixed success of detecting features led to it being side lined in the UK. There was better success with GPR in the USA, mainly because the drier conditions of the mid-west gave far better results in detecting features, thus spurring development of better sets. Studies in the USA show the details they can detect. Surveys using multi frequency antennas were conducted along an 875m section of the Lollie Levee near Conway, Arkansas, USA, to assess the levee's structural integrity. Many subsurface animal burrows, water-filled cavities, clay clasts, and metallic objects were investigated and identified. These anomalies were located at different depths and have different sizes. To validate the observations, hand dug trenches were excavated to confirm several anomalies with the results showing matches between GPR interpreted anomalies and the observed features. (Chlaib et al 2014) In the UK the urban use of GPR has been most prevalent in surveying graveyards, because graveyards and cemeteries globally are increasingly designated as full, there is a growing need to identify unmarked burial positions to find burial space or exhume and re-inter if necessary. GPR can reveal unmarked burials, extra/missing individuals from parish records and a variety of burial styles from isolated, brick-lined, to vertically stacked individuals. (Hansen et al, 2014). The ability to keep up to date records of space aligns itself nicely to the ethos of Smart infrastructure planning. Also GPR surveys can be conducted inside buildings to reveal features within and beneath the building which can aid in tracking and mapping the locations of utility features.

2.1.3 Magnetometers

Magnetometers used in geophysical survey use a single or two spatially separated sensors to measure the gradient of the magnetic field (the difference between the sensors). In most archaeological applications the gradiometer configuration is preferred, where one sensor is high up and the second is below

it near to the ground, because it provides better resolution of small, near-surface phenomena. Magnetometers may also use a variety of different sensor types giving varying degree of resolution detail. Every kind of material has unique magnetic properties, even those that we do not think of as being "magnetic." Different materials below the ground can cause local disturbances in the Earth's magnetic field that are detectable with sensitive magnetometers. Magnetometers react very strongly to iron and steel, brick, burned soil, and many types of rock, thus archaeological features composed of these materials are very detectable. Where these highly magnetic materials do not occur, it is often possible to detect very subtle anomalies caused by disturbed soils or decayed organic materials for example ditches, pits and sometime post hole. The chief limitation of magnetometer survey is that subtle features of interest may be obscured by highly magnetic geologic or modern materials. They can be used in fields beyond the archaeological as shown in 1993, during the removal of a diesel and a gasoline underground storage tank at the municipal garage in Wisconsin, soil testing revealed environmental contamination at the site. A site investigation revealed the possibility of a second on-site source of petroleum contamination. Limited historical data and the present usage of structures within the suspected source area precluded the use of most invasive sampling methods and most geophysical techniques. A fluxgate magnetometer survey, followed by confirmatory excavation, was conducted at the site. The fluxgate magnetometer survey identified nine possible magnetic anomalies within the 18 × 25 m area. The subsequent excavation near the anomalies revealed the presence of five paired and two individual 2000 litre underground storage tanks. (Biersel T.P et al, 2002)
This geophysical survey technique is a considered a standard method of collecting information about features hidden beneath the ground. It is a quick method of surveying a site and offers good resolution when it reveals features it reveals beneath the ground. Motorised mobile systems have offered the ability to uncover past landscapes covering many square miles to be revealed, for example dried river beds and lost villages can be revealed which would not appear in small one man surveys. This can be achieved with the use of a "Superconducting quantum interference device" (SQUID). The system provides mapping of large areas while allowing high resolution as well as lateral precision to be obtained. The properties of the SQUID were tested intensively at the large Neolithic double-ring ditch enclosure of Niederzimmern near Weimar, Germany. It revealed the wider historic landscape and details high previous had not been observed from smaller surveys. (Schultze et al, 2008) While the focus is on detecting historic features it will show modern features such as, pipes and electrical wires which would be usual for utility maps though these features are generally ignored on archaeology surveys or if included are marked to prevent there damage when a site is being excavated.

2.1.4 Earth Resistivity (E.R)

Is the oldest technical geophysical method and could be considered the simplest. It involves creating a current in the ground and allows travelling between probes and then measuring the resistance. The changes in resistance allow features to be identified. Using Ohm's Law, the resistivity of a material is derived using current and voltage values as measured by the electrodes. Differences in the amount of resistance are in response to the differing conductive properties of the sample body. This is determined by the amount of interstitial water and various salts held within. For example high resistance is indicative of stone features as the current has further to travel while low resistance is indicative of moist features such as ditches and large pits. The technique will also pick up geological features and modern man made features in the ground. However if there is little or too much moisture in the ground will prevent clear measurements from being taken. This enables archaeologists to non–invasively ‘read’ what is below the surface, either as a two-dimensional plan or, more recently, as a three-dimensional profile or section of the stratigraphy. (Mol and Preston, 2010) Due to the reliance on being able to create a current through the ground there are optimal times of the year to carry out surveys depending on soil types and land use. Generally surveys now use combined techniques which allows for corresponding data to be compared and correlated. This was used in a geophysical survey in Iznik/Nicaea, Turkey which led to the discovery of a Byzantine Church using GPR, E.R and microgravity. This allowed them to identify and differentiate the foundations, voids and the natural bed rock. (Rabbel et al, 2014) It is used to inform about the natures of site before any excavations are started to ensure there is no blind digging into the ground so only features that are shown on the survey results are targeted for further study.

2.2 Forensic Science

The application of these surveying methods to forensic science uses the same techniques and methodology but there is a greater detail in recording information to ensure the highest standards when the evidence is presented to the judicial and police service and to maintain integrity of the chain of evidence. This has led to search methodologies being established for specific geology’s for example detection of clandestine graves in coastal beach environments. (Pringle et al, 2012). The application in forensic science is mainly in the identification of human graves. Due to the nature of these methods they can also detect buried objects. This allows a detailed record of sites to be created and that affords the opportunity to build intelligence maps of locations and crimes. Though only for major or serious crime are these geophysical techniques brought in to be utilized. Any standards that are created for forensic science can be easily adapted to archaeology which means in turn there should be any standards to indoor and outdoor mapping standards.

3 Evolution to Geo

The uses and applications of these techniques can be applied to civic planning in both urban and rural areas with relative ease since they are ready used in a limited way already. It can be taken that they have the potential to have an expanded use. One of the issues that limit greater use is the problem of information dissemination of their potential application. Since companies and intuitions are not obliged to make their data open source. Encouraging the data to be shared in an open way would aid in potential speeding up the introduction to greater areas of use. These methods and techniques are a form of intelligence gathering with the raw data processed to sought the needs of the specialized area that they are being applied to so to apply these methods to indoor and outdoor mapping should cause too much of a problem.

The application of the techniques to Indoor GML would be to provide data either for utilitarian or provide a record of location of features. The methods that could most easily adapted be used are GPS and possibly GPR. Though GPR it would be most likely to ensure continued coverage between blocks of ground as GPR can penetrate most modern materials.
Surveys aimed at populating a City GML record can then rely mainly on GPS and GPR because they are the two methods which can effectively penetrate through modern building materials, though in open green space there is a possible use of E.R and Magnetometer but if there too many cars or the equipment is used too close to buildings they could affect the results that could be gained.
Any rural version of GML if implemented could use all methods with fewer problems though results would vary due to localized soil and geological conditions.

4 Conclusion

Already widely used this paper has illustrated a few cases where with careful re-use the existing geo-data gathering techniques in use in archaeology and forensic science can augment and enhance the data models using in smart cities and general mapping.
The benefits of these techniques is that trained personnel exist who are trained in their use although often in domains such as archaeology. In applying these techniques to other geo-domains such as civic planning and smart city initiatives the data models can be expanded quickly and build a much richer geo-data resource than if such sources were ignored. Furthermore the alignment of these data sources could come about in a short span of time if the resources were made available.
The methods and techniques described here offer a means to bring additional data resource to integrate into navigation and mapping with a view to maximising the data and information value of smart cities.
While these techniques and methodologies can be applied to smart city it is considered that maintaining the lines as independent, with archaeology, forensic science and geo-planning forging their own paths is counterproductive. The examples illustrate that convergence is possible and the assertion of this paper is that if convergence is seized upon then all domains will benefits.

5 Acknowledgements

The work reflected in this paper has been supported by the; i-locate project, and by each of the SUNSHINE and i-scope projects and has incorporated some of the findings from the i-tour project. The authors are solely responsible for the content and assert that the content does not represent the opinion of the Community and that the Community is not responsible for any use that might be made of information contained therein.

5.1 i-SCOPE

The project has received funding from the European Community, and it has been co-funded by the CIP-ICT Policy Support Programme as part of the Competitiveness and Innovation Framework Programme by the European Community, contract number 297284.

5.2 SUNSHINE

This project is partially funded under the ICT Policy Support Programme (ICT PSP) as part of the Competitiveness and Innovation Framework Programme by the European Community.

5.3 i-locate

The project has received funding from the European Community under contract number 621040.

References

✔ Bowman, M., Debray, S. K., and Peterson, L. L. 1993. Reasoning about naming systems. ACM Trans. Program. Lang. Syst. 15, 5 (Nov. 1993), 795-825. DOI= http://doi.acm.org/10.1145/161468.16147.

✔ Biersel, T. P, Bristoll, C, Taylor, R. W, Rose, J. 2002, Abandoned Underground Storage Tank Location Using Fluxgate Magnetic Surveying: A Case Study. Ground Water Monitoring & Remediation. Feb2002, Vol. 22 Issue 1, p116-120. 5P

✔ Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE. In Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology (Vancouver, Canada, November 02 - 05, 2003). UIST '03. ACM, New York, NY, 1-10. DOI= http://doi.acm.org/10.1145/964696.964697.

✔ Capra A, Gandolfi S Laurencich L, Mancini F, Minelli A, Orsini C, Rodríguez A, 2002, Multidisciplinary approach for archeological survey: exploring GPS method in landscape archeology studies, Journal of Cultural Heritage 3 93–99

✔ Chlaib H K, Mahdi H, Al-Shukri H, Su M M, Catakli A, Abd N, 2014, Using ground penetrating radar in levee assessment to detect small scale animal burrows. Journal of Applied Geophysics. Apr2014, Vol. 103, p121-131. 11p.

✔ Ding, W. and Marchionini, G. 1997. A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

✔ Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

✔ Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 01 - 06, 2000). CHI '00. ACM, New York, NY, 526-531. DOI= http://doi.acm.org/10.1145/332040.332491.

✔ Hansen J.D, Pringle J.K, Goodwin J, 2014, GPR and bulk ground resistivity surveys in graveyards: Locating unmarked burials in contrasting soil types, Forensic Science International , Vol. 237, pe14-e29. 16p.

✔ Killam, E. W. 2004. The Detection of Human Remains. English. Electronic books. 2nd ed. Springfield: Charles C Thomas Publisher. DOI= http://capitadiscovery.co.uk/bournemouth-ac/items/753468

✔ Liefferinge K.V, Stal C, Wulf A.D, 2011, The Thorikos Excavations 1963 – 2010 in Maps, Thorikos 10 Reports and Studies,5-14

✔ Mol L and Preston P.R, The Writing`s in the Wall: A review of new Preliminary Applications of Electrical Resistivity Tomography within Archaeology, Archaeometry 52, 6 1079–1095.

- ✔ Pringle J.K, Holland C, Szkornik K, Harrison M, 2012, Establishing forensic search methodologies and geophysical surveying for detection of clandestine graves in coastal beach environments, Forensic Science International 219 e29–e36.
- ✔ Rabbel W, Erkul E, Stumpel H, Wunderlich T, Pasteka R, Papco J, Niewohner P, Baris S, Cakin O and Peksen E, 2014, Discovery of a Byzantine Church in Iznik/Nicaea, Turkey: an Educational Case History of Geophysical Prospecting with Combined Methods in Urban Areas, Archaeological Prospection DOI:10.1002/arp.1491.
- ✔ Sannella, M. J. 1994. Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.
- ✔ Schultze V, Linzen S, Schuler T, Chwala A, Stolz R, Schulz M, Meyer H-G, 2008, Rapid and Sensitive Magnetometer Surveys of Large Areas using SQUID – the Measurement System and its Application to the Niederzimmern Neolithic Double-ring Ditch Exploration, Archaeological Prospection 15, 113-131.
- ✔ Smedt P.D, Mervenne M.V, Saey T, Baldwin E, Gaffney C, Gaffney V, 2014, Unveiling the prehistoric landscape at Stonehenge through multi-receiver EMI, Journal of Archaeological Science 50 16-23.
- ✔ Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender, Ed. ACM Press Frontier Series. ACM, New York, NY, 19-33. DOI= http://doi.acm.org/10.1145/90417.90738.
- ✔ Tavel, P. 2007. Modeling and Simulation Design. AK Peters Ltd., Natick, MA.
- ✔ Yu, Y. T. and Lau, M. F. 2006. A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions. J. Syst. Softw. 79, 5 (May. 2006), 577-590. DOI= http://dx.doi.org/10.1016/j.jss.2005.05.030.

Contextual processing for pedestrian tracking in GPS-denied environments

Enrico de Marinis, Michele Uliana, Fabrizio Pucci, Fabio Andreucci

DUNE srl, via Tracia 4, Roma

Abstract

This paper introduces the ConTeXtual Processing (CTXP), a novel and powerful concept for pedestrian tracking in GPS-denied environments. Its major advantages are: no need of external, ad-hoc infrastructures, no need of floorplan, low cost/weight/size, no need of calibrations or fingerprinting measurement campaigns, accuracy independent of the walked distance. In addition, CTXP processing is light enough to be hosted in a pocket-size commercial smartphone/tablet. CTXP has been extensively tested by Italian and International Agencies and industries in a widespread ensemble of scenarios (e.g. battleships, vessels, large industrial plants, malls) with experiments durations up to 3 hours and walked distance up to 6 km, always providing end-to-end results compliant with the target requirements.

Keywords

MEMS, tracking, processing, inertial, IMU, geolocation

1 Introduction

Substantial efforts and resources have been steered in the past decade toward Inertial Navigation Systems (INSs) for human tracking and localization based on Inertial Measurement Units (IMUs) based on Micro Electro-Mechanical Systems (MEMS) technology (Foxlin, 2005), (Leppäkoski et al. 2013). The major attractive is that these devices might provide low-cost, low-power, miniaturized, lightweight and infrastructure-less solutions for the navigation in GPS-denied scenarios. However they suffer significant bias, noise, scale factors, temperature drifts and limited dynamic ranges, resulting into position deviation and magnification of the angular Abbe error. These drawbacks de-facto prevent the use of MEMS IMUs for a long-range localisation. As a consequence, it is not surprising that in the recent years a widespread ensemble of techniques have been proposed to improve the localization capabilities. Most of the techniques rely on the Pedestrian Dead Reckoning (PDR) (Foxlin, 2005), where the walking behaviour is exploited to reset the INS errors by adopting an ECKF (Extended Complementary Kalman Filter). Other approaches achieve better performance by exploiting the presence of ancillary sensors, such compass (Jimenez et al., 2010), or by visual-inertial odometry (Li & Mourikis, 2013). Also pre-existing and independent sources of information are exploited, such as RFID tags (Jiménez et al., 2012) or map-matching techniques (Kaiser, Khidera, Robertson, 2013). The recent trends jointly exploit multiple-sensors readings (e.g. compass, barometer, RFID tags) into UKF (Unscented Kalman Filter) structures (Romanovas et al., 2013).

Scrutinizing the current State of the Art (SoA), the better performing solutions are based on the exploitation of two major sources of information: intrinsic (human-centric) and extrinsic (infrastructure-centric) information. The former is relevant to the operator himself and is basically limited to the information coming from body-mounted sensors (e.g. IMUs, barometer, compass, laser scanners); the latter is the information:

a) extracted from external sources (e.g. radio beacons, UWB devices, 3G/4G base stations, WiFi hotspots) that can be either ad-hoc deployed or can be just sources of opportunity (or both);
b) a priori information (e.g. floorpan, map).

The major advantage of the intrinsic information is that no position-aiding infrastructures are needed (expensive to deploy and maintain), but this comes at the cost of poor performance, as the available body-mounted sensor technology is not yet mature enough to insure an acceptable accuracy in the long term (hours). Moreover, some sensors need a calibration phase before the operations to avoid excessive drifts and errors. Despite the plethora of calibration methods (Ilewicz & Nawrat, 2013), some MEMS-based IMUs and compasses also suffer a long-term obsolescence of the calibration, from a few months to even 1 week. This would imply a re-calibration performed on a regular basis: an unacceptable task from the end-user perspective.

The solely use of extrinsic information has weak points as well, as the infrastructure cost, deployment time and maintenance can hardly meet the CApital Expenditure / Operating Expense (CAPEX/OPEX) budget and the set-up time constraints.

The better performing systems employ intrinsic information augmented by extrinsic information (e.g. joint use of IMU body-mounted sensors and Ultra Wide Band (UWB) transmitters deployed in the environment). On one hand the accuracy boost can be significant, on the other hand the drawbacks and limitations due to the CAPEX/OPEX and deployment time of the infrastructure(s) can easily jeopardize the real adoption of such systems. In addition, not all the areas/buildings are allowed to provide a clearance to deploy a location-aiding infrastructures (for safety or privacy constraints).

The authors, starting from the results of the RESCUE, PATH-SAFE and EXPLORERS projects (co-funded by Italian research initiatives), beside the intrinsic and extrinsic domains, introduce a novel information domain: the contextual domain that has shown (more than 120 on-field experiments) unprecedented performances in terms of accuracy, cost, weight.

2 The Contextual Processing

2.1 The Contextual domain

The contextual domain is the information coming from the behavioural features extracted from the walked path. For instance, when a pedestrian, after some time, crosses again a point previously crossed, this generates a contextual information. Indeed, a typical PDR processing, because of mismatches and drifts, would estimate these points (actually matching in the ground truth) as separate points in a 3D space. The same happens by analysing the spatial pattern of specific elements of the walked path; for instance, when some stairs are walked upward and then, after some time, are walked downward, the two typical “stairs-shaped” patterns are estimated, because of the drifts and scale factors, as dislocated in space by a PDR processing.

In other words, the contextual information is related to points or elements of

the path crossed again after some time: each time a contextual point is detected, the DUNE proprietary ConTeXtual Processing (CTXP) software estimates the (many) mismatches and drift factors, providing a backward compensation of the estimated path and, at the same time, a superior compensation of the forthcoming track, still to be walked.
On a statistical base, the presence of contextual points has been verified in practically all the scenarios and situations analysed so far. In some scenarios the presence of contextual points is very high (e.g. ships and vessels); in other scenarios is pretty good (e.g. malls). Only in some very specific cases are hardly to happen (e.g. inspector visiting a nuclear plant), but in this special case the cooperating inspector purposely crosses again a few previously visited points, so to generate the contextual information.

2.2 The DUNE system

The whole system is made of simple elements: a foot-mounted MEMS-based IMU sensor unit (20 g, 30x30x25 mm) also hosting the PDR processing; the uncompensated, drift-affected, estimated path is transmitted via Bluetooth (BT) link to a commercial smartphone/tablet, where the DUNE proprietary Android App provides a first drift compensation based on barometer, compass and GPS data (if available and if reliable) as well as the logging of the detected context points (that can be automatic or manual, depending on the application). The uncompensated track and detected contextual points are then fed into the CTXP software, which provides a final drift removal.

2.2.1 System architecture

In Figure 1, the general overview of the system, along with the processing blocks, is presented. The architecture can be conceptually split into three major blocks: the Local Processing Block (LPB) and the Remote Post-processing Block (RPB) and the Contextual Processing (CTXP). The core of the former is an ECKF and its processing cycle is performed at the same sample rate of the IMU signals (typically from 100 Hz to 1 kHz); the RPB is based on a iterative drift compensation driven by the information coming from the ancillary sensors (e.g. compass, GPS, Altimeter), if available and reliable. The CTXP block is then in charge of providing the estimation of the drift factors coming from the contextual information.
The terms “local” and “remote” indicate the suggested, but not mandatory, physical location of the processing blocks: the data rate of raw IMU data (input of the LPB) ranges from 30 to 150 kbps/user; whereas the output of the LPB (that is the input of the RPB) is approx 0.1 kbps/user. This suggests that, in case of a radio link between LPB and RPB, is by far better to perform the LPB on the same device hosting the IMU, thus requiring a very limited bandwidth to transmit the necessary data to the device where the operations of the RPB are performed (e.g. a smartphone).
The Local Processing Block: it performs the basic (almost “standard”) processing to get the (drift-affected) walked path from the solely inertial data.
The sampled signals (temperature-compensated accelerometers and gyroscopes) from the foot-mounted IMU unit are first fed into a step-detection algorithm (Benzerrouk, 2014); the detected stance, swing and stride phases of the foot, along with the sampled signals, are the input of a 15 states ECKF (Foxlin, 2005) that performs the PDR and thus provides a first, rough estimation of the track.

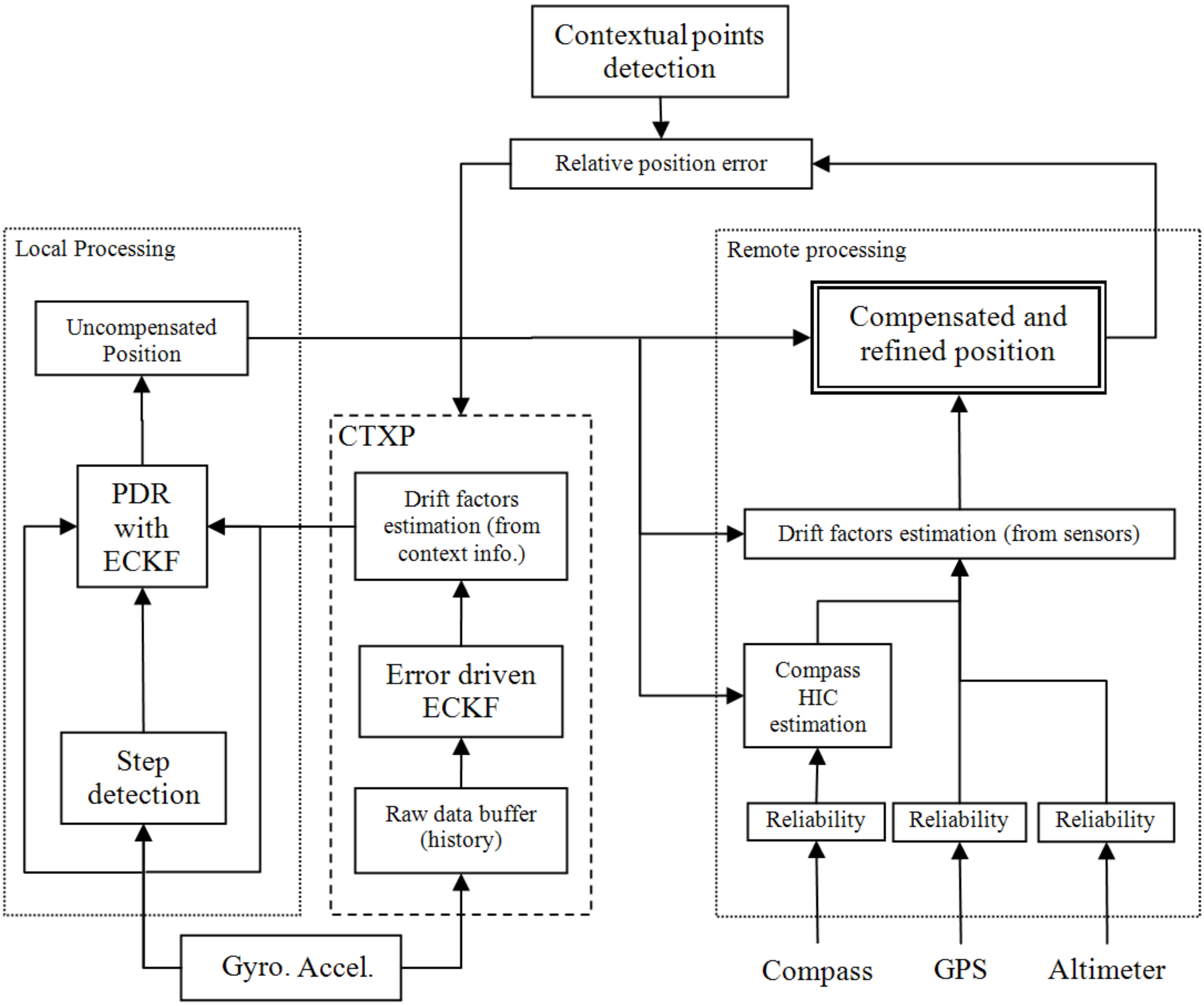


Figure 1: Overview of the complete system, with the relevant processing blocks.

The Remote Processing Block: the estimated track from the LPB is the major input to the RPB. The RPB is in charge of providing, if possible, a drift estimation and compensation by processing the information coming from the other body-mounted sensors (GPS, altimeter, compass). The data from these sensors are first fed into a reliability check process, providing a weighting metric about the quality of the sensed information: the lower a metric, the less that information is accounted for the drift estimation process. Employing a proprietary algorithm of DUNE, the compass data are also processed to provide a runtime Hard Iron Calibration (HIC) of the sensor itself. Finally, the multi-sensor fusion of the drift-affected track and the weighted information coming from the ancillary sensors provides a first level of drift compensation by comparing: a) the absolute attitude estimated by the LPB and the one provided by the compass; b) the differential horizontal position estimated by the LPB and the one coming from the GPS; c) the differential vertical position from the LPB and the same information given by the altimeter. At this stage, with no CTXP activated, the error between the estimated path and the ground truth is still linearly growing as 1% of the walked distance (estimate at the 95th centile).

The Contextual Processing Block: from the conceptual point of view, the

CTXP is "between" the LPB and the RPB, exploiting the information of both. Once a context point is detected, the analysis of the track of the RPB (embedding a first drift removal, if possible) gives the estimated error between the locations that should match in the ground truth, but are estimated as dislocated in the 3D space, because of the residual, still uncompensated drifts. This error information is fed into another ECKF that, based on the re-processing of the buffered past IMU data, provides the equivalent Recursive Least Squares (RLS) estimation of the drift factors that would minimise the estimated error between all the detected context points. As these drift factors come from a stronger information, they are employed to refine the already walked path (via backward processing) and supersede the factors previously estimated by the LPB, thus providing a better estimation of the path still to be walked.

3 Experimental results

The CTXP has undergone so far to more than 120 experiments and in all of them (purposely) no information coming from the GPS or the compass has been ever employed. The scenarios range from malls, large urban areas, industrial plants to battleships and vessels (in harbour and in navigation) and underground caves. The experiments have been carried out by different walkers, with or without still phases as long as 1 hour (e.g. emulating a meeting or a dinner) and with no need of calibrations or mandatory starting path shapes, with durations up to 3 hours and walked distance up to 6 km.

Most experiments have been performed in "double blind", as the path to be walked was not planned by the walker himself (who had no a priori knowledge of the planned path) and the operator in charge of the processing was not aware about the track actually walked.

The next Figure 2 and 3 show the results of a mixed indoor/outdoor experiment, (1h:49mins., 5.300 m), with 5 context points detected. Figure 2 represents the path estimated by a state of the art PDR algorithm, without GPS or compass information; the significant residual drift makes the estimated path useless after just a few hundred meters. Figure 3 represents the same path, but with the adoption of the contextual information; start-end error is 0,5 m for the 5,3 km path.

Figures 4 and 5 show details of the CTXP results of Figure 3. The former is the detail of an indoor/outdoor segment: in the multi-floors mall the GPS is totally obstructed and in the outdoor track is affected by significant errors (multipath). A similar situation is depicted in Figure 5, where 7 floors are walked up and down inside the building and the outdoor part is again affected by a significant multipath, leading to very poor GPS fixes.

As additional example, Figure 6 shows an experiment carried out in a multi-deck vessel (36 mins., 1.420 m, GPS absent, compass data always unreliable). Again, the path without CTXP is affected by a significant drift (start/end error: 34.5 m) that has been fully compensated by the CTXP (start/end error: 0.9 m).

Similar results have been obtained from all the experiments performed so far, from which it has been possible to estimate that the achievable 3D accuracy is largely independent of the walked distance, hardly exceeding 3-5 m or 1 m in the horizontal plane (without and with minimal presence of infrastructures, respectively) and 0.5 m in the vertical plane.

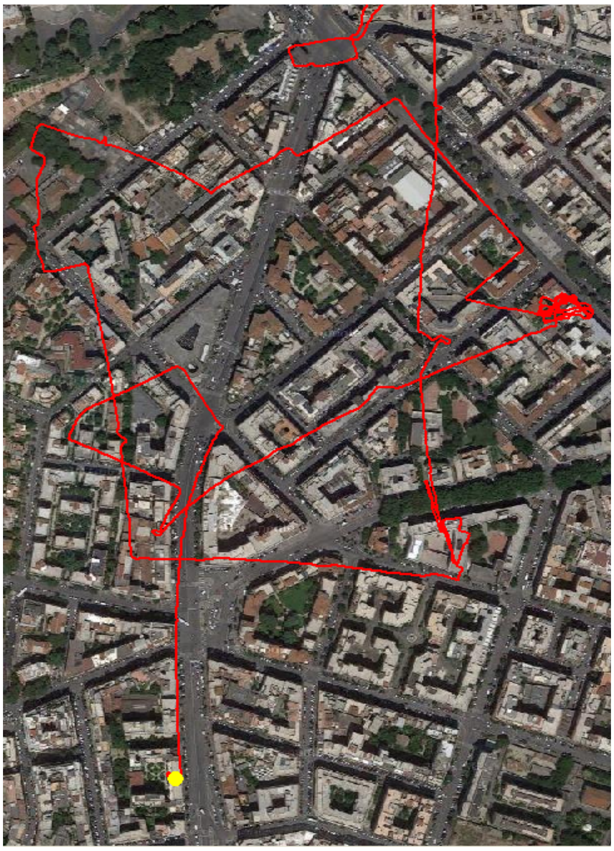


Figure 2: Path estimated by a PDR based on a 15 states ECKF.

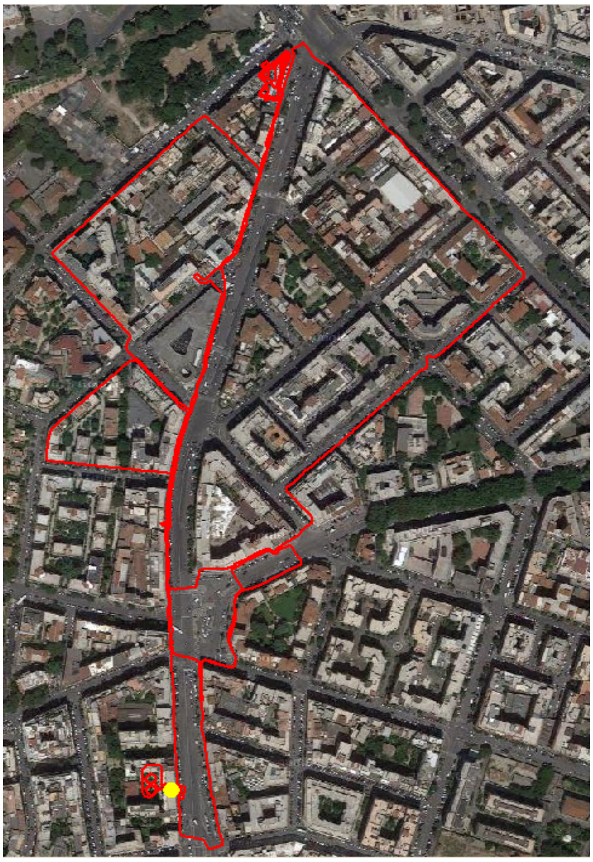


Figure 3: The same path as Figure 2, but processed with the CTXP.

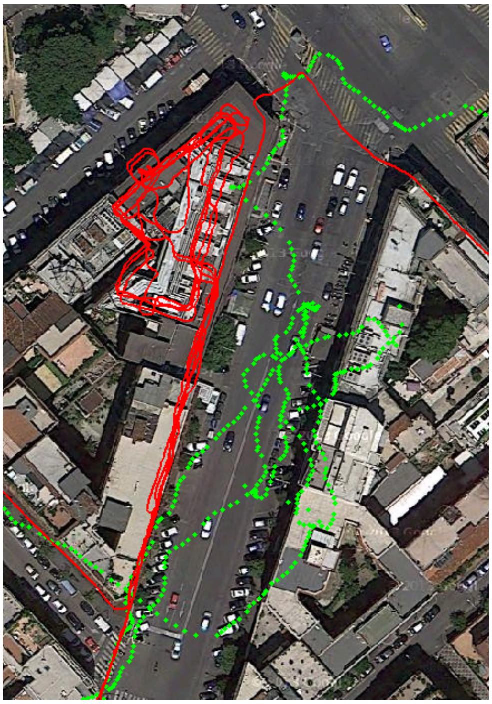


Figure 4: Detail of Figure 3: path walked indoor/outdoor. The green dots represent the GPS fixes.

Figure 5: Detail of Figure 3: path walked indoor/outdoor. The green dots represent the GPS fixes.

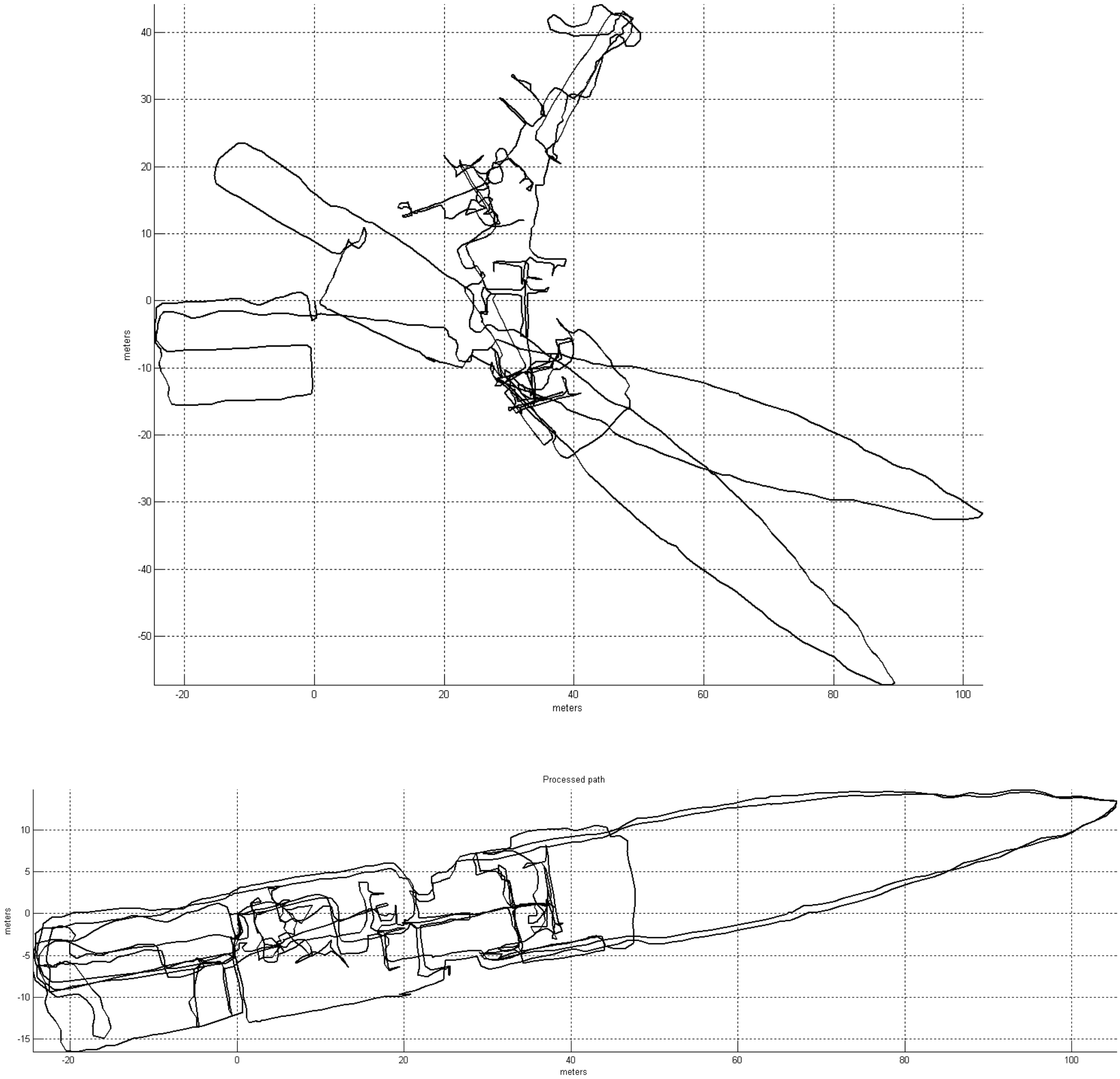


Figure 6: Experiment in a multi-deck vessel; upper plot: path estimated by a PDR based on 15 states ECKF; lower plot: path estimated by CTXP.

The following Figure 7 summarises the indicative accuracy with and without the CTXP block. The figures refer to the 95th centile, 1 hour walk in a dense urban environment (i.e. poor GPS performance and high magnetic pollution). More specifically:

- Without CTXP
  - In the absence of GPS, the processing provides an error below 1% of the walked distance (i.e. unbounded error);
  - When GPS is available, as a function of its reliability, the processing provides an error independent of the walked distance, ranging from 0.5 to 5 m.

- With CTXP

- without any external infrastructure, (thus requiring manual interaction on the operator’s side), the accuracy can be below 5 m, independently of the walked distance.
- When a minimal infrastructure (e.g. BT tags) is allowed, the CTXP works in a drop-and-forget mode (no interaction with the operator). In this configuration, the accuracy can be below 1 m, independently of the walked distance.

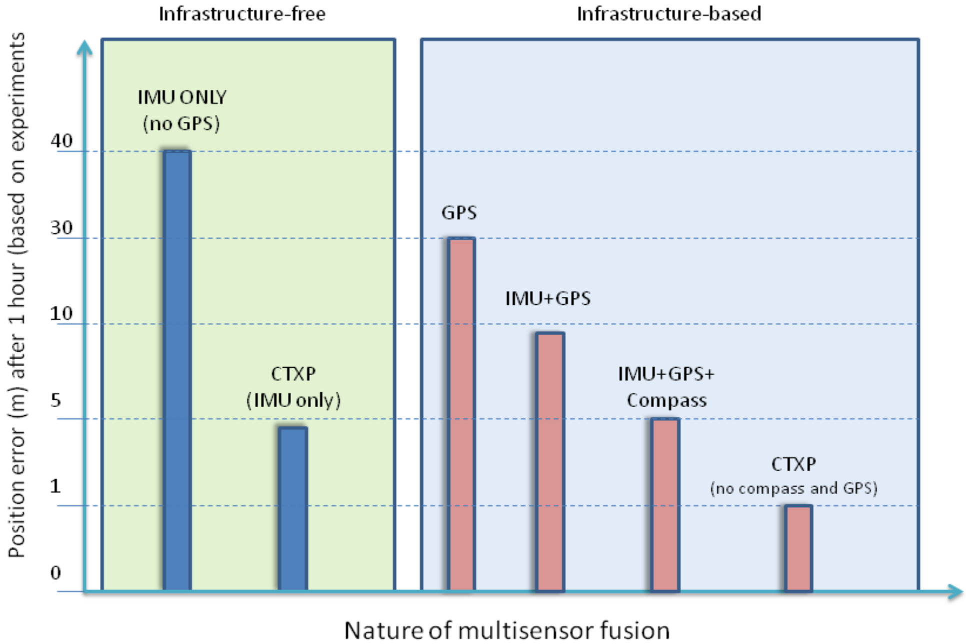


Figure 7: Indicative accuracy achieved after 1 hour walk in a dense urban environment (i.e. poor GPS performance and strong magnetic pollution).

The robustness of the CTX has been also tested against the accuracy of the detected context points (i.e. when a context point has an error w.r.t. the unknown ground truth). This performance has been estimated over an ensemble of 20 on-field experiments, each processed 10 times; a set of context points has been given a uniform random error ranging from 2 to 10 m (by adding the error in the processing software), whereas a different set has been geo-located and used as a benchmarking set. The results show that the maximal error on the benchmarking points never exceed the maximal error injected on the context points and that at the 95th centile the mean value of the errors on the benchmarking set approximately equals the mean value of the errors given on the mislocated context points.

4 Conclusions

The CTXP has been designed to cope with the major drawbacks coming from the infrastructure-free and infrastructure-based localisation approaches. It exploits the contextual (behavioural) information domain to achieve a localisation error independent of the walked distance (typical of the infrastructure-based systems) with no actual need of a pre-existing location-aiding infrastructure.

Tested so far in more than 120 on-field experiments (covering a widespread ensemble of different conditions), in all of them the CTXP has achieved results fully compliant with the target requirements, thus demonstrating to provide a real and interesting improvement w.r.t. the state of the art of the localisation systems in GPS-denied environments.

References

- ✔ Benzerrouk, H., Nebylov, A., Salhi, H., & Closas, P. (2014). MEMS IMU/ZUPT Based Cubature Kalman Filter applied to Pedestrian Navigation System. In Proceedings of International Electronic Conference on Sensors and Applications.
- ✔ Foxlin, E. (2005). Pedestrian tracking with shoe-mounted inertial sensors. IEEE Comput. Graph. Appl., 25(6), 38–46. doi:10.1109/MCG.2005.140
- ✔ Ilewicz W., Nawrat A., (2013). Direct Method of IMU Calibration. In Nawrat A., Simek K., Świerniak A. (Ed.) Springer Berlin Heidelberg. Advanced Technologies for Intelligent Systems of National Border Security. (pp 155-171), doi:10.1007/978-3-642-31665-4_13
- ✔ Jiménez A.R., Seco F., Prieto J.C., Guevara J. (2010), Indoor Pedestrian Navigation using an INS/EKF framework for Yaw Drift Reduction and a Foot-mounted IMU. WPNC 2010: 7th Workshop on Positioning, Navigation and Communication, 135-143. doi:10.1109/WPNC.2010.5649300
- ✔ Leppäkoski, H., J. Collin, J. Takala (2013). Pedestrian Navigation Based on Inertial Sensors, Indoor Map and WLAN Signals. Journal of Signal Processing Systems, 71(3), 287-296. doi:10.1007/s11265-012-0711-5
- ✔ Li, M., Mourikis A. I., (2013). High-precision, consistent EKF-based visual–inertial odometry. International Journal of Robotics Research, 32(6). doi: 10.1177/0278364913481251
- ✔ Kaiser S., Khidera M., Robertson P. (2013). A pedestrian navigation system using a map-based angular motion model for indoor and outdoor environments. Journal of Location Based Services, Special Issue: Indoor Positioning and Navigation. Part III: Navigation Systems, 7(1), 44-63. doi:10.1080/17489725.2012.698110
- ✔ Romanovas M., Goridko V., Klingbeil L., Bourouah M., Al-Jawad A., Traechtler M., Manoli Y. (2013). Pedestrian Indoor Localization Using Foot Mounted Inertial Sensors in Combination with a Magnetometer, a Barometer and RFID. In Jukka M. Krisp (Ed.) Springer Science & Business Media. Lecture Notes in Geoinformation and Cartography: Progress in Location based Services, (pp. 151-172).

Geotagging Software for media content based on EGNOS/GPS Location Based Service (LBS)

Angelos Anagnostopoulos[1], Anastasios Trypitsidis[1], Nikolaos Papadakis, Marc Bonazountas, Despina Kallidromitou

[1] EPSILON International SA | Monemvasias 27, 15125, Marousi, Greece | www.epsilon.gr

Abstract

This paper describes the design and implementation process adopted by Geotagging in c-Space project (http://www.c-spaceproject.eu/). This software module is utilizing different technologies, developed during two EU projects. Mainly:

- Tagging metadata to media content by c-Space project,
- GPS LBS module developed during i-locate project (http://www.i-locate.eu/),
- EGNOS module – Enhances GPS position by c-Space.

Basically, this module governs tagging video streaming or images with location information and informative metadata provided either automatically via a repository such as Wikipedia or by the users themselves.

Keywords

Geotagging; Location Based Services; EGNOS; GPS; Video stream; Images; c-Space; i-locate.

1 Introduction

One of the main objective of c-Space project is to benefit of nowadays trend - 14% of Internet users regularly upload and share media content (photos/videos) in a mobile scenarios – and deliver a low-cost solution to reconstruct 4D scenes (3D plus time) of indoor or outdoor events from casual real-world footage (e.g. captured by mobile phones) through Video Based Rendering (VBR) techniques.

In order to carry through the abovementioned objective many tasks have been defined. In our case we will focus on the development of a software module which will be used to tag media content (i.e. video stream or image) generated by end-user mobile application with meta-data. This meta-data includes spatio-temporal information and more “traditional” static values. It is either produced automatically (e.g. via smart-phone sensors and server-side logic) or is manually entered by the end-user.

The importance of this module is vital for the project because it intends to gather spatial information from mobile devices and store within a video. Afterwards, our intention is to create a standard, as it already stands for

photos.
From a literature standpoint, there are not many references about geotagging of videos – enhance videostreams with spatial information (lat, long, elevation, time). There are very few examples regarding geotaging of videos and in most cases they are just storing a single location as metadata for searching purposes (i.e. geo-referenced queries – range), as can be found (Seo, Cui & Zimmermann, 2012). There are also some interesting thoughs by the Safe Software Blog (blog.safe.com), where actually they are refering to our "problem"/goal . Specifically, Harper S. indicates the importance to have a standard which will store location information in each frame of a video. Specifically, "A standard which allows us to very easily tie together individual frames with location. On a more technical level, I want to be able to programmatically extract the latitude and longitude for any frame in the video. If we had that then it would allow us to do some amazing things" (Harper, 2012). In addition, there are also some commercial solutions where outcomes are unknown (i.e. Garmin, Go Pro) and semi-automatically procedures are required.
In the following paragraphs, data is categorized into types. For each media content type, the standards selected to describe it are explained. The technical process followed is described along with the resulting API. Finally, considerations are reported and future actions are enumerated.

2 Metadata categories

It was determined that meta-data falls under two major categories: Spatio-temporal; Static. Furthermore, in terms of the different media content types generated by the end-user - mobile application, two more categories arise: Image meta-data; Video meta-data.
All four combinations of these two classifications are valid and were considered when designing the technical solutions.

2.1 Spatio-temporal

From project's perspective, spatio-temporal metadata is the most important as it contains all meta-data describing the media content location in space and time. This information is critical for the project since the 4D (3D models + time) content reconstruction is relying on it.
Taking the above into consideration we determine that at minimum the following must be stored in a media file:

1. Spatial meta-data (at the time of shooting)
    a. Location: The camera location, expressed using World Geodetic System coordinates, including elevation (measured in meters).
    b. Bearing: The camera bearing (where it's pointing to), expressed in degrees.
    c. Speed (optional): The camera speed, expressed in km/h.
2. Temporal meta-data
    a. Time: The camera time, measured in millisecond accuracy and expressed in Greenwich Mean Time (GMT/UTC). Special considerations regarding the time are detailed in the following document section.

Temporal Considerations:
During the design of the software module, we came across with specific requirements from the 4D reconstruction engine. Specifically, the media content will come from different devices (smartphone, tablets, etc.), which will definitely not have fully synchronized internal clocks between them. This immediately creates a problem when trying to synchronize media on the server and especially in the 4D reconstruction engine.
The issue was successfully faced by querying a centralized time server (European NTP Pool) and getting back both a common time, but more importantly, the time offset between the device's internal clock and the common time. That offset is then stored into the media, and taken into account during meta-data extraction.

2.2 Static

For this set of meta-data we refer to information that apparently does not change over time/space and characterizes and media content. Basically part of this meta-data is automatically added to any media file by the device whist another part can be manually added by the user.
Specifically, automatically added values include:

- Image content: File type (i.e. JPG in the c-Space scope); File size; Image resolution (width and length); Bits per sample; Colour space;
- Video content: File type (i.e. MP4 in the c-Space scope); File size; Frame resolution (width and length); Number of frames; Compression format (e.g. H.264)

Manually added values include

- Image content: Title/description; Author/creator/artist; Document name; Copyright; Maker note; Keywords; User comment
- Video content: Title/description; Author/creator/artist; Keywords; Album info; Copyright; Rating.

3 Standards-Compliance

3.1 EXIF Standard

Well known standards were used for storing meta-data in media content as for image content EXIF Standard, as can be found (Huggel, 2014). The Exchangeable image file format is an established, commonly used, standard that specifies meta-data for images and sound. It is based on tags, following a specific structure, borrowed from TIFF files.
Various tools exist for quick EXIF data inspection on any given file, such as the free ExifTool tool, available for Linux, Windows and MacOS.

3.2 ISO base media file format (ISO/IEC 14496-12)

The primary reason for selecting the ISO/IEC 14496-12 standard was its capability of containing time-based media file formats such as video. Specifically, it is able to specify the structure and use the ISO base media file format (ISO, 2012).
In addition, the file structure is object-oriented; and as a consequence “a file can be decomposed into constituent objects (or “boxes”) very simply, and the

structure of the objects inferred directly from their type". The standard is extended to cover the MP4 and 3GP container formats, used by Android video capture (and, consequently, in the c-Space context), based on International Organization for Standardization (2013).

4 technical Implementation

The technical implementation presents solutions used the standards and basically it consisted of two parts: the client – Android and the server side implementation.
For the Android side, research was performed in order to locate open-source software libraries and/or tools fit to cover the project's needs. The final result is an API style set of code (including usage example mini applications) which can be included in any third-party Android application, such as in our development case. For the server side, an HTTP REST web service was developed, responsible for meta-data extraction and/or enhancement after uploading the content to the server.

4.1 Metadatalibrary

In this subsection, the main Android classes and methods are described. The figure below illustrates a step by step tagging process activity, both for images and video.

Figure 1: Geo-tagging process activity diagram.

4.2 The TimestampClient class

This class is used to access the remote Europe NTP Pool (via the internet) and fetch any remote time related data. Mainly, it queries:

- the europe.pool.ntp.org time server and returns it's time in UTC format (dd/MM/yyyy HH:mm:ss.SSS z), in millisecond accuracy,
- Queries the 'europe.pool.ntp.org' time server and returns the offset (in milliseconds) between the local device internal clock and the NTP server.

4.3 The ImageTagger class

At first and for image tagging, we built into the Android SDK the ExifInterface (Android, 2015). Since, it provides access to a limited subset of the EXIF tag set; therefore, we build upon the Android Exif Extended library (Crungola, 2014). At the end an Android API was developed providing easy to use methods, via the ImageTagger class, for setting values to specific tags. The methods provided along with the parameters and their descriptions, include:

- setTitle: A text title that sets the image description.
- setAuthor: A text that sets the image author/creator.
- setLocation: Coordinates (in WGS84) which sets camera’s location during image capture.
- setBearing: Sets the camera bearing, in degrees.
- setSpeed: Sets the device's speed at the time of capture.
- setDateTimeCaptured: sets the date and time (in millisecond accuracy) when the image was captured. See section 'Temporal Considerations' for more details on how to set a proper value.
- setTimeOffset: The time offset, measured in milliseconds

4.4 The VideoTagger class

For the video content tagging we used the open-source mp4parser library, which allows full control over the MP4 container format, including retrieving/setting meta-data values (or boxes, as they are called in the ISO/IEC 14496-12 standard) (ISO, 2012; Annies, 2014). Build upon this library and choosing specific tags from the full set (pertaining project needs), the VideoTagger class contains the following methods:

- setTitle: Sets the video title/description.
- setAuthor: Sets the video author/creator.
- setKeywords: Sets an array of keywords related to the video
- setLocation: Sets a "general" location describing the entire video. Not suitable for obtaining tracking info. For this reason we are using also 3GPP technical specifications (3GPP, 2005)
- setTrackingData: Parses a GPX GPS/EGNOS tracking file and embeds it into the video.
- setTrackingData: Parses a GPX GPS/EGNOS tracking file and embeds it into the video. See the relevant section for a detailed explanation of GPS/EGNOS meta-data tracking.
- setTimeOffset: Sets the time offset between the Android device's internal clock and the online Europe NTP time server, as described in

section “Temporal Considerations”.

4.5 Tracker Library

The task of tagging video content is more complicated than an image. For images is simply a matter of adding a simple set of coordinates and date/time info at the time of capture.

For video content, however, the device position is very likely to change during capture (consider, for example, a UAV drone recording a concert). Positional tracking must, therefore, take place throughout the entire duration of the video recording, in order to provide an adequate quantity of information.

Since the MP4 video container (or any other container, for that matter) does not provide a built-in mechanism for recording & storing location tracking data (aside for a single 'location' set of coordinates, as described previously) a custom solution have been implemented.

In respect to the GPS/EGNOS tracking mechanism implemented, it was based on the open-source (licensed GPLv3) library Open GPS Tracker (Open GPS Tracker, 2013). This library offers an Android Service component able to track and store GPS data in real-time, including any device sensor data (such as speed and bearing). The service runs in the background, so any other Android activity can take place at the same time, in our case a recording video.

The stored data is saved in XML following the GPX Exchange Format Schema, a common GPS data format for software applications. GPX allows for easy reviewing of captured data, via a large number of available GPX viewing applications.

Open GPS Tracker captures GPS data using Android LocationProvider. Additionally, we integrated with the EGNOS module (wherever available and applicable) an Android MockLocation, that was implemented and registered with the LocationProvider.

The resulting GPX file can then be easily embedded into a video file, by using the suitable VideoTagger method, as described previously.

At the end, a specific, re-factored, version of the library was created, cSpace_Tracker, which can be included in any other Android project (along with its dependencies). An excerpt from a GPX file is as follows:

```
<metadata>
      <time>2014-10-09T08:05:32Z</time>
</metadata>
<trk>
      <name>Track 2014-10-09 11:05</name>
      <trkseg>
            <trkpt lat="37.979397447779775" lon="23.783715711906552" >
                  <ele>256.29998779296875</ele>
                  <time>2014-10-09T08:05:38Z</time>
                  <extensions>
                        <ogt10:accuracy>5.0</ogt10:accuracy>
                        <gpx10:course>42.0</gpx10:course>
                  </extensions>
            </trkpt>
            …......
      </trkseq>
</trk>
```

4.6 Workflow for third-party consumers

The steps below summarize the workflow for creating a video recording Android application. The application initiates the GPS/EGNOS tracking service when the user starts the application. The application requests the Europe NTP time offset and stores it along with the current local device timestamp. Then, the user through the application begins recording a video and after while stops recording. Thereafter, the application deactivates the GPS/EGNOS tracking service application and it exports the captured tracking data to a GPX file. At the end, the application embeds the GPC file to the captured video, along with any other metadata desired.

Results

The following figures illustrate a sample application which we developed for testing reasons. The first figure depicts the interface of the application. We can either start the tracking service using GPS or use the EGNOS. The second figure illustrates the status of the receiver after the mobile device has been connected to a receiver.

Finally, the last figure depicts the tracker result, namely the GPX file. The GPX file is visualized through a well-known application (GPX Viewer).

Figure 2: Tracker application.

Figure 3: Tracker application-2.

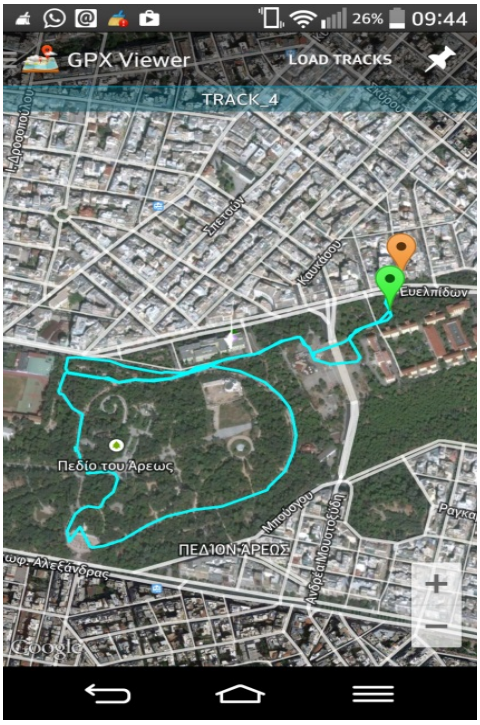


Figure 4: Tracker result - GPX file.

5 GPS & EGNOS Technical Specifications

As indicated the application supports both GPS and EGNOS as the reference location based service.

5.1 GPS APIs

With respect to the GPS LBS we developed a native Android library that retrieves the required location information from the Android OS.

Additionally, the system provided the functionality for a single location update, from the GPS (if available), or alternatively from the Network provider, and to notify the given instance. For this purpose, a locationListener “parameter” instance is used notifying on location updates. A LocationProvidersUnavailableException is thrown if neither one of the GPS/NETWORK nor other Location providers are available.

Below are the most important APIs:

- The getLocationByProvider endpoint is used to obtain last known location from a specific provider (network/gps). The location information is returned as a String object.
- The getLastKnownBestLocation endpoint is used to obtain the “best” last known location selected from all providers. A threshold in milliseconds is used (defaults to 5000 ms) to determine whether a Location is to be considered old. The endpoint returns a Location instance. The following assumptions have to be made: a location update is considered 'old' if its older than the configured; GPS is preferable than network location; If GPS data is old it cannot be considered reliable and the system will opt for network location; If both data are old then the system should return the newest of those two.

5.2 EGNOS APIs

We based the EGNOS integration on the EGNOS SDK. EGNOS SDK is a complex tool which basically comprises a plurality of APIs. That module is, in essence, a 'wrapper' around the EGNOS core source code, hiding any “unnecessary” EGNOS complexity from the module consumer, whist combining at the same time the features of the various EGNOS toolkits, as they become available.
In order for the SDK core algorithms to apply corrections to the EGNOS signals received from satellites or SISNeT, pseudorange/ephemeris GPS input data is required. At the moment, smartphone GPS chipsets (as well as the Android Location API) do not provide such information. It can only be obtained using an external GNSS chipset. Thereunder, basic API’s will be presented.

- The getReceiverType endpoint is used to provide the type of the receiver that the mobile is connected. This class returns the connected receiver type
- The getEDAS function informed if EDAS service is on or off. EDAS is the technical core of the EGNOS Commercial Data Distribution Service (CDDS) and provides the opportunity to deliver EGNOS data to users who cannot always view the EGNOS satellites (such as in urban canyons) or to support a variety of other value added services, applications and research programs. This class returns 0 or 1 based on EDAS ON or OFF from settings.
- The getPosition function computes the position as a table of 7 x 1 values. Rows 0 to 9 as GPS Latitude, GPS Longitude, GPS Altitude, EGNOS Latitude, EGNOS Longitude, EGNOS Altitude, HPL, R&D Latitude, R&D Longitude and R&D altitude respectively. The class returns the current position.
- The getinitialEGNOSPosition function computes the initial position acquired. The position in the Earth-Centered, Earth-Fixed (ECEF). Used as a first estimation to obtain a position. This class returns the initial EGNOS position.
- The getCurrentLocation function: sets mode of data source for EGNOS position through EGNOS Signal in Space or SISNeT; Checks for network in case data source is SISNeT; Gets the GPS Latitude, GPS Longitude, GPS Altitude; EGNOS Latitude, EGNOS Longitude, EGNOS Altitude and HPL from the EGNOS SDK. This class returns GPS and EGNOS position from EGNOS SDK.
- The checkNetwork function checks if the mobile device is connected to a network via 3G or Wi-Fi. This class returns “0” if no network is available, “1” if it is available.

6 Future steps

Next steps will include enhancing the meta-data library, to include additional meta-data tags/fields, as needs multiply. We are really interested to integrate also with the upcoming Galileo Open Service (OS), when it will be released. Meta-data extraction on the server side will be fully implemented. Finally, if the need arises, more sources will be added to the automatic data extraction.

References

- 3GPP TS 26.244 V6.4.0 , (2005), "Technical Specifications: 3rd Generation Partnership Project", Retrieved from: http://www.qtc.jp/3GPP/Specs/26244-640.pdf
- Android, (2015). ExifInterface", Retrieved from: http://developer.android.com/reference/android/media/ExifInterface.html
- Annies, S. (2014). "mp4parser on GitHub", Retrieved from: https://github.com/sannies/mp4parser
- Crungola, A. (2014). "Android Exif Extended on GitHub", Retrieved from: https://github.com/sephiroth74/Android-Exif-Extended
- Harper, S. (2012). "GPS Meets Video. Do We Need a Standard for Geotagging Videos?", Retrieved from: http://blog.safe.com/2012/04/gps-meets-video-do-we-need-a-standard-for-geotagging-videos/
- Huggel, A. (2014). "Exiv2 Image metadata library and tools", Retrieved from: http://exiv2.org/tags.html
- International Organization for Standardization, (2012), "Information technology -- Coding of audio-visual objects -- Part 12: ISO base media file format", Retrieved from: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61988
- International Organization for Standardization/IEC, (2013), "Sustainability of Digital Formats. Planning for Library of Congress Collections", Retrieved from: http://www.digitalpreservation.gov/formats/fdd/fdd000079.shtml
- Open GPS Tracker, (2013). "Tracker, Open GPS by Google Code", Retrieved from: https://code.google.com/p/open-gpstracker/
- Seo, B., Cui, W., & Zimmermann, R. (2012). "An experimental study of video uploading from mobile devices with HTTP streaming", Retrieved from: http://dl.acm.org/citation.cfm?doid=2155555.2155589

Open source web tool for tracking in a low cost mobile mapping system

Francesca Fissore[1], Andrea Masiero[1], Marco Piragnolo[1], Alberto Guarnieri[1], Antonio Vettore[1], Francesco Pirotti[1,2]

[1] Interdepartmental Research Center of Geomatics (CIRGEO), University of Padova, Viale dell'Università 16, Legnaro (PD) 35020, Italy

[2] Department of Land Environment Agriculture and Forestry (TESAF), University of Padova, Viale dell'Università 16, Legnaro (PD) 35020, Italy

Abstract

To support via web a mobile mapping system, open source libraries for spatial representation over web-based platforms can be of added value. In the following article we report our web tool which uses open source libraries to represent incoming data from the mobile mapping system and integrating all this data in a map baselayer and represent each part in graphs and plots. This enhances visualization and comprehension of the data in real time, and allows also to implement tools to distribute data to other stakeholders.

Keywords

Web-tool, mobile mapping, low-cost, open source, inertial navigation system

1 Introduction

A Mobile Mapping System (MMS) is a technology able to acquire in an efficient way three dimensional data able to describe with accuracy the surrounding environment (Guarnieri et al. 2008; Schwarz K. P., El-Sheimy N. 2004). Such systems are composed of an integrated array of time-synchronised navigation sensors and imaging sensors mounted on a mobile platform.

Today most of the MMSs available on the market have a "high cost" (Ellum and El-Sheimy 2002; Piras et al. 2008). Despite the fact this solutions are complete, very specific and particularly related to the typology of survey, their price is a factor that restricts the number of users and the possible interested sectors.

This paper will describe a new model of Mobile Mapping System, developed at the University of Padua (TESAF, Department of Land Environment Agriculture and Forestry) by the research team in CIRGEO. The objective of this paper is to report on the development of a prototype of MMS for the collection of geospatial data based on the assembly of low cost sensors managed through a web interface developed using open source libraries. The main goal is to build a system accessible by any type of user, and flexible to any type of upgrade or introduction of new models of sensors or versions thereof.

2 Description of the low cost MMS

The goal of our research is the development of a prototype of MMS composed by the assembling of low cost sensors managed through a web interface, designed specifically for our MMS using open source libraries. All the sensors, installed on a rigid platform, are kind of low cost. This choice is done to reduce the price of MMS making it accessible to a wider range of users, while the management of the system through web interface will facilitate its usage during the data acquisition, storage and synchronization. Furthermore, the use of open source software for the web interface dedicated to processing enables easier integration of new sensors in our MMS at later stages.

2.1 Hardware

Our proposed Mobile Mapping System is a modular measurement system consisting in two Logitech c920 webcams, low cost GNSS the RoyalTek MBT-1100, and the inertial navigation system 3DM-GX3-45 from Microstrain. It should be emphasized that the limiting factor in the number of sensors integrated in our MMS is linked to the hardware resources of our system and not those of the software.

3 Software

In order to collect and synchronize the data coming from navigation and imaging sensors, we have developed a tool based on a web platform that, using open source technology, is able to acquire, manage, aggregate and allow visualization thereof of real time data collected by sensors of MMS. Furthermore a historical track of the data can be logged if necessary.
The web platform is organized on a modular software architecture so that each sensor of the MMS is managed independently with respect to the others through dedicated modules. These collection of modules embody the back-end structure of the web tool and their main roles are summarized in the following keywords: connection, acquisition, interpretation and archiving.
In order to facilitate the interaction between the user and the MMS data and to increase the intuitive usage and access, we developed a web interface with clear panels for each part of the MMS information. The web interface represents the front end of our project and it is the means by which users manage and control in real time the data coming from the MMS.
The web platform schema is organized hierarchically with web pages accessible from the initial homepage. The homepage is designed using independent blocks, identified by different colors and showing the main parameters measured by the sensors (see Figure 1).
In order to georeference the position of the MMS, a continuously updated map keeps track of the distance covered during the survey mission of mobile mapping. Through the navigation menu the users have access to the web pages devoted to the single type of sensors where the values of all parameters acquired by the device are visible in real time. An example is shown in Figure 2.
Finally in order to facilitate data recovering and use in post processing as open data, we have created a special section of the web tool, accessible from the voice ‘Download’, where the user may get the measures recorded by the single

sensors in compressed text files.

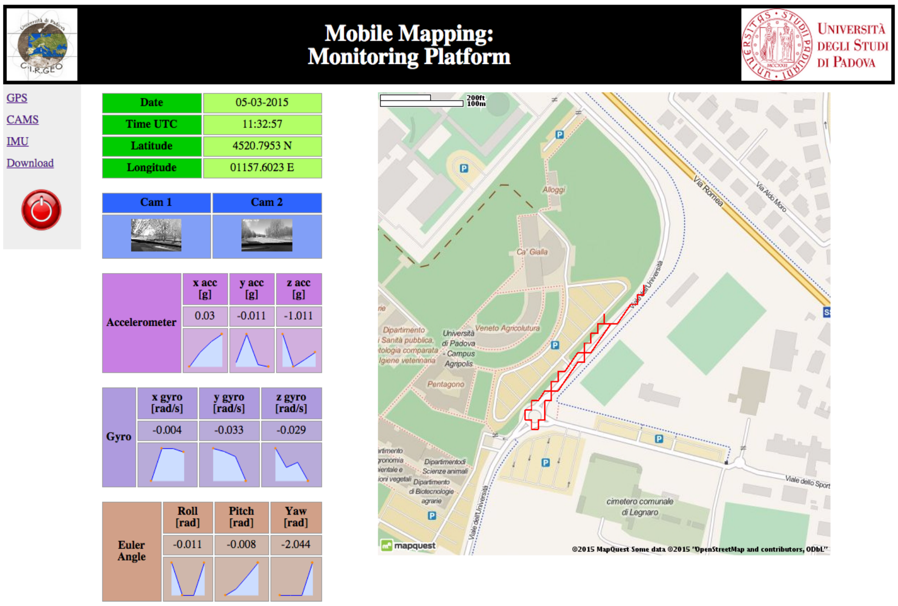


Figure 1: The main page of the Mobile Mapping web tool.

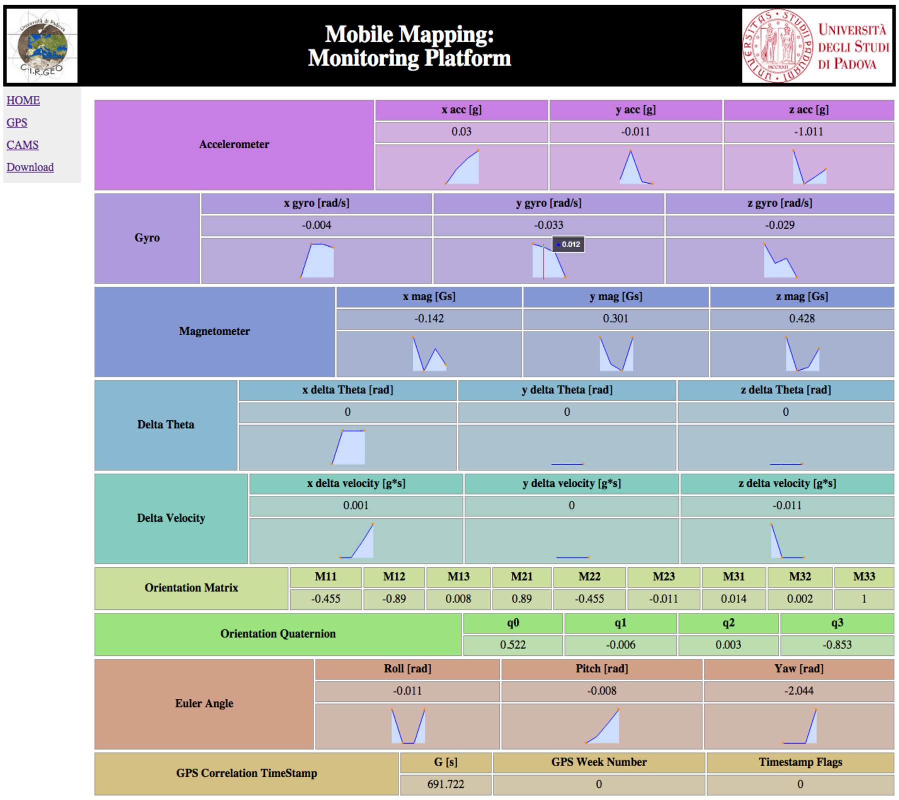


Figure 2: Web page dedicated to the inertial navigation system.

4 Future developments and conclusions

This paper presented the project conducted by the CIRGEO group at Padua University to create a low cost Mobile Mapping System and build a platform to manage its data. The goal of this project is to achieve low cost mobile mapping system leveraging where possible open source libraries for data management. The use of open source technologies and of a modular architecture have made our system highly flexible and open to the integration of other sensors and to be customized on the base of scientific area of application. Moreover a friendly web tool for its management, has opened potential usage of the data to others, starting a sharing principle and a collaborative potential (Pirotti et al. 2011).
Possible future developments of our project will focus on the integration of other sensors in our low cost MMS in order to differentiate the performance of our low cost system and to allow an accurate reconstruction of the environment observed during the survey mission in post processing phase. Moreover we will integrate other modules devoted to the calibration of the sensors in order to validate different methods of calibration and develop new ones. Finally the synchronization of the signals from sensors will be integrated inside the web tool and not carry out in post processing, as now it is doing.

References

- ✔ Ellum, Cameron, and Nasser El-Sheimy. “Land-Based Mobile Mapping Systems.” *Photogrammetric Engineering and Remote Sensing* 68, no. 1 (2002): 13–17.
- ✔ Guarnieri, A., Menin, A., Pirotti, F. Vettore, A., (2008). Low Cost System: GPS/MEMS for Positioning. In *Proceedings of FIG Working Week*, 14–19.
- ✔ Piras, M., Cina, A. Lingua. (2008). A. Low Cost Mobile Mapping Systems: An Italian Experience. *In Position, Location and Navigation Symposium*, 2008 IEEE/ION, 1033–45.
- ✔ Pirotti, F., Guarnieri, A., Vettore, A. (2011). Collaborative Web-GIS Design: A Case Study for Road Risk Analysis and Monitoring. *Transactions in GIS* 15, no. 2: 213–26
- ✔ Schwarz, K.P., El-Sheimy N.. (2004). Mobile Mapping Systems–state of the Art and Future Trends. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35, no. Part B: 10.

Cohesive routing service for indoor and outdoor navigation

Tao Feng, Joran Jessurun, Theo Arentze

Eindhoven University of Technology, The Netherlands

Abstract

Routing services plays a fundamental role in the development of navigation systems. For indoor routing specifically, it allows the navigation among multiple floors and generates turn-by-turn navigation guidance. In the paper we present the proposed routing system, which has been developed in the context of the European i-Locate project. We discuss the extensions required for seamless outdoor-indoor routing and an application to illustrate the new system.

Keywords

IndoorGML, Turn-by-turn Navigation, Multi-modal routing, Navigation Graph, Indoor Geocoder

1 Introduction

Routing services play a fundamental role in the development of navigation systems. With given start and end locations, these services calculate the optimal route. Similar to outdoor route planning and navigation which has been well developed, indoor navigation has great potential of market benefits regarding way finding in public buildings. However, the development of indoor navigation has been delayed for long because of the different limitations in techniques for indoor localization. Nowadays, with the emergence of new localization techniques, it becomes highly interesting for different stakeholders to have the indoor navigation service on board to better address desires and needs of people regarding indoor navigation.

The main barriers which limit the use of hybrid indoor and outdoor location-based service (LBS) are the i) lack of indoor maps available as open data, ii) lack of technological ecosystems that can use this data for innovative location, routing and asset management services within indoor & outdoor scenarios, and iii) limited support to indoor/outdoor LBS by current GI standards._iLocate consortium is such a team who tackles all these issues in a comprehensive manner. Within the scope of iLocate project, we developed the routing service for both outdoor and indoor environments based on OpenTripPlanner (OTP) an open source platform for multi-modal (outdoor) routing. The new routing service developed in the context of i-Locate works for both indoor and outdoor routing, with specific added value to indoor routing and indoor graph building. The routing service supports multimodal routing from door-to-door in a seamless way across outdoor and indoor environments. It allows the navigation among multiple floors by incorporating stairs, elevators, open spaces, etc., and generates turn-by-turn navigation guidance which is specific for indoor routing.

2 Indoor and outdoor routing service

The routing service is built upon the combination of different datasets, including the Open Street Map (OSM), General Transit Feed Specification (GTFS) and indoorGML. The OSM and GTFS compromise the bases of multimodal routing for outdoor environment, while the indoorGML serves the routing for indoor environment. An integrated graph combining the three types of data is built for the cohesive outdoor and indoor routing service.
Specific to the outdoor routing, the indoor routing service we developed incorporates the standard of indoorGML data defined by OSG. A structured metadata set which includes all possible elements in an indoor environment is pre-defined according to the requirement of navigation graphs and path-finding algorithms. A graph builder is made to let the OTP import the indoorGML data, combine multiple indoor/outdoor graphs and build the navigation graph for indoor /outdoor navigation. The graph builder reads a configuration file using Spring Framework, allowing the flexibility of setting various parameters. It also embeds a special functionality which connects to the online i-Locate portal and catches the needed information for the indoor graphs. In this way, the indoorGML files can be on the local disk or if a URL is used it can be downloaded from a server.

2.1 Navigation graphs

An integrated navigation graph is then built by connecting the outdoor and indoor networks through pre-defined anchor nodes (building entrances). Depending on the different attributes of the anchor nodes, multiple indoor/outdoor graphs are interconnected to obtain one single graph. We implement new vertices and edges for indoor graphs as an extension of known types for outdoor routing, and apply the existing routing algorithm to find routes in the combined outdoor-indoor network.

2.2 Indoor geocoder

As additional support of indoor routing, we also developed an indoor geocoding service. Extending the existing outdoor geocoder, the indoor geocoding facility uses an extended indoor address to find the unique indoor location in terms of building floor level and X, Y coordinates which can be used for indoor routing. The extended geocoder looks into the properties stored in the indoorGML file.

3.3 Compliance with OpenLS

The routing service receives input parameters from clients as a routing request, finds the optimal route and sends it back together corresponding turn-by-turn directions to clients as a response. The request parameters are input to the routing API which is bundled as a Java Servlet on the web server. The routing algorithm requires time and location information from clients to find the specific optimal route. The start and end locations are represented as longitude, latitude and floor identifier (for indoor navigation). In addition, the routing service also represents preference parameters and/or constraints one may have in reality. We present the proposed routing system, which has been developed in the context of the European i-Locate project. We discuss the

extensions required for seamless outdoor-indoor routing and an application to illustrate the new system.
A framework of the system for indoor and outdoor routing service is shown as below.

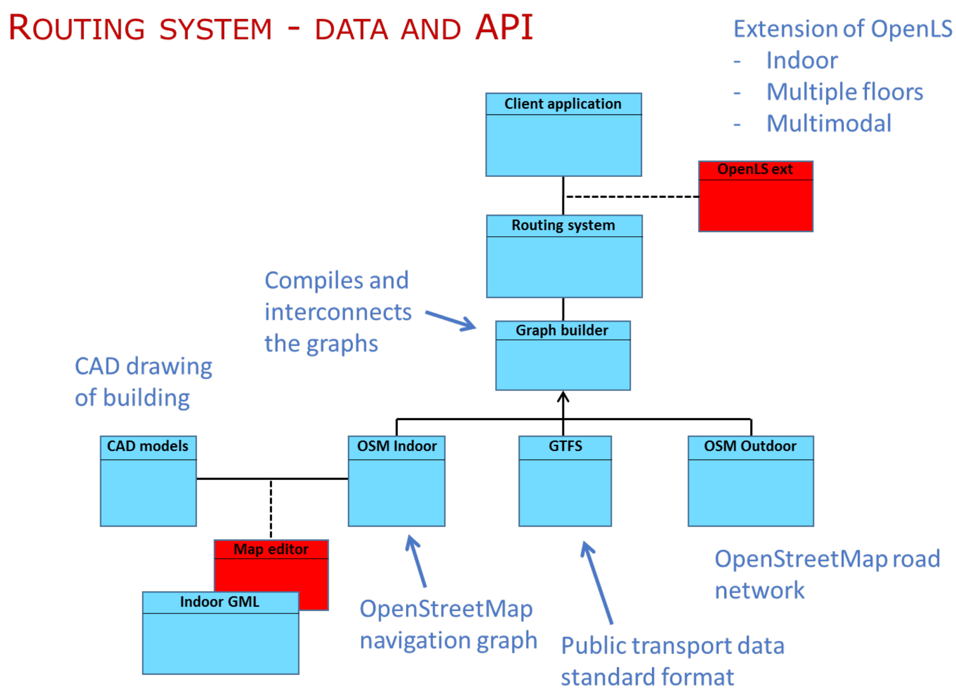


Figure 1: Framework of the indoor and outdoor routing service.

References

- ✔ OTP, OpenTripPlanner, an open source multi-modal trip planner: http://www.opentripplanner.org/
- ✔ OSM, OpenStreetMap: http://wiki.openstreetmap.org/wiki/Main_Page
- ✔ Deliverable: i-locate Toolkit v1.0 (Accompanying report) http://www.i-locate.eu/data/uploads/2015/01/D.3.1-i-Locate-toolkit-v1.0-rev-2015-01-201.pdf

Network real-time kinematic GNSS positioning assisted by tablets: the new frontier of the open source positioning and mapping

Paolo Dabove, Ambrogio Maria Manzino

Politecnico di Torino

Abstract

Kinematic satellite positioning consists in estimating the trajectory of a moving receiver, by processing observations from Global Navigation Satellite Systems (GNSS). This is a well-known techniques if geodetic or GIS receivers are considered but nowadays it is also possible to make a precise positioning with other devices, such as smartphones and tablets. This study shows the performances of an external mass-market GPS L1 receiver connected to a tablet in NRTK positioning for GIS purposes.

Keywords

Open source programs, GNSS positioning, RTKLIB, tablet, OpenStreetMap

1 Network real time positioning: concepts and principles

The Network Real-Time Kinematic (NRTK) positioning has got a great development in recent years, thanks to the appearance of GNSS CORSs (Continuous Operating Reference Stations) networks.

A network of permanent stations for real-time positioning is an infrastructure consisting of three parts: one part consists of all CORSs (more or less extended), with accurately known position, that transmit their data to a control center in real-time (Landau et al., 1995). The second part consists of a control center which receives and processes the data of the stations in real- time, fixes the phase ambiguities for all satellites of each permanent station and calculates the GNSS biases (e.g. ionosphere, tropospheric, clock delays). The third part is the set of network products that can be provided from the control center to the user: these products consist or in raw measurement file of each permanent station that the user may require for post processing purposes (Euler et al., 2001; Euler et al.,2003) or in streams (called differential corrections) that contain information related to the biases previous cited that allow the rover to make the real-time positioning (Manzino and Dabove, 2013). Moreover, the control centre is composed by a database that contains the calculated biases and the state of the network (i.e. the ambiguities, ionospheric and tropospheric delays for each station, …) for post-processing purposes, in order to be able to create the so called Virtual RINEX (Cina et al., 2014). Real time approach is instead traditionally considered when we desire to obtain a better quality, in terms of accuracy, of our positioning with respect the stand-alone position (i.e. with WAAS correction) or with a centimetric level of accuracy (Dabove and Manzino, 2014). The differential corrections which are generated by each GNSS permanent station are valid only in a limit area around the single site, considering a limited space: the main hypothesis is that if the bias remain almost the same in the base station and rover they can be eliminated by a process differential or relative (Fortes et al., 2000). In general the control center provides four types of network products, summarized as follows:

- Virtual Reference Station (VRS);
- Multi-Reference Station (MRS);
- Master Auxiliary (MAC or MAX)
- Flächen Korrektur Parameter (FKP)
- Nearest station (NRT).

While in the second, third and fourth case it is necessary to have a double frequency GNSS receiver as rover (Landau et al., 2011; Wübbena et al., 2002), in the first and last case it is possible to use also a single frequency receiver: this factor permits to perform an NRTK positioning also with mass-market receivers or with sensors that are able to track both the pseudorange and carrier phase measurements on L1 GPS frequency.

2 New instruments for GNSS positioning: smartphones and tablets

Nowadays, thanks to new technologies, the information about our position is available in almost every moment and almost everywhere thanks to smartphones, for example. These devices may include GPS/GNSS chipset that enable users to plan their activities (for example to know the time that it is necessary to wait a bus) or to share their location on social networks (e.g. Facebook). The accuracy of these sensors are quite bad: in some previous study (Bauer, 2013; Dabove 2014) was demonstrated that these internal sensors provide an accuracy greater than 5 m and it is not possible to extract any raw data from their internal sensors. This fact is the necessary requirement to obtain accuracies useful for GIS applications (about 10 cm). Moreover these devices are not equipped with external antenna connections, so it is not possible to perform a precise positioning also because it is unknown the measurement point. For these reasons we have decided to connect the considered device to an external GNSS mass-market instruments to obtain better results.

A Samsung Galaxy 10.4 tablet was considered in this paper, in order to show what are the performances obtainable if an external GPS instrument (u-blox LEA-7P) is connected.

In Figure 1 are shown both considered devices: we have considered this GPS receiver because is one of those that is able to provide in output the raw measurements described in the previous section. On the Samsung tablet, the RTKGPS+ app was installed: this free app, available at https://play.google.com/store/apps/details?id=gpsplus.rtkgps, can compute precise GPS positions using RTKLIB algorithms (RTK or PPP) and it supports external GPS with raw carrier phase output, such as for example the u-blox one. So, with these instruments is possible to perform an NRTK positioning obtaining accuracies of 5-10 cm and it is also possible to display the real-time position on some maps coming from IGN Geoportail (cadastral parcels, satellite, roads), OpenStreetMap Mapnik, Bing (satellite, roads).

Figure 1: Instruments used in this work.

3 Tests performed and future developments

Some preliminary tests were made: in order to compute a better positioning, an external mass-market antenna (Garmin GA-29F) was considered. These instruments are assembled to perform a trajectory outside the Politecnico di Torino, as it is possible to see in Figure 2 – right: this picture is captured directly by the RTKGPS+ software that is able to show the user's position of different cartographies, as described previously. In this case, all positions are shown on the OpenStreetMap map.

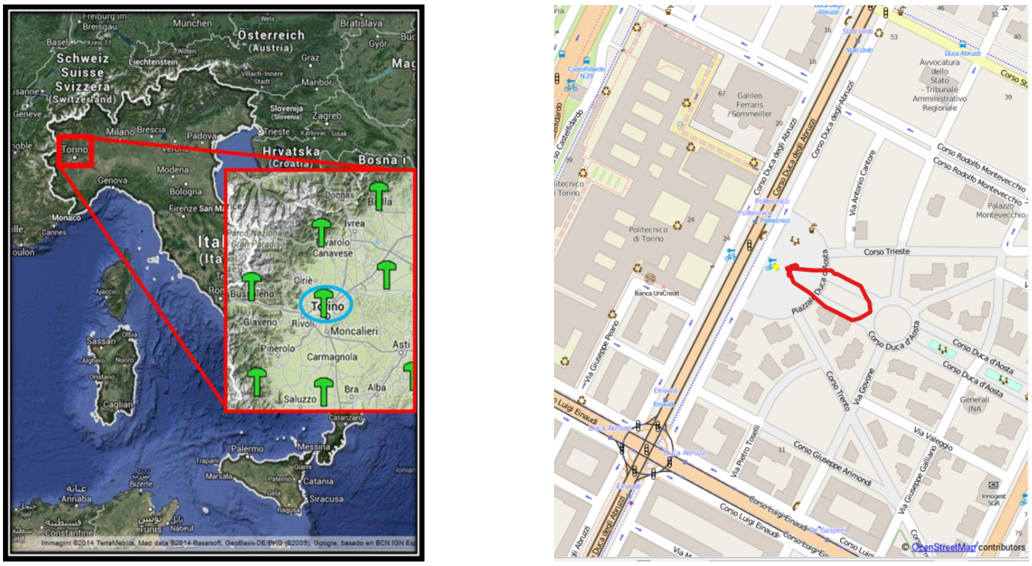


Figure 2: The Regione Piemonte CORSs network (left) and the trajectory performed by the rover (right).

For these real-time experiments, we considered the NRTK Regione Piemonte network shown in Figure 2 - left. The figure shows the area where the measurements were carried out. This network has a typical Italian mean inter-station distance (about 60 km) and can be used as a representative network in order to evaluate the performances of the receivers previously cited. The session length of this test is about 10 mins and the nearest CORS station was about 800m far from the rover: the results shown in this section consider epochs only with fixed phase ambiguity. Particular attention was paid also for the so called time-to-fix (TTF) that in this case was about 73 s: this means that a generic user must to wait less than 2 mins before to have an accuracy of about 10 cm in real time. The obtained accuracies are less than 15 cm (compared with well-known points measured with geodetic double frequency GNSS receiver) as shown in Figure 2, so these type of sensors are useful for many applications, such as navigation purposes and mapping.

Environment	Mean (m)			Standard Deviation (m)			Fixed Ambiguities
	E	N	V	E	N	V	
open area	-0.04	0.03	0.05	0.02	0.02	0.05	YES (98% of epochs)

Table 1: Results of performed test.

These interesting aspects will be investigate in the future, also considering a session length of 24 h in order to have more robust statistic parameters. Moreover some tests will be conducted considering the distance of the rover about 1, 5 and 10 km far from the nearest CORS in order to evaluate the efficiency of the network products for these devices.

References

- ✔ Bauer, C. (2013). On the (In-)Accuracy of GPS Measures of Smartphones: A Study of Running Tracking Applications, Proceedings of 11th International Conference on Advances in Mobile Computing & Multimedia (MoMM 2013), pp. 335-340
- ✔ Cina, A., Dabove, P., Manzino, A.M., Piras, M. (2014). Augmented Positioning with CORSs Network Services Using GNSS Mass-market Receivers. 2014 IEEE/ION Position, Location and Navigation Symposium (PLANS), 359-366
- ✔ Dabove, P. (2014). What are the actual performances of GNSS positioning using smartphone technology? Inside GNSS, 9 (6), 34-37.
- ✔ Dabove, P., Manzino, A.M. (2014). GPS & GLONASS Mass-Market Receivers: Positioning Performances and Peculiarities. Sensors, 14 (12), 22159-22179.
- ✔ Euler, H.-J., Keenan, C.R., Zebhauser, B.E., Wübbena, G. (2001). Study of a simplified approach in utilizing information from permanent reference station arrays. Proceedings ION GPS 2001, 379-391.
- ✔ Euler, H.J., Zelzer, O., Takac, F., Zebhauser, B.E. (2003). Applicability of Standardized Network RTK Messages for Surveying Rovers. Proceedings ION GPS/GNSS 2003.
- ✔ Fortes, L.P., Lachapelle, G., Cannon, M.E., Ryan, S., Marceau, G., Wee, S., Raquet, J. (2000). Use of a Multi-Reference GPS Station Network For Precise 3D Positioning in Constricted Waterways. International Hydrographic Review, 1 (1), 15-29.
- ✔ Landau, H., Euler, H.J. (1992). On-the-fly ambiguity resolution for precise differential positioning, Proceedings of the 5th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 1992), 607-613.
- ✔ Landau, H., Pagels, C., Vollath, U. (1995). Development and Implementation of a Centimeter-Accurate Real-Time-Kinematic Positioning System. Proceeding ION-GPS-1995.
- ✔ Landau, H., Vollath, U., Chen, X. (2002). Virtual Reference Station Systems. Journal of Global Positioning Systems, 1 (2), 137-143. Retrieved October 25, 2011, from http://www.gnss.com.au/JoGPS/v1n2/v1n2pH.pdf.
- ✔ Manzino, A.M., Dabove, P. (2013). Quality control of the NRTK positioning with mass-market receivers. Global Positioning Systems: Signal Structure, Applications and Sources of Error and Biases. Ya-Hui Hsueh, Hauppauge NY, U.S.A. , 17-40.

✔ Wübbena, G., Bagge, A. (2002). RTCM Message Type 59-FKP for trasmissine of FKP. Version 1. GEO++ White paper n. 2002.01. Retrieved from: http://www.geopp.de/download/geopp-rtcm-fkp59.pdf.

Geodetic monitoring experiment by low-cost GNSS receivers and goGPS positioning engine

Stefano Caldera[1], Eugenio Realini[1], Daisuke Yoshida[2]

[1] Geomatics Research & Development (GReD) srl, c/o ComoNExT, via Cavour 2, Lomazzo, Como, Italy

[2] Graduate School for Creative Cities, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka, Japan

Abstract

This paper describes an experiment of GNSS-based monitoring by means of different models and brands of low-cost receivers. Single-frequency observations were processed by goGPS software, in relative positioning mode with respect to a virtual reference station and three permanent stations. The experiment was carried out in Osaka, Japan; the results here described cover a timespan of two months. The daily solutions show standard deviations ranging from about 1 mm to 1 cm, depending on the distance of the base station and its observation quality.

Keywords

GNSS, low-cost receivers, monitoring, goGPS

1 Introduction

Monitoring the deformations of critical infrastructures, such as large buildings, bridges, dams or high voltage electricity towers, is important to guarantee both the safety of the population and the functioning of national/regional services. In this regard, GNSS-based monitoring provides advantages over other techniques, since it is continuous in time, it has a global coverage, and it does not require the work of operators directly on the site to be monitored. Traditional GNSS monitoring is performed by means of geodetic dual-frequency receivers; however, such receivers have high costs that make the widespread adoption of GNSS monitoring impractical. Recently, low-cost GNSS receivers are being experimented by the geodetic community also for high-precision static positioning, with good results (Buchli et al., 2012; Benoit et al., 2014).

2 Experiment setup

The GNSS monitoring equipment, composed of a low-cost patch GNSS antenna, three GNSS receivers and a splitter, was installed in the Media Center, Osaka City University, Japan. The Media Center is ten-story building, built in 1996, and it is one of the tallest building in the neighborhood (Figure 1, left), guaranteeing a good sky visibility for receiving GNSS satellite signals. The GNSS antenna, a

NV2410 by NVS Technologies AG, was placed on its rooftop (Figure 1, center); the low-cost GNSS receivers, namely a u-blox 4T, a u-blox 6T and an NVS NV08C-CSM, were connected to a 4-way splitter, which, in turn was connected to the antenna by a 30 meter long cable (Figure 1, right). The data logging was carried out by goGPS Java programs on a server machine, storing binary data and creating RINEX files on-the-fly.

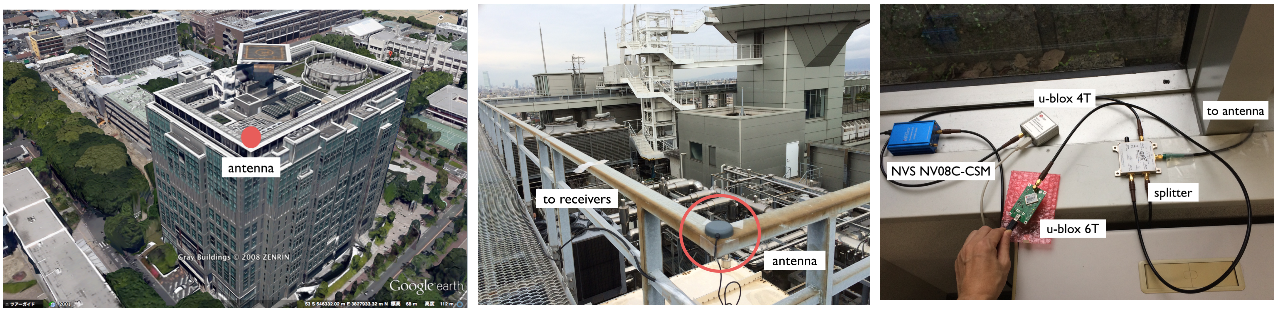


Figure 1: Antenna position on the rooftop of the Osaka City University Media Center building (left); NVS NV2410 GNSS antenna installed on the building rooftop (center); low-cost GNSS receivers and splitter (right).

3 Results

The results reported in this paper refer to a period of 60 days, from 25 December 2014 to 22 February 2015. Code and phase observations by each of the three low-cost receivers was post-processed by the relative positioning algorithm of goGPS, MATLAB version, which applies an extended Kalman filter (for this experiment, the dynamic model was set to "static") on double-differenced code and phase observations, estimating the antenna coordinates and float carrier phase ambiguities. The LAMBDA3 algorithm (Verhagen and Li, 2012) was applied epoch-by-epoch on the float ambuigity solutions. As base stations, a Virtual Reference Station (VRS), generated by the JENOBA positioning service, and three permanent stations belonging to GEONET (the nation-wide GNSS network of Japan) were used. The VRS was generated at a distance of about 60 m from the low-cost antenna. The three GEONET stations, with marker names 0336, 0337 and 0761 were located at about 9.5 km, 7.5 km and 11 km, respectively. In particular, station 0336 is located in a densely urbanized area of Osaka, and it is surrounded by tall buildings and nearby trees, degrading its observation quality. The other two GEONET stations have relatively similar surroundings, with low buildings and open sky visibility above the elevation cutoff used in this work (15 degrees). All the processing was done on GPS-only observations, although the NVS receiver and all the base stations were in fact logging also GLONASS data; GLONASS-only and GPS+GLONASS solutions are going to be investigated in future tests.

3.1 Comparison between low-cost receivers

Figure 2 shows the daily solutions obtained by processing the three low-cost receivers with respect to the VRS. The three solutions do not highlight significant differences in terms of repeatability, although the older u-blox model (the LEA-4T) shows a few more spurious oscillations.

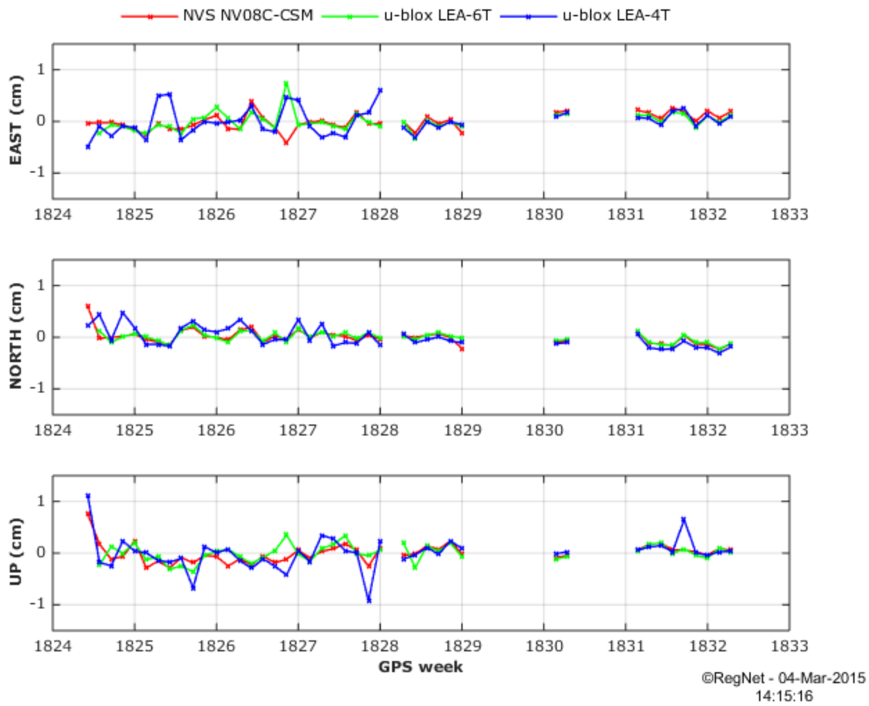


Figure 2: Daily solutions obtained by processing each low-cost receiver with respect to the VRS.

3.2 Comparison between base stations

Figure 3 shows the daily solutions obtained by processing the u-blox LEA-6T with respect to the VRS and the three GEONET stations. The GEONET station 0336 confirms to provide the worst repeatability, due to the aforementioned bad quality of its observations.

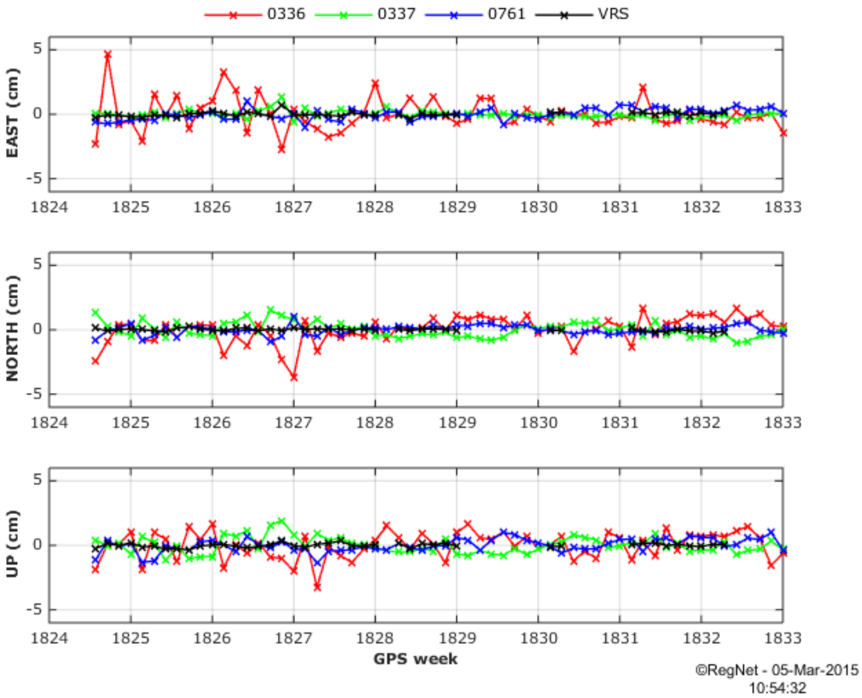


Figure 3: Daily solutions by processing the u-blox LEA-6T receiver with the four base stations.

4 Conclusions

The results confirm that the observation quality of low-cost GNSS receivers, even with a low-cost patch antenna, are sufficient to perform precise daily positioning by the double-difference method with respect to a base station. The precision level, ranging from 1 mm to 1 cm in standard deviation in this experiment (Table 1) is influenced by the distance of the base station and its observation quality. The rightmost column of Table 1 reports a stability index expressed as $SI=\sqrt{\sigma^2(E)+\sigma^2(N)+\sigma^2(U)}$.

Master station (approx. baseline length)	Low-cost receiver	St.Dev. (EAST) [mm]	St.Dev. (NORTH) [mm]	St.Dev. (UP) [mm]	SI [mm]
VRS (60 m)	u-blox 4T	2.5	1.5	3.0	4.2
	u-blox 6T	1.7	0.9	1.6	2.5
	NVS	1.4	1.2	1.8	2.6
0337 (7.5 km)	u-blox 4T	3.1	5.0	7.0	9.1
	u-blox 6T	2.7	5.4	6.3	8.7
	NVS	2.6	6.0	6.7	9.4
0336 (9.5 km)	u-blox 4T	12.9	8.1	12.3	19.6
	u-blox 6T	12.7	9.1	10.5	18.8
	NVS	9.4	9.3	11.7	17.7
0761 (11 km)	u-blox 4T	4.3	3.5	5.2	7.6
	u-blox 6T	3.5	3.5	4.7	6.8
	NVS	3.3	4.2	4.7	7.1

Table 1: solution repeatability for all the combinations of base stations and low-cost receivers.

Acknowledgments

The authors would like to thank the JENOBA Co., Ltd service for their assistance with the VRS tests. This work was supported by JSPS KAKENHI Grant Number 26730156.

References

✔ Benoit, L., Briole, P., Martin, O., & Thom, C. (2014). Real-time deformation monitoring by a wireless network of low-cost GPS. Journal of Applied Geodesy, 1-10.

✔ Buchli, B., Sutton, F., & Beutel, J. (2012). GPS-equipped wireless sensor network node for high-accuracy positioning applications. In Wireless Sensor Networks (pp. 179-195). Springer Berlin Heidelberg.

✔ Verhagen, S., Li, B. (2012). LAMBDA - Matlab implementation, version 3.0. Delft University of Technology and Curtin University.

Extending goGPS for Snapshot Positioning

Emanuele Ziglioli[1], Eugenio Realini[2]

[1] Sirtrack Ltd, 8A Goddard Lane, Havelock North 4130, New Zealand

[2] Geomatics Research & Development (GReD) Srl, c/o ComoNExT, via Cavour 2, 22074, Lomazzo, Como, Italy

Abstract

In order to estimate the coordinates of a GNSS device, navigation satellites transmit a precise reference clock and their orbit parameters. However, if the received signal is very weak or too short, traditional positioning algorithms cannot be applied. A technique known as Coarse Time Navigation or Snapshot Positioning addresses this problem by providing coarse time and ephemeris by means other than satellite transmission.

The goGPS library is an open source library in MATLAB and Java for precise positioning by various stand-alone and relative techniques. This paper presents how Coarse Time Navigation can be added to goGPS. Problems that are unique to this technique will be discussed, as well as some solutions.

Keywords

goGPS, snapshot GPS, assisted GPS, indoor positioning, wildlife tracking

1 Introduction

Indoor positioning and wildlife tracking are very different applications of satellite navigation, although they have at least one thing in common: they are cases where computing a location via traditional methods may be impossible.

For a Global Navigation Satellite System (GNSS) receiver to work, two sets of variables need to be known: a reference clock (in this work we will focus on the NAVSTAR Global Positioning System, thus we will refer to the GPS clock) and satellite orbit parameters.

The GPS clock is used to measure pseudoranges, based on the time it takes for the signal from each satellite to reach the receiver. Ephemeris are used to compute the precise location of satellites and that also requires a clock with a few millisecond precision. GPS satellites can move up to ±3.3 km/s in Earth-Centered, Earth-Fixed (ECEF) coordinates, corresponding to ±800m/s along the satellite-receiver line-of-sight (fig.1) so a clock error of a few seconds translates into a computed position error in the order of kilometres. Both sets of parameters are transmitted by satellites as part of the navigation message. In the case of GPS, the Time-of-Week (TOW) is transmitted every 6s. Ephemeris are transmitted every 30s, with each satellite transmitting their own. With no prior knowledge of ephemeris (a situation that is known as *cold start*), a GPS receiver has to synchronize with the incoming signal and decode it until ephemeris are received from a minimum of 4 satellites. That can take multiples of 30s, depending on the signal strength.

In the case of indoor positioning, the satellite signal may be so weak that the navigation message would be impossible to decode without errors. In the case of wildlife tracking, marine animals, such as penguins or turtles (an example of tracking device attached to the shell of a sea turtle is shown in fig. 2), often surface for less than a second, a much shorter period than what is required to decode GPS clock and ephemeris. Despite this, Snapshot Positioning is commonly used by providing the required information by means other than satellite communication.

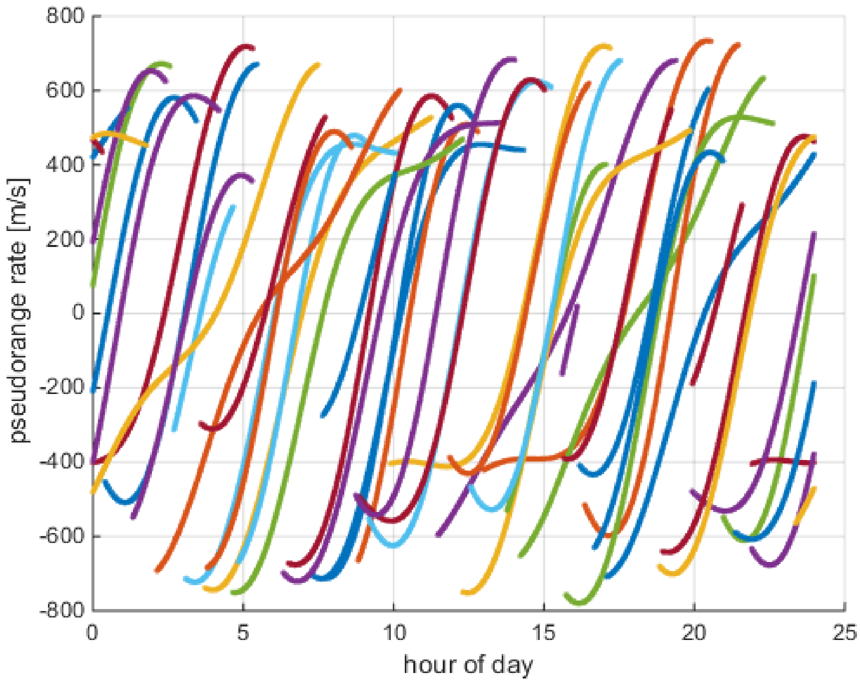


Figure 1: Pseudorange rate as captured by a fixed station in Japan (latitude: 30.55 degrees) over a period of 24h.

Real-time use cases of Snapshot Positioning fall into the category of *Assisted GPS*, for which coarse time and ephemeris are provided to the receiver, for example by the mobile network (Van Diggelen, 2009). In the case of wildlife tracking, where animals can be out of reach for several months, GPS snapshots can be stored or processed on board. Locations are computed in post-processing using archival data about satellite orbits.

In section 2 we will review the equation systems for satellite navigation, first for the traditional case then for the snapshot case. We will review in particular the region of convergence for a solution and the problem of identifying a suitable a priori location and time within that region.

Figure 2: Example of GNSS tracking device attached to a sea turtle shell.

In section 3 we will describe how we have modified the goGPS Java library in order to support coarse time navigation. Finally, results of a simulation of a snapshot series will be presented. We will compare the performance of a location computed from Snapshot GPS as opposed to standard GPS in terms of HDOP.

2 Equation Systems for Satellite Navigation

We will first describe the navigation system in a traditional sense followed by a description for Snapshot Positioning.

2.1 The 4-state system for traditional GPS

A minimum of 4 pseudoranges is required to solve a system containing 4 unknown variables, that are receiver position coordinates (*X*, *Y*, *Z*) and common bias (i.e. the combination of the receiver clock offset and other biases that all the observations have in common, such as antenna cable delays, etc.) multiplied by the speed of light (*b*):

$$\boldsymbol{x}=\begin{bmatrix} X \\ Y \\ Z \\ b \end{bmatrix}$$

The pseudorange residual, i.e. the difference between measured and a priori pseudorange, can be expressed as:

$$\delta\rho^{(k)}=\rho-\hat{\rho}=-\boldsymbol{e}^{(k)}\cdot\delta\boldsymbol{x}_{XYZ}+\delta\boldsymbol{x}_b+\epsilon^{(k)} \quad \text{for each satellite } (k)$$

where

$e^{(k)}$ is the line-of-sight vector between the a priori position and the satellite,

$\delta\boldsymbol{x}_{XYZ}$ is the update vector of the a priori position,

$\delta\boldsymbol{x}_b$ is the common bias (as a distance),

$\epsilon^{(k)}$ accounts for the observation error.

This 4-state system is commonly expressed in the matrix form:

$$\delta\boldsymbol{\rho}=\boldsymbol{H}\cdot\delta\boldsymbol{x}+\boldsymbol{\epsilon}$$

where, for n satellites, with $1 \leq k \leq n$

$$\boldsymbol{H}=\begin{bmatrix}-\boldsymbol{e}^{(1)} & 1\\ \vdots & \vdots\\ -\boldsymbol{e}^{(n)} & 1\end{bmatrix}$$

is the design matrix. The system requires a minimum of 4 measured pseudoranges and is solved for example using the least squares method:

$$\delta\boldsymbol{x}=(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\delta\rho$$

2.2 The 5-state system for Snapshot Positioning

In all cases of Snapshot Positioning, the GPS clock is a variable that needs to be solved for, in order to eliminate the need to decode the TOW (Sirola & Syrjärinne, 2002). The system of navigation equations now presents an additional unknown state variable, i.e. the system coarse time:

$$\boldsymbol{x}=\begin{bmatrix}X\\ Y\\ Z\\ b\\ t_c\end{bmatrix}$$

so that in principle a minimum of 5 pseudoranges needs to be measured.

An additional constraint linking satellite speed and system time update is introduced (Van Diggelen, 2009):

$$\dot{\rho}^{(k)}=\boldsymbol{e}^{(k)}\cdot\boldsymbol{v}^{(k)}$$

where

$\boldsymbol{v}^{(k)}$ is the satellite speed

$\dot{\rho}^{(k)}$ is the radial component of the satellite speed, i.e. the pseudorange rate

The pseudorange residual is now:

$$\delta\rho^{(k)}=-\boldsymbol{e}^{(k)}\cdot\delta\boldsymbol{x}_{XYZ}+\delta\boldsymbol{x}_b+\dot{\rho}^{(k)}\cdot\delta\boldsymbol{x}_{tc}+\epsilon^{(k)}$$ for each satellite (k)

where

$\delta\boldsymbol{x}_{tc}$ is the update of the a priori value of the system time.

And the design matrix is now:

$$\boldsymbol{H}=\begin{bmatrix}-\boldsymbol{e}^{(1)} & 1 & \dot{\rho}^{(1)}\\ \vdots & \vdots & \vdots\\ -\boldsymbol{e}^{(n)} & 1 & \dot{\rho}^{(n)}\end{bmatrix}$$

The pseudorange rate can be computed from measured Doppler:

$$\dot{\rho}^{(k)}=\frac{\Delta f}{f}c$$

where

c is the speed of light

Δf is the measured Doppler frequency

f is the carrier frequency

If the measured Doppler is not available, the satellite speed can be calculated from ephemeris (Remondi, 2004), although that would not take into account the speed of the receiver.

For simplicity of description we have not explicitly taken into account correction factors such as the known satellite clock biases, tropospheric and ionospheric errors and the rotation of the earth. Our implementation of Snapshot GPS though does include those corrections.

2.3 Convergence region for the Snapshot Positioning system and millisecond code ambiguity resolution

Because of the lack of synchronization with the GPS clock, a snapshot receiver can only measure satellite code phases, that are times of arrival within a 1ms period of the GPS coarse code signal. The full pseudorange $\rho^{(k)}$ is the sum of an unknown number of integer cycles $N^{(k)}$, a fractional pseudorange $z^{(k)}$ and the unknown common bias from the state vector:

$$\rho^{(k)}=N^{(k)}+z^{(k)}+b$$

The relationship between measured code phase α and fractional pseudorange $z^{(k)}$ is:

$$z^{(k)}=1-\alpha^{(k)}$$

The fractional nature of observed pseudoranges means that the a priori pseudorange $\hat{\rho}$ must be within 0.5ms (about 150km) from the true pseudorange:

$$\left|\delta\rho^{(k)}\right|=\left|\rho-\hat{\rho}\right|<150\,km$$

In a way that is reminiscent of carrier phase tracking: we will have to estimate $N^{(k)}$. A first step could simply round the difference between a priori pseudorange and observed fractional pseudorange:

$$N^{(k)}=round\left(\hat{\rho}-z^{(k)}\right)$$

As in carrier phase tracking, where carrier phase cycle slips occur and need to be monitored, 1ms code slips can also occur as an effect of the common bias and other biases.

It is worth pointing out that, as opposed to the carrier phase slip case, where a slip of 1 cycle introduces an observation error of about 20 cm (depending on the carrier wavelength), a 1ms code slip translates into an error of about 300km.

A couple of strategies have been suggested in order to detect a 1ms cycle slip.

The first strategy (Van Diggelen 2009) involves selecting one of the observed pseudoranges as a reference and then subtracting the others from the reference in order to eliminate the effect of the common bias.

The second strategy (Othieno 2012) estimates $N^{(k)}$ for all satellites, then it considers the difference between observed and a priori pseudorange for all satellites and uses the smallest one as a reference. All other differences should not be further than ±150km from the reference. If that occurs, a 1ms cycle slip has occurred and the value of 300km is added to or subtracted from that pseudorange. For this implementation of Snapshot Positioning, we have chosen the second strategy. It seemed both simple and robust, not requiring an arbitrary choice of reference satellite.

Once a solution is found we can verify that the a priori state was within the convergence zone to the true location by computing a posteriori residuals:

$$\delta\rho_{ps}^{(k)}=\left|\rho^{(k)}-\hat{\rho}^{(k)}\right|$$

In the case of an overdetermined system with 6 or more satellites, a true

solution will return a posteriori residuals in the order of meters while a solution outside the convergence zone is likely to produce a very large residual for at least one satellite (Van Diggelen 2009).
A strategy for increasing the system redundancy is to add an altitude soft-constraint (i.e. a pseudo-observation with its associated “observation error”). goGPS MATLAB provides a module that loads a Digital Terrain Model, interpolates it and uses the interpolated value as a pseudo-observation; this module could be adapted to goGPS Java and used for snapshot positioning as well.

2.4 Effects of extra state over Dilution of Precision

Dilution of Precision (DOP) is commonly used as an indicator of quality of GPS positioning. Its value is derived from the matrix G:

$$\boldsymbol{G}=(\boldsymbol{H}^{\boldsymbol{T}}\boldsymbol{H})^{-1}$$

The geometric component of G can be expressed in global (ECEF) or in local (East-North-Up, ENU) coordinates for standard GPS:

$$\boldsymbol{G}=\begin{bmatrix} g_X & \vdots & \vdots & \vdots \\ \vdots & g_Y & \vdots & \vdots \\ \vdots & \vdots & g_Z & \vdots \\ \vdots & \vdots & \vdots & g_b \end{bmatrix} \qquad \boldsymbol{G}_{\boldsymbol{ECEF}}=\begin{bmatrix} g_X & \vdots & \vdots \\ \vdots & g_Y & \vdots \\ \vdots & \vdots & g_Z \end{bmatrix} \qquad \boldsymbol{G}_{\boldsymbol{ENU}}=\begin{bmatrix} g_E & \vdots & \vdots \\ \vdots & g_N & \vdots \\ \vdots & \vdots & g_U \end{bmatrix}$$

These ECEF and ENU forms are linked by a transformation matrix (Strang G & Borre K, 1997). The values g_{XYZ} and g_{ENU} are affected by the geometry of the observations.
The GDOP and HDOP values in the 4-state case are calculated as:

$$GDOP_4=\sqrt{tr(\boldsymbol{G})}=\sqrt{g_E+g_N+g_U+g_b} \qquad HDOP_4=\sqrt{g_E+g_N}$$

We can see intuitively that in the 5-state case of Snapshot GPS, the GDOP will be at least as large as in the 4-state case:

$$GDOP_5=\sqrt{g_E+g_N+g_U+g_b+g_{tc}}\geqslant GDOP_4$$

The extra-state theorem (Van Diggelen, 2009) states that adding an extra state makes all DOPs greater than or equal to their original value. While this difference can be large with 5 satellites, it becomes smaller and smaller as the number of satellites increases. Interestingly, the difference in HDOP between the 5 and the 4 state case also decreases as the latitude increases, both towards the North Pole and the South Pole.

2.5 Doppler Based Initialization

The a priori position can be calculated from the measured Doppler. This allows for a position resolution where only satellite ephemeris and coarse time are needed.
For the static case, we're seeking to resolve a 4 state system, i.e. user position and receiver clock rate or frequency offset (Othieno, 2012):

$$\boldsymbol{x}=\begin{bmatrix} X \\ Y \\ Z \\ \dot{b} \end{bmatrix}$$

The system is

$$\boldsymbol{H}\boldsymbol{x}=\boldsymbol{B}$$

The design matrix is now:

$$\boldsymbol{H}=\begin{bmatrix}\boldsymbol{v}^{(1)}|\boldsymbol{s}^{(1)}|\\ \vdots \quad \vdots\\ \boldsymbol{v}^{(n)}|\boldsymbol{s}^{(n)}|\end{bmatrix}$$

where

$\boldsymbol{v}^{(k)}$ is the satellite speed

$\boldsymbol{s}^{(k)}$ is the satellite position

The ***B*** vector is

$$b^{(k)}=\boldsymbol{s}^{(k)}\cdot\boldsymbol{v}^{(k)}+\dot{\rho}^{(k)}\cdot|\hat{\boldsymbol{x}}-\boldsymbol{s}^{(k)}|+\hat{\dot{b}}\left(|\boldsymbol{s}^{(k)}|-|\hat{\boldsymbol{x}}-\boldsymbol{s}^{(k)}|\right)$$

where

$\dot{\rho}^{(k)}$ is the radial component of the satellite speed, i.e. the pseudorange rate

$\hat{\boldsymbol{x}}$ is the position as computed from the previous iteration

$\hat{\dot{b}}$ is the receiver clock rate from the previous iteration

For the first iteration we can initialize position and clock rate to zero. The static system can be generalized to the dynamic case. It then becomes a 7 state system with the addition of the receiver speed.

The performance of Doppler based positioning is much worse than in the case of GPS and is the order of km/Hz (Van Diggelen, 2009). However, even an error of a few km is sufficient to provide an a priori position within the region of convergence of Snapshot GPS system.

3 Extending goGPS

goGPS is a mature library that implements state of the art RTK algorithms in two languages, MATLAB and Java. For this paper though we have developed the GPS Snapshot algorithm exclusively in Java.

Despite the MATLAB version of goGPS being more advanced, the Java version includes most of the building blocks needed for the Snapshot GPS algorithm and it is reasonably efficient (Realini et al, 2010). Java is also an attractive choice because of the JVM, a very popular platform for web services, and of Android, the most popular mobile platform in the world.

Finally, the authors are aware of an open-source implementation of a Snapshot Algorithm in C++[1] but they are unaware of any open source implementation in Java.

3.1 The Snapshot Module

A key architectural characteristic of goGPS is its modularity (Realini, 2013). These are its observational modules:

Code Module. C/A code pseudoranges

Phase Module. Carrier phase tracking

DTM Module. Digital Terrain Model

In the case of Snapshot GPS, phase tracking is not available so the implementation is restricted to a Code Module, while we have left the use of a DTM module for future development.

The *codeStandalone()* method was chosen as a starting point for the snapshot

1 http://fastgps.sourceforge.net/

implementation, and these are the major changes:

- addition of an extra state: the coarse time
- computation of pseudorange rate (either from Doppler or satellite speed)
- estimation of integer number of 1ms cycles
- detection and compensation of 1ms rollovers

3.2 Snapshot Workflow

The goGPS library can run in standalone GPS mode, with a single receiver or in a master/rover configuration. It can also compute updates by individual epochs or by using an Extended Kalman Filter (EKF). All that is achieved by adding or swapping modules but without changing the basic workflow of goGPS:

- initialization

and for each set of observations:

- update computation
- error estimation

For the initialization we can either get an a priori location from AGPS (Assisted GPS) or compute it from Doppler (the static case only was implemented for this work). The Snapshot module currently runs epoch by epoch and we have left the option of running as a Kalman Filter for future updates.

3.3 Simulation of Snapshot positioning

We have tested the Snapshot algorithm using a reference track, computed by an epoch-by-epoch solution by least-squares adjustment of code observations. These were collected by a u-blox receiver, with its standard patch antenna mounted on the rooftop of a car driven in northern Norway.
In order to simulate Snapshot data we have truncated the pseudoranges down to their fractional part, added a random common bias to each epoch, added a random error to the timestamp between -/+ 5 seconds. Finally we've limited the observations to groups of 5, 6 and 7 satellites. Figure 3 shows the ENU components of the difference between the simulated Snapshot positioning results and the reference track; Table 1 reports the corresponding statistics.

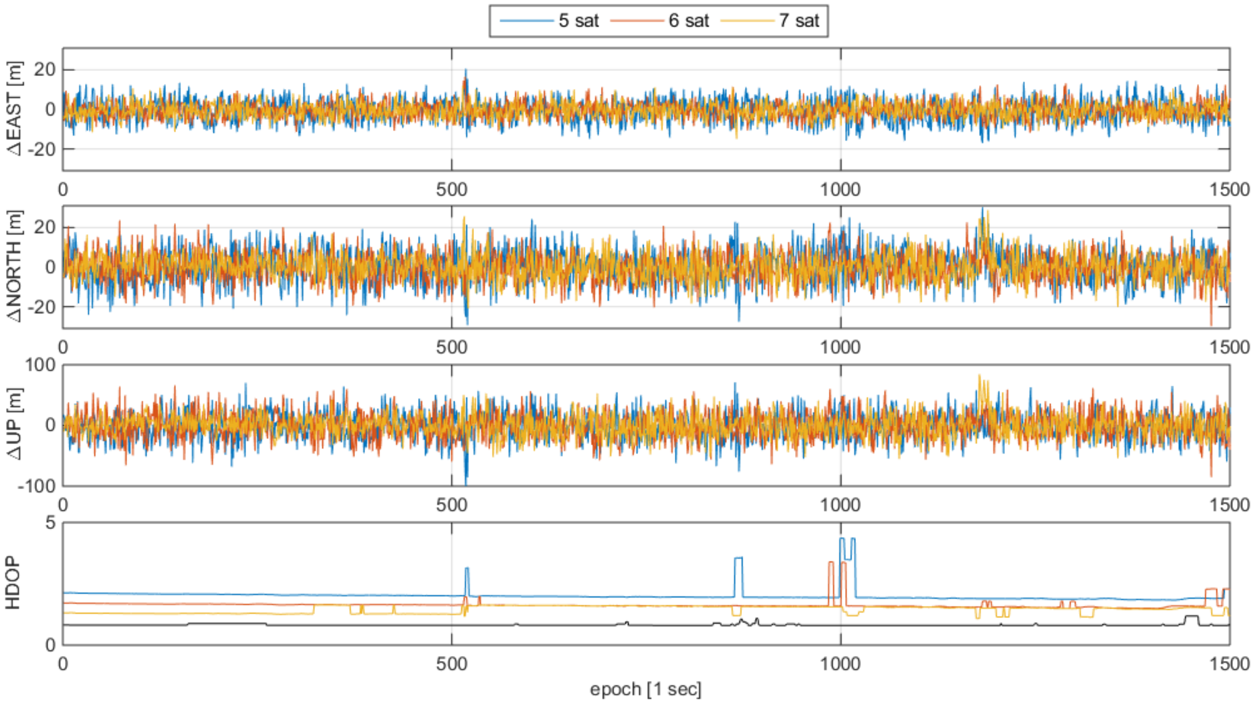


Figure 3: The top three panels show the difference between the results obtained by the implemented Snapshot algorithm using 5, 6 and 7 satellites and a reference track; the bottom panel shows the corresponding HDOP values, compared to the reference HDOP (black line).

	5 sat.			6 sat.			7 sat.		
	E	N	U	E	N	U	E	N	U
mean [m]	-0.7	0.2	-0.9	-0.4	-0.4	-1.1	-0.5	0.2	0.3
st.dev. [m]	6.2	9.0	24.0	4.2	7.8	23.1	4.1	6.9	19.1

Table 1: Mean and standard deviation of the difference between the Snapshot positioning results and the reference track.

3.4 Doppler Based Positioning Experiment

The Doppler based positioning currently implemented in goGPS Java is limited to the stationary receiver case (i.e. it does not take into account the receiver velocity). Figure 4 shows an example of the results achieved by processing the Doppler shift measured by a stationary u-blox receiver in Osaka, Japan. The plotted time series are the difference between the Doppler based positioning results, in their ENU components, and the known position of the receiver antenna (estimated as a daily solution in relative positioning mode with respect to a virtual reference station generated at about 60 m of distance).

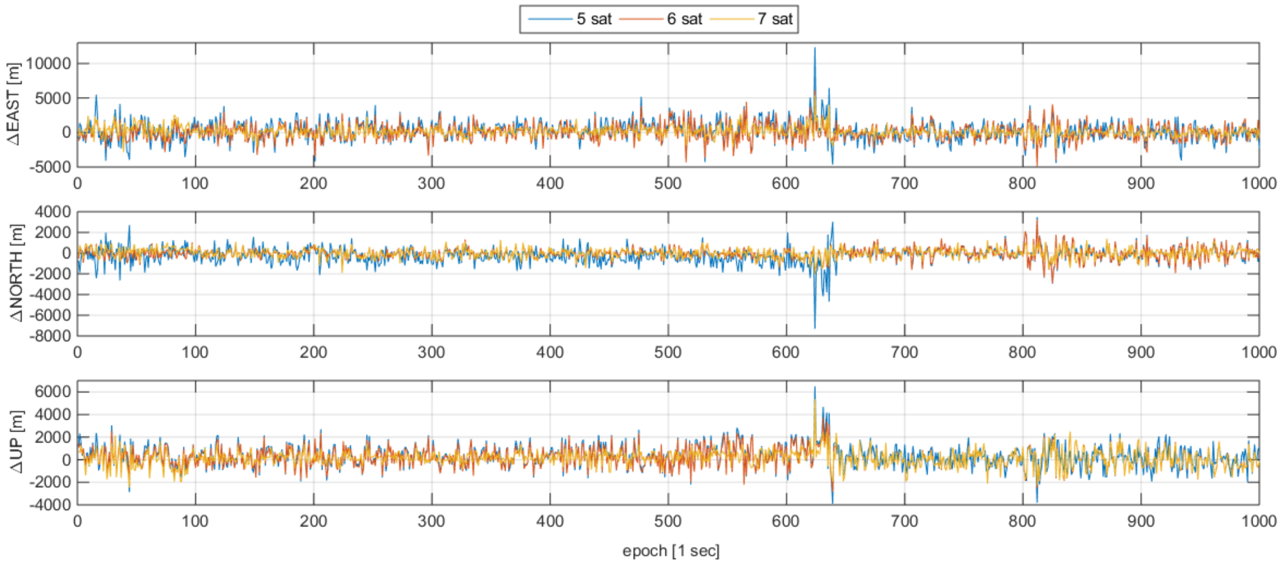


Figure 4: Difference between the Doppler based positioning results and the known position.

	5 sat.			6 sat.			7 sat.		
	E	N	U	E	N	U	E	N	U
mean [km]	0.210	-0.216	0.279	0.640	-0.42	0.184	0.157	-0.006	0.123
st.dev. [km]	1.513	0.818	1.024	1.216	0.564	0.903	0.747	0.476	0.740

Table 2: Mean and standard deviation of the difference between the Doppler based positioning results and the antenna known position.

4 Conclusion and Future Work

We have shown how goGPS could be extended to cover the case of Snapshot GPS by adding an extra state, i.e. the coarse time. We have implemented a new code module that applies a Snapshot positioning algorithm, by reusing or modifying existing building blocks from goGPS Java. A way to compute a priori location from Doppler was also provided.
Future work may include using the DTM module as a way to improve HDOP for the Snapshot module. A Kalman filter implementation could used for smoothing the errors especially with observations sets of only 5 satellites. Finally, the differential GPS aspect of goGPS could be exploited also in this case, using a reference station.

5 Acknowledgments

The authors would like to thank:
Mirko Reguzzoni, Daniele Sampietro and Andrea Gatti for providing tracking data from Norway.
Andrei Marinescu for sending the decompose_speed Matlab routine.
Professors Per Enge and Van Diggelen for their fantastic MOOC (www.gps-lab.org).

References

- ✔ Marinescu A & Cătălin D (2009). Assisted GPS System Design, Technical Report, Group 815, Aalborg University
- ✔ Othieno N (2012). Combined Doppler Time-free Positioning for Low Dynamics GNSS Receivers, Masters thesis, Concordia University
- ✔ Realini E, Yoshida D, Patocchi L, Reguzzoni M & Raghavan V (2010), Porting goGPS from MATLAB to Java: performance analysis, tests and results. Proceedings of FOSS4G 2010, Barcelona, Spain, 6-9 September 2010
- ✔ Realini E, Yoshida D, Reguzzoni M & Raghavan V (2012), Enhanced satellite positioning as a web service with goGPS open source software, Applied Geomatics, 4(2), 135-142
- ✔ Realini E & Reguzzoni M (2013). goGPS: open source software for enhancing the accuracy for enhancing the accuracy of low-cost receivers by single-frequency relative kinematic positioning, Measurement Science and technology, 24(11), 115010
- ✔ Remondi B (2004). Computing satellite velocity using the broadcast ephemeris, GPS Solutions
- ✔ Sirola N & Syrjärinne P (2002). GPS positioning without navigation data. Proceedings of ENC-GNSS2002, 27-30 May 2002, Copenhagen
- ✔ Strang G & Borre K (1997). Linear Algebra, Geodesy, and GPS, Wellesley-Cambridge Press
- ✔ Van Diggelen F (2009). A-GPS: Assisted GPS, GNSS, and SBAS, Artech House, first edition

GuideMe: an outdoor/indoor navigation app based on the i-locate open toolkit

Daniele Miorandi[1], Erinda Jaupaj[1], Carlo Caprini[1], Andrea Zito[1], Rossana Bartolacelli[1], Iacopo Carreras[1], Angelos Anagnostopoulos[2], Anestis Trypitsidis[2], Marc Bonazountas[2], Giuseppe Conti[3], Stefano Piffer[3], Nicola Dorigatti[3], Tao Feng[4], Joran Jessurun[4], Theo Arentze[4]

[1] U-Hopper srl, v. A. da Trento 8/2, 38100 - Trento, Italy. Email: {name.surname}@u-hopper.com

[2] EPSILON International SA, Monemvasias 27, 15125 Marousi, Greece. Email: bonazountas@epsilon.gr

[3] Trilogis Srl, Via Zeni 8, 38068 Rovereto (TN), Italy. Email: {name.surname}@trilogis.it

[4] Eindhoven University of Technology, Vertigo 8.16, 5600MB, Eindhoven, The Netherlands. Email: t.a.z.arentze@tue.nl

Abstract

GuideMe is an outdoor/indoor navigation app for Android. Based on open standards (e.g., indoorGML) and open data (from OpenStreetMap), it was built using the open source i-locate toolkit, which allows researchers, Web entrepreneurs and innovative startups to quickly and easily design and deploy indoor/outdoor location-based services. The GuideMe app provides end-to-end routing and turn-by-turn navigation through outdoor and indoor environments. Different positioning technologies are supported for both the outdoor and indoor parts. The app is released under an Apache v.2 license, leaving it open for the community to extend and enhance it.

Keywords

navigation, routing, indoor, indoorGML, i-locate

1 Introduction

Outdoor navigation services can nowadays be considered a commodity. Boosted by the ubiquitous adoption of GPS technologies in smartphones and by the emergence of popular mapping service, a large number of applications supporting routing and turn-by-turn navigation have been developed for various platforms.
But what about indoor environments?
The lack of a widely adopted standard for the representation of indoor spaces and the inherent difficulty in achieving accurate indoor positioning have long made this unfeasible. The recent emergence of a number of indoor positioning technologies and the arising of the indoorGML standard[1] are now changing the picture.

1 http://indoorgml.net/

The i-locate project[2] is a cooperative R&D&I endeavour whose mission is to develop an open source toolkit for fostering the creation of an ecosystem of indoor/outdoor location-based services. One of the services developed is GuideMe, an app for indoor/outdoor routing and navigation. The app is released as open source under a permissive license (Apache v.2) and is released to the community for being enhanced and tailored to specific service needs.
The remainder of this paper is organised as follows. Sec. 2 introduces briefly the i-locate toolkit. Sec. 3 presents the GuideMe app. Sec. 4 concludes the paper.

2 The i-locate toolkit

The i-locate toolkit is a flexible, open and extensible middleware meant to facilitate the development and deployment of location-based services by providing standardised interfaces and open-source technology enablers.

2.1 Architecture

The toolkit is based on a layered architecture, whose high-level logical structure is reported in Fig. 1.

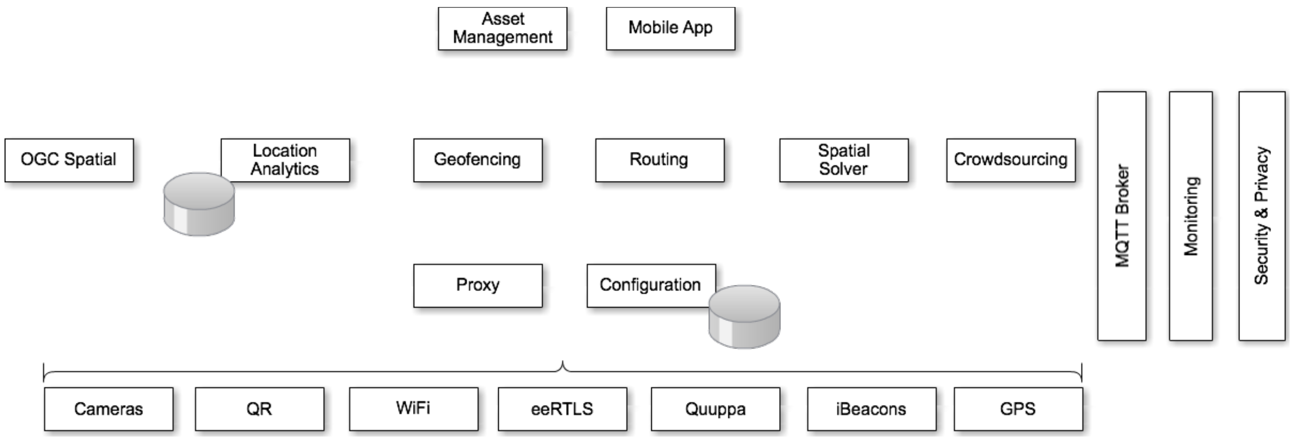


Figure 1: i-locate toolkit architecture (logical view).

The toolkit includes the following components:

- Core localization services
  - Proxy: represents the unique entry point for LBS enablers to get localization data about system entities. It connects with a number of positioning technologies, including WiFi, Quuppa, eeRTLS, cameras etc.
  - Configuration: provides access to additional information of specific devices within various indoor localization technologies like battery status or firmware revision etc.
- Generic LBS Enablers:
  - Spatial solver: the module provides an interface to the Open Data Repositories and to the current position of assets; it provides a spatial query interface, providing processes data to the caller.

2 http://www.i-locate.eu/

    - OGC Spatial: this module provides access to the geographical information in a standardised way, enabling third party software to ingest the (spatial) information without any conversion.
    - Routing: is the component capable to generate a route plan and corresponding turn-by-turn directions for a trip with given origin and destination location.
    - Geofencing: this module is responsible to ingest the location information of an asset (person) and on that verify spatial roles (entry/exit from a room) defined case by case.
    - Location analytics: this module computes a number of statistics related to the occupancy of indoor spaces. Indicators computed include total dwell time, frequency of visits to a given area, time spent within a given area, transitions among indoor areas etc.
    - Upload/Download: allows upload of arbitrary files attached to a site in the portal,and download through a web service. IndoorGML navigation data can also be downloaded.
    - Crowdsourcing: this module provides i-locate users with the opportunity to enter and validate geo-localised information.
- Specific LBS Enablers:
    - Asset management: the module provides the ability to accurately represents assets and the definition of maintenance processes.
    - Mobile app: this refers to the GuideMe app subject of this paper.
- Cross-cutting modules:
    - Security & Privacy: based on OpenAM[3] is the component allowing a relying party to identify the principal and determine if service is to be offered.
    - Monitoring: monitors the functioning of the toolkit runtime providing diagnostics and statistics.
    - Mqtt broker: a pub/sub communication system used to let enablers access location updates from the proxy.

More details can be found in Anestis et al. (2014) and Piffer et al. (2014).

2.1 Relevant Components

The following i-locate toolkit components are used to support the app backend operations.

2.1.1 Proxy

The goal of the proxy component is to provide location information relative to i-locate entity(ies), specified by a given unique ID. The key functionality provided by this module is to return the position of a given object; this is offered in both push and pull mode. At the moment it supports the following positioning technologies: WiFi (through Combain APIs[4]), Quuppa Intelligent Locating System[5], ZigPos eeRTLS[6], cameras, QR codes, GPS, EGNOS, iBeacons, UWB. See Miorandi et al. (2015) for more details.

3 https://www.forgerock.com/en-us/products/access-management/
4 https://combain.com/
5 http://quuppa.com/
6 http://www.zigpos.net/zigpos/index.php?c=2&a=36

2.2.2 Outdoor Localization

The outdoor localization component currently supports localization through GPS and WiFi (through the Combain APIs). It is shipped in the form of an Android native library, which is integrated in the app for ease of usage.

2.2.3 Routing

The routing service works for both indoor and outdoor routing, with a specific definition of the indoor data structure. Given the start and end locations specified by longitude and latitude coordinates, the service gives the optimal route. The routing service supports multimodal routing from door-to-door in a seamless way across outdoor and indoor environments. In this first release implementation, the routing service currently allows only a single level/floor for indoors. It will allow the navigation among multiple floors by incorporating stairs, elevators, open spaces, etc. Furthermore, in this first release turn-by-turn navigation guidance adopts the same structure as for outdoor routing.
The routing service of integrated indoor and outdoor navigation is built upon the OpenTripPlanner (OTP)[7], an open source software specifically designed for multimodal routing.
The data used for outdoor routing are based on OpenStreetMap[8] and GTFS data[9]. The i-Locate routing system extends the OTP by incorporating indoor routing using the indoorGML standard for the representation of indoor spaces.

2.2.4 OGC Spatial

The spatial services are supported by the use of Geoserver API. In the first version of the toolkit, following OGC standards are provided:

- WMS, Web Map Service;
- WFS, Web Feature Service.

3 The GuideMe app

The GuideMe app comes in the form of a flexible prototype which supports:

- Display of a map with current location;
- Ability to search the route to a given location (across indoor/outdoor);
- Display the computed route on a map;
- Navigate the user to the intended destination;
- Notify the user that it is entering the destination building;
- Display an indoor map with the current location;
- Signal the user of the arrival at the intended destination.

The mobile client has been developed using the Titanium Appcelerator SDK v.3.4.1[10] for ensuring cross-platform support. In the initial phase of development, and given operating system-level restriction in terms of access to raw GPS and WiFi data (see Outdoor Localization component description above), we have decided to focus on Android as target operating system.
The mobile client interoperates with the following other toolkit components running in the backend:

---

7 http://www.opentripplanner.org/
8 http://www.openstreetmap.org/
9 https://developers.google.com/transit/gtfs/
10 http://www.appcelerator.com/titanium/

- Proxy: for retrieving indoor/outdoor position based on the provided client-measured parameters (e.g., WiFi access points BSSID and RSSI);
- Outdoor localization: for retrieving GPS and WiFi data;
- OGC Spatial: for accessing reverse geocoding functionality;
- Routing: for computing route to a destination (outdoor/indoor);
- Upload/Download: to dynamically download indoor map of the reference building.

The GuideMe app is released under Apache v.2 license at https://gitlab.com/ilocate/ilocate-app

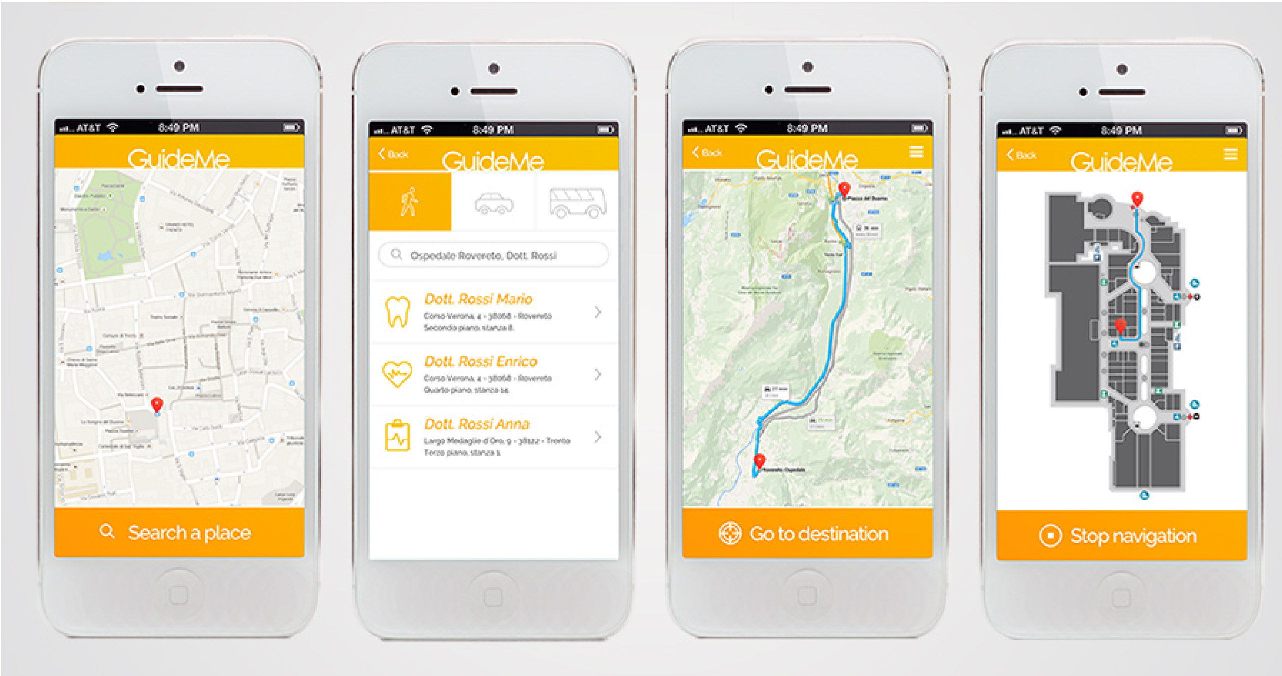


Figure 2: GuideMe app screenshots.

4 Conclusions

GuideMe is meant to act as a template for the community to quickly and easily develop navigation apps able to seamlessly work in both outdoor and indoor environments. The app builds upon the open source i-locate toolkit. While the app already supports a number of different positioning technologies and presents basic functionalities, we strongly encourage the community to contribute to its further development and to use it, properly tailored, to support additional use cases.

References

✔ Trypitsidis, A. et al. (2014). System Architecture. i-locate project deliverable D.1.4. Retrieved from htttp://www.i-locate.eu/public-deliverables/

✔ Piffer, S. et al. (2014), i-locate toolkit v.1.0. i-locate project deliverable D.3.1. Retrieved from http://www.i-locate.eu/public-deliverables/

✔ Miorandi, D. et al. (2015). Localisation and tracking systems. i-locate project deliverable D.3.2. Retrieved from http://www.i-locate.eu/public-deliverables/

Development of Indoor Environments with a Novel Indoor Mapping Approach for OpenStreetMap

Thomas Graichen[1], Sven Quinger[2], Ulrich Heinkel[3], Marek Strassenburg-Kleciak[4]

[1] thomas.graichen@etit.tu-chemnitz.de

[2] sven.quinger@etit.tu-chemnitz.de

[3] ulrich.heinkel@etit.tu-chemnitz.de

[4] Marek.Strassenburg-Kleciak@elektrobit.com

Abstract

The community project OpenStreetMap (OSM), which is well-known for its open geographic data, still lacks a commonly accepted mapping scheme for indoor data. Most of the previous approaches show inconveniences in their mapping workflow and affect the mapper's motivation. In our paper an easy to use data scheme for OSM indoor mapping is presented. Finally, by means of several rendering examples from our Android application, we show that the new data scheme is capable for real world scenarios.

Keywords

OpenStreetMap, indoor maps, data scheme, mapping

1 Introduction

The vast amount of projects and applications of the OpenStreetMap (OSM) has shown that Volunteered Geographic Information (VGI) is able to compete with proprietary and commercial solutions, such as Google Maps, in outdoor environments. The main reason for this is the extensive and detailed coverage of the project's maps for the most parts of the world.

Unfortunately, the map quality for interiors of buildings is not as high as for outdoor maps in OSM. The OSM project community has still not agreed on an unified data scheme for indoor maps, while global players, such as Google and Microsoft, have already mapped the indoor environments of a vast amount of shopping malls, airports and other public buildings.

In order to achieve a coverage level for OSM indoor maps that is comparable to the high quality of the outdoor maps and to develop a real alternative to existing commercial solutions, it is absolutely necessary to establish a mapping scheme for indoor environments that is easy to use and to understand.

Therefore, we describe the principles and practical application problems of existing indoor mapping schemes (*2. Related Work*) and propose a novel approach for mapping indoor environments for OSM (*3. Indoor Data Scheme*). Finally, we present a rendering example of our Android map application (*4. Conclusions and Further Work*).

2 Related Work

A precondition for mapping building interiors and explaining the existing approaches is the understanding of the basic OSM mapping. Map data can be stored in XML format and is comprised of three basic data elements: nodes, ways and relations.
The purpose of nodes is to describe geometrical points in the world in the WGS84 projection and provide, if available, additional information with so-called tags. A tag always consists of one key-value pair. By connecting nodes to ways, it is possible to describe lines (unclosed way) and polygons (closed way). With appropitate tags, the purpose of a way is defined (e.g. street, building, etc.). Relations are used to group nodes and ways to one semantic meaning by specific tags. An example for relations are bus routes, which consist of several streets (ways) and bus stops (nodes).
All these data elements are used by one of the best known and most used mapping schemes for indoor data, the IndoorOSM proposal (Goetz, 2011 & 2012). The concept of this approach is the construction of a hierarchical data structure (see simplified visualisation in Figure 1), which is able to represent the 3-D structure of buildings.

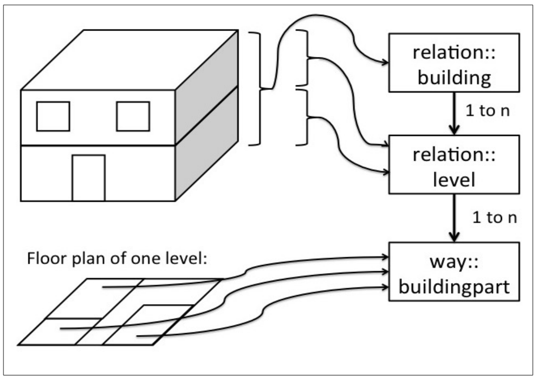


Figure 1: IndoorOSM proposed data scheme.

The top level element of this data structure is a relation, which represents the *building* itself and contains the *levels* and the outer shell of the building. Levels are modelled as relations as well and contain so-called *building parts*. These building parts are spatial elements that represent rooms, floors, stairways, etc. They are mapped as ways and describe the geometrical layout. Furthermore, Points of Interest (POIs), such as entrances, doors, etc., are represented as nodes and may be members of *buildings*, *levels* or *buildingparts*. Another proposal (Hubel, 2014) uses a similar data scheme.
However, both approaches have one drawback: They are based on a high number of necessary relations for the description of a building. Therefore it proves difficult for the mapper to handle the data with the available editors, such as JOSM. For those reasons, both approaches are rarely used in the community.
In the following part of the paper, a way of obtaining a structured description without such a number of relations is introduced.

3 Indoor Data Scheme

The goal of the presented data scheme is to avoid as much mapping overhead as possible and to reduce the barriers for new indoor mappers. Therefore, the usage of relations must be rigorously reduced.

3.1 Proposed Data Scheme

The basic question is whether relations are really needed for the description of buildings. The top level relation of the previously named approaches (see Figure 1, building relation) is used to describe the building and its inner elements, but this could also be achieved by a geometrical analysis of the outer shell of a building and all inner elements. This analysis may be performed by clipping of polygons (O'Brien, 2005). By discarding the top level relation, a preprocessing step for the assignment of rooms to buildings is now required for rendering and searching applications.

The second relation, i.e. the relationship between levels and building parts, can also be expressed by a simple level tag for each level element (ways and POIs). With this approach, it is possible to avoid all relations and provide a much easier mapping approach (see Figure 2) at the expense of more complex processing steps in rendering and searching algorithms.

A short example of this new approach applied to an indoor area, such as a room, is given by the following XML snippet (Figure 3).

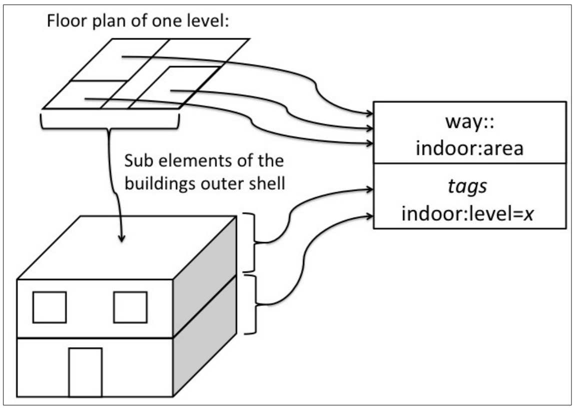


Figure 2: Visualisation of the novel indoor mapping scheme.

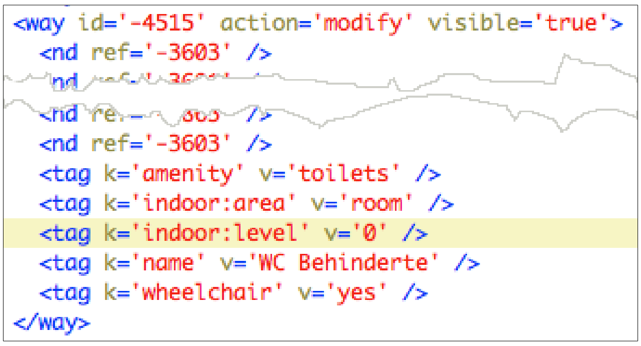


Figure 3: An example way, which describes a room at level 0 in the novel data scheme syntax (high-lighted line marks the level tag).

3.2 Rendered Map Examples

The new scheme was used to map several buildings of the Chemnitz University of Technology and an underground car park with the editor JOSM. With a specially defined MapCSS file ((OpenStreetMap-Wiki, 2015) and appropriate data filters, the rendering of the indoor data was tested in JOSM. Thereby the MapCSS file was used to control the rendering, e.g. by specifying the colour a room is filled with.

Furthermore, an Android application was developed to display both outdoor and indoor maps. The rendering of outdoor data was realised by the Android library *mapsforge* (Mapsforge Project, 2015), whereas the indoor data rendering was realized by an own developed rendering library (Graichen, 2014). Figures 4 and

5 show rendering examples of this Android application.

Figure 4: Rendering example of an university building.

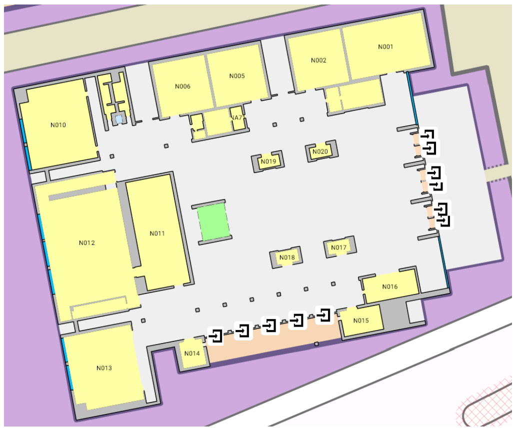


Figure 5: Rendering example at higher zoom level.

4 Conclusions

In this paper, a novel approach for mapping indoor environments is presented. Compared to previous proposals, this new mapping scheme renounces the use of relations and therefore is easy to understand and to use with available OSM editors. However, the simplification of the indoor mapping requires additional preprocessing steps in rendering, routing or searching applications in order to reconstruct the structure in buildings (building → level → room).
By means of an own developed Android map rendering application, it has been shown that such a preprocessing step is feasible and that our approach is suitable for real word scenarios.

References

- ✔ Goetz, M., & Zipf, A. (2011). *Extending OpenStreetMap to indoor environments: bringing volunteered geographic information to the next level*. CRC Press: Delft, The Netherlands.
- ✔ Goetz, M. (2012). IndoorOSM. Retrieved March 9, 2015, from https://wiki.openstreetmap.org/wiki/IndoorOSMReference
- ✔ Graichen, T. (2014). *A Combined In- and Outdoor Map for Android.* State of the Map EU 2014. 13 to 14. June 2014, Karlsruhe, Germany.
- ✔ Hubel, A. (2014). Proposed features/indoor. Retrieved March 9, 2014, from https://wiki.openstreetmap.org/wiki/Proposed_features/indoor
- ✔ Mapsforge Project. (2015). Rendering of OpenStreetMap data. Retrieved March 9, 2015, from https://github.com/mapsforge/mapsforge
- ✔ O'Brien, J., (2005). Clipping and Hidden Surfaces. Retrieved March 9, 2015, from http://www-inst.eecs.berkeley.edu/~cs184/fa05/lectures/lecture-06.pdf
- ✔ OpenStreetMap-Wiki. (2015). MapCSS Implementation. Retrieved March 9, 2015, from https://josm.openstreetmap.de/wiki/Help/Styles/MapCSSImplementation

Software Defined Receivers in GNSS scientific applications: variometric approach to exploit GNSS-SDR phase observations

Mara Branzanti[1], Javier Arribas[2], Carles Fernandez-Prades[2], Mattia Giovanni Crespi[1]

[1] Geodesy and Geomatics Division, University of Rome "La Sapienza"

[2] Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)

Abstract

The use of GNSS for many high accuracy applications is continuously increasing. Most of the methodologies of GNSS data processing are based on the observations collected by conventional (geodetic and low cost) receivers. These devices are rather sophisticated, but appears as black boxes to the user: basically, the user does not have any access to the internal structure of the receiver. In this respect, a new class of devices has been recently proposed: the so-called Software Defined Receivers (SDRs). Up today, the use of SDR is still limited for experimental receivers, but their development, their very low cost and their flexibility have opened new possibilities including high precision applications. In this work, the variometric approach (implemented in the VADASE software) is applied to single frequency phase observation obtained from the open source multiconstellation GNSS-SDR software receiver.

The final aim of this work is twofold: one hand it evaluates the reliability of GNSS-SDR phase observations, on the other hand it asses VADASE reachable accuracy when applied to phase observation retrieved by a SDR. To these aims some tests were performed and an accuracy at the decimeter level was found.

The obtained results show that in the next future SDR will have an important role in GNSS scientific applications, going beyond the limits imposed from conventional GNSS receivers.

Keywords

GNSS-SDR, VADASE, real-time, multi-constellation, high-accuracy, phase observations

1 Introduction

In the last years the technological development (high performances GNSS receivers) and the progress in data processing, have been opening new opportunities for the use of GNSS in scientific applications where a positioning accuracy at the few centimeters level is required: precision farming, vehicle precise navigation, Earth monitoring and others. Unfortunately, existing solutions for these ambitious applications are typically very costly. In most cases dual frequency geodetic receiver are needed to reach a centimeter accuracy, but emerging investigation projects are moving toward the use of low cost single frequency receiver. Another important aspect to consider is that most of the proposed methodologies are based on the observations collected by conventional (geodetic and low cost) receivers. These devices are rather sophisticated, but appears as black boxes to the user. Basically, the user does not have any access to the internal structure of the receiver. In this respect, a new class of devices has been recently proposed: the so-called Software Defined Receivers (SDRs).

Differently from the classic ones, SDRs allow a complete interaction user-receiver and any kind of customization according to specific needs. SDRs are finding a great response in the current GNSS scenario particularly now that new signals (e.g. GPS L5 frequency) and new constellations (Galileo, GLONASS, Beidou) have been introduced and consequently receivers architecture must be continuously updated. When considering a SDR, the user can chose to extend the receiver functionalities to different constellations, allow phase observations in addiction to code observations, or integrate GNSS observations with other devices (such as accelerometers).

Up today, the use of SDRs is still limited for experimental receivers, but their development, their very low cost and their flexibility have opened new possibilities, including high precision applications.

This work aims to investigate the potentialities of the open source GNSS-SDR Software Defined Receiver. Its phase observations are exploited for the first time, in order to demonstrate the potentialities of SDRS in scientific applications. In particular, the reliability of GNSS-SDR phase observations is assessed exploiting the single frequency capabilities of the variometric approach implemented in the VADASE software, an innovative strategy for GNSS data processing.

In Section 2 GNSS-SDR is introduced and the state of art of its implementation is reported. The Sub-section 2.1 proposes a methodology to produce hybrid GPS-Galileo observables within the receiver. An overview of the variometric approach is given in Section 3. Finally, in Section 4, the first VADASE solution over GNSS-SDR code and phase observations are presented, considering simulated and real life signals.

2 GNSS-SDR

GNSS-SDR, available at *http://gnss-sdr.org* and mainly developed at CTTC (Centre Tecnologic de Telecomunicacions de Catalunya, Barcelona, Spain), is an open source GNSS software receiver (Fernández–Prades, Arribas, Closas, Avilés and Esteve, 2011).

The software, written in C++, is able to work either from raw signal samples stored in a file, or in real-time with a radio frequency front-end. It provides, under General Public License (GPL), the implementations of all the algorithms required by a GNSS receiver: raw sample reading and conditioning, execution of signal processing block performing acquisition, code and phase tracking, lock detectors, demodulation and decoding of the navigation message, observable computation and PVT (Position, Velocity, Time) solution. Moreover it makes availabale to the user observation and navigation files in the standard

RINEX format. The general diagram of the proposed software receiver is shown in Figure 1.

2.1 A multi-constellation software receiver

The GNSS-SDR project started in 2010 with the implementation of all the steps to be operative with the GPS constellation (Arribas, 2012). The receiver was later extended to the Galileo constellation in a work developed within the Google Summer of Code program (Fernández–Prades, Arribas, Esteve, Pubill and Closas, 2012): in 2013, GNSS-SDR allowed to achieve for the first time a stand-alone Galileo-only position fix, in real-time, with an open source software defined GNSS receiver (GPS World staff, 2014).

At present it is a completely operative single-frequency GPS(L1)-Galileo(E1) receiver; for the next future its extension toward the GLONASS and Beidou constellations is already planned, as well as its extension to multiple bands.

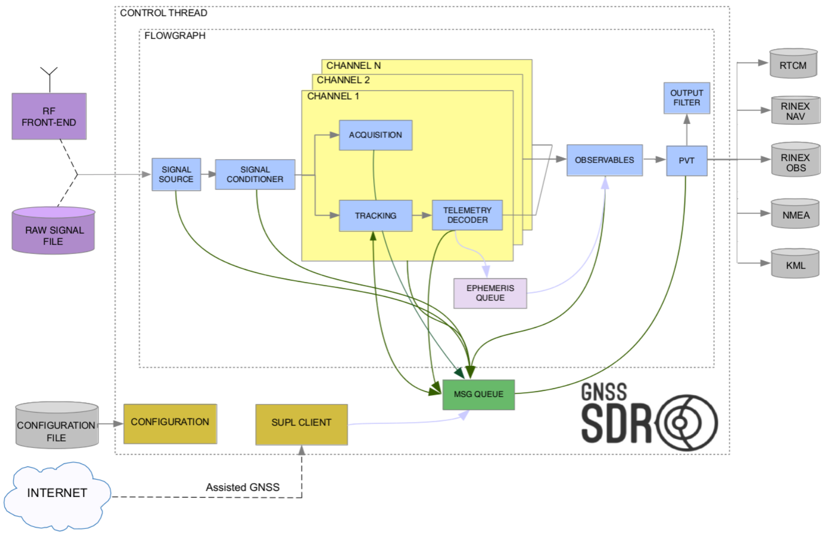


Figure 1: General diagram of the GNSS-SDR control plane and signal processing flow graph. Users can invoke a particular implementation and parameters for each processing block (blue boxes) via a single configuration file.

2.1.1 GPS-Galileo hybridization at observables level

In 2014 the receiver was further developed to work in a hybrid mode with respect to the GPS and Galileo constellations.

As a matter of fact, Galileo System Time (GST) and GPS Time (GPST) do not use the same time reference, and hence a time offset exists between both systems. From a PVT computation stand view, this means that the pseudoranges determined with Galileo are referenced to the GST, while the ones from GPS use the GPST. The difference between GPST and GST, if unknown, requires an extra equation and thus an extra in-view satellite in order to compute a PVT solution using both GPS and Galileo measurements in the same set of equations. Thus, five or more in-view GPS and Galileo satellites would be needed to obtain a hybrid PVT solution.

Those two internal times are derived independently on one another but, luckily enough, the GPS-Galileo Time Offset (GGTO) is distributed through Galileo's

navigation message (European Union, 2010). This means that, if the GGTO is properly retrieved and included in the navigation solution, a total number of four in-view satellites (in any combination: 3 GPS and 1 Galileo satellites, 1 and 3, or 2 and 2) would be enough to compute a PVT fix, effectively allowing interoperability between both systems.
Exploiting the availability of the GGTO in the Galileo navigation message, an hybrid configuration of GNSS-SDR at the observables level was implemented (Arribas, Branzanti, Fernández–Prades and Closas, 2014). In the hybrid configuration, GNSS-SDR can combine GPS and Galileo data at the observable level, providing better capabilities than the ones that would be achieved by relying solely on one system. The final result is an expanded set of observations, that can be considered coming from a unique (expanded) constellation.

3 VADASE

The Variometric Approach for Displacements Analysis Stand-alone Engine (VADASE) was originally proposed (Colosimo, Crespi and Mazzoni, 2011) as innovative solution to estimate in real-time rapid movements of GPS receivers. In this respect, it was thought to be applied in the seismology field.
The approach, based on time single differences of carrier phase observations (Hoffman-Wellenhof, Lichtenegger, and Wasle, 2008), only needs observations collected by a unique GNSS receiver and broadcast product available in real-time. Moreover, differently from other data processing approaches (i.e. differential positioning (DP) and precise point positioning (PPP)), VADASE does not require phase ambiguity solving (Teunissen and Keusberg, 1996), thus it is also able to work with single-frequency data only.
VADASE is implemented and continuously devoleped at Geodesy and Geomatics Divsion of “La Sapienza” University of Rome. It is subject of an international patent and was awarded the German Aerospace Agency (DLR) Special Topic Prize and the Audience Award at the European Satellite Navigation Competition 2010.
The effectiveness of VADASE was proved for seismology aims, through its application to the catastrophic Tohoku-Oki earthquake (Mw = 9, March 11, 2011, 02:03:51 UTC) (Branzanti, Colosimo, Crespi and Mazzoni, 2013) and to the Emilia earthquake (Mw = 6.1, May 20, 2012, 02:03:51 UTC) (Benedetti, Branzanti, Biagi, Colosimo, Mazzoni and Crespi, 2012). In the latter, the accuracy level of VADASE when processing L1 observations only was evaluated and an accuracy at few centimeters level was found.
Such high quality results, obtained from single-frequency observations of geodetic receivers, encouraged further experimentations considering observations collected by low-cost GPS and Galileo receivers. The results obtained exploiting the first four Galileo satellites in orbit were submitted to ESA (Branzanti, Benedetti, Colosimo, Mazzoni and Crespi, 2014) and the VADASE team received one of the 50 ESA Certificate for Galileo In-Orbit-Validation Fix 2014.
Finally, a kinematic implementation of the variometric approach was recently proposed, in the so-called Kin-VADASE (Branzanti, 2015), in order to extend its application to the navigation field and to retrieve kinematic parameters

(accelerations, velocities and positions) of vehicles in motions (land, maritime or air vehicles)

4 GNSS-SDR and VADASE: software receiver to scientific applications exploiting the variometric approach

GNSS-SDR is one of the most advanced open-source Software Defined Receiver thanks to its multi-constellation approach, the continuous improvements in the implementation and for the availability of public bibliography.
Intense work was carried out to evaluate its functionality at the level of signal processing (Fernández–Prades et al. 2011) or PVT solutions based on code observations (Arribas et al. 2014). However, its intermediate products (RINEX observations and navigation files), mainly as regards the recent availability of phase observations, have never been exploited for scientific purposes.
At present, not all the standard low-cost hardware receivers allow the user to extract raw data (RINEX) for post processing purposes. Furthermore, Galileo-GPS low-cost receivers are hardly available in the current market, being the Galileo constellation still under development (its full operation capability is expected for the end of this decade). Thus, it is not so common to experiment and test scientific algorithms exploiting observations coming from both constellations.
As matter of fact, the possibility to extract GPS and Galileo RINEX files within GNSS-SDR is a precious opportunity for the GNSS scientific community.
This section wants to experiment and evaluate the potentialities of GNSS-SDR in scientific applications, exploiting the single frequency capabilities of the variometric approach. To this aims, some tests were performed on the bases of simulated and real life signal. The experimentations have a twofold purpose: on one hand they evaluate the reliability of GNSS-SDR phase observations, on the other hand they assess the accuracy that can be achieved in processing with a scientific software (VADASE) the observables produced by a software defined receiver.

4.1 Experiments with simulated signal

The first experiments were carried out considering synthetic data produced by the IFEN GmbH NavX-NCS Simulator. GPS and Galileo signals were generated simulating the whole Galileo constellation, in its final configuration. The geometry of the scenario is shown in Figure 2: six Galileo satellites and eight GPS satellites were visible in the sky.

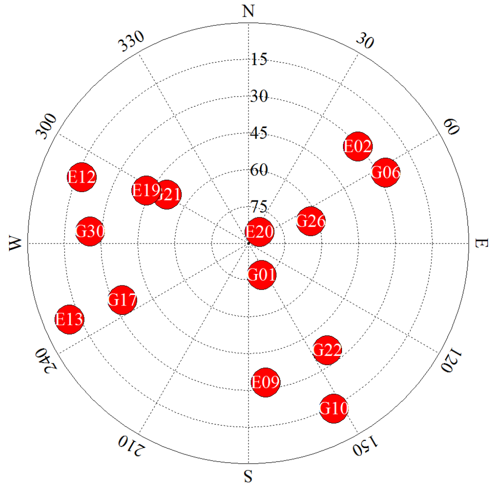


Figure 2: Visible GPS (G) and Galileo (E) satellites for the synthetically generated GNSS signals.

GNSS-SDR was set in its hybrid configuration to acquire, track, decode and produce observables of both constellations. Data were collected for a time interval of 3 minutes.
This experiment wanted to: *a)* evaluate the implementation of code and phase observables generation block in GNSS-SDR *b)* compare VADASE GPS solutions with VADASE Galileo solutions (using the same number of observations for each constellation), in order to make a comparison between constellations) *c)* exploit the hybrid configuration of GNSS-SDR to compute GPS-Galileo combined VADASE solutions *d)* compare the accuracy of the solutions obtained considering just one constellation at a time to the ones obtained from the hybrid approach.
To these aims, GNSS-SDR code and phase observations, collected in the standard RINEX format, were processed with the VADASE software in the following configuration: *1)* Galileo solution with six in view satellites *2)* GPS solutions with eight in view satellites *3)* GPS solutions with six satellites (satellites G22 and G26 were not used in the processing) in order to have GDOP values as similar as possible to the ones of the Galileo processing, and make a comparison between constellations under the same conditions *4)* GPS and Galileo combined solutions exploiting the fourteen satellites in view.

<u>4.1.1 Code solutions</u>

VADASE solutions over GNSS-SDR code observations are graphically represented in Figure 3, while statistics are reported in Table 1. GPS and Galileo solutions are at the same accuracy level when a similar geometry, with six satellites in view (GDOP values are 2.47 and 2.44 for GPS and Galileo respectively). The East and North accuracy is between one and two meters, while the RMSE value in the Up component is 6.2 m for GPS and 6.7 m for Galileo. In the hybrid solutions the number of satellites increases from six to fourteen, reducing the GDOP value to 1.55. In this satellite configuration the East and North accuracy in terms of RMSE of the solutions is 1.61 m and 1.97 m: it does not change significantly with respect to GPS-only and Galileo-only solutions, but with code observations higher performances can hardly be achieved. Differently, a significant improvement (of about 50%) occurs in the Up component, reaching a final accuracy of about 3 m.

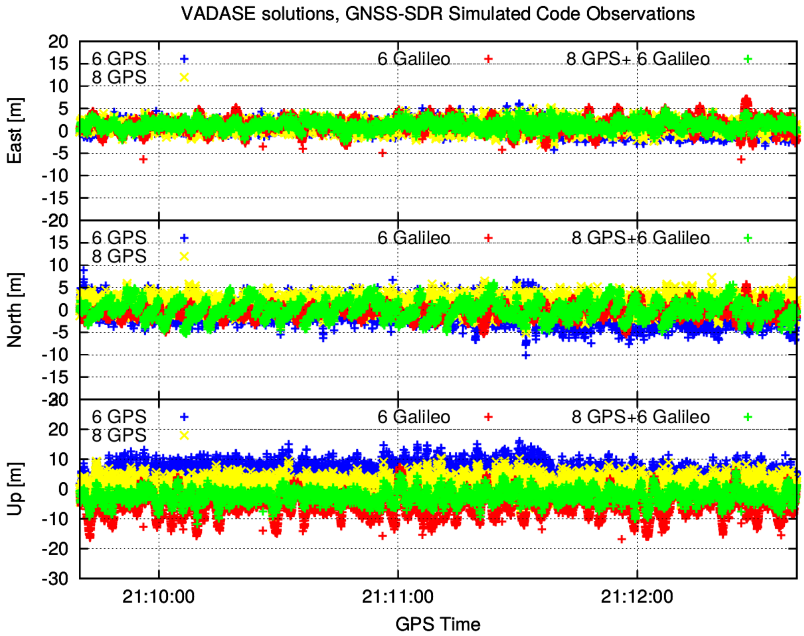


Figure 3: Experiment with simulated signal: VADASE solutions over GNSS-SDR code observations under different configurations (GPS, Galileo, GPS+Galileo).

4.1.2 Phase solutions

The variometric approach was also applied to phase observations produced by GNSS-SDR. Even in this case, GPS only, Galileo only and combined GPS-Galileo solutions were carried out. Probably due to a problem in the tracking algorithm implemented in the receiver, GPS phase observations have a validity interval of only 35 seconds. Differently from GPS, Galileo solutions are stable for the entire time interval. This issue will be surely faced and solved in the future, but it is not the aim of this work opening a detailed discussion about the GPS tracking algorithm of the receiver. Thus, VADASE accuracy with GPS only and GPS-Galileo observations was assessed over the first 35 seconds (from 21:09:40 to 21:10:15, Figure 4, while Galileo accuracy was evaluated both over the first 35 seconds (to make a comparison between constellations) and over the entire interval.

Statistics over the reduced time interval (35 seconds) are presented in Table 2. Under the same satellite geometry (6 satellites used in the processing for both constellations), Galileo results are much more accurate than GPS. RMSE of GPS solutions is at the decimeter level in all the three components; RMSE of Galileo solutions is at the millimeters level in East and North and of few centimeters in the Up component. In the hybrid configuration, the final statistics improve with respect to the GPS only solutions. RMSE is 0.05 m, 0.03 m. 0.08 m in the East, North and Up components respectively. Galileo statistics over the entire three minutes interval are reported in Table 3: the accuracy is at the centimeter level in planimetry (0.007 m and 0.032 m in East and North) and of three decimeters in the Up component.

GPS				
	East [m]	North [m]	Up [m]	
Average	0.995	1.323	1.178	**8 Sat GDOP=2.27**
St. dev	1.055	1.471	2.568	
RMSE	**1.450**	**1.979**	**2.862**	
Average	0.683	-0.434	5.489	**6 Sa GDOP=2.47**
St. dev	1.240	2.238	3.003	
RMSE	**1.415**	**2.281**	**6.257**	
Galileo				
	East [m]	North [m]	Up [m]	
Average	1.252	-0.506	-5.547	**6 Sat GDOP=2.44**
St. dev	1.617	1.475	3.874	
RMSE	**2.045**	**1.567**	**6.707**	
GPS + Galileo				
	East [m]	North [m]	Up [m]	
Average	1.260	0.120	-2.462	**14 Sat GDOP=1.55**
St. dev	1.068	1.967	2.262	
RMSE	**1.651**	**1.971**	**3.344**	

Table 1: Experiments with simulated signal: statistic over a 3 minutes time interval of VADASE solutions with GNSS-SDR code observations under different satellite configurations (GPS, Galileo, GPS+Galileo).

GPS				
	East [m]	North [m]	Up [m]	
Average	0.062	-0.045	0.095	**8 Sat GDOP=2.27**
St. dev	0.038	0.043	0.06	
RMSE	**0.073**	**0.063**	**0.113**	
Average	0.058	-0.082	0.063	**6 Sat GDOP=2.47**
St. dev	0.035	0.103	0.096	
RMSE	**0.068**	**0.132**	**0.115**	
Galileo				
Average	0.001	-0.005	0.039	**6 Sat GDOP=2.44**
St. dev	0.002	0.003	0.024	
RMSE	**0.002**	**0.006**	**0.045**	
GPS + Galileo				
	East [m]	North [m]	Up [m]	
Average	0.042	-0.02	0.066	**14 Sat GDOP=1.55**
St. dev	0.026	0.022	0.040	
RMSE	**0.05**	**0.029**	**0.077**	

Table 2: Experiments with simulated signal: statistic over a reduced time interval (from 21:09:40 to 21:10:15) of VADASE solutions with GNSS-SDR phase observations under different satellite configurations (GPS, Galileo, GPS+Galileo).

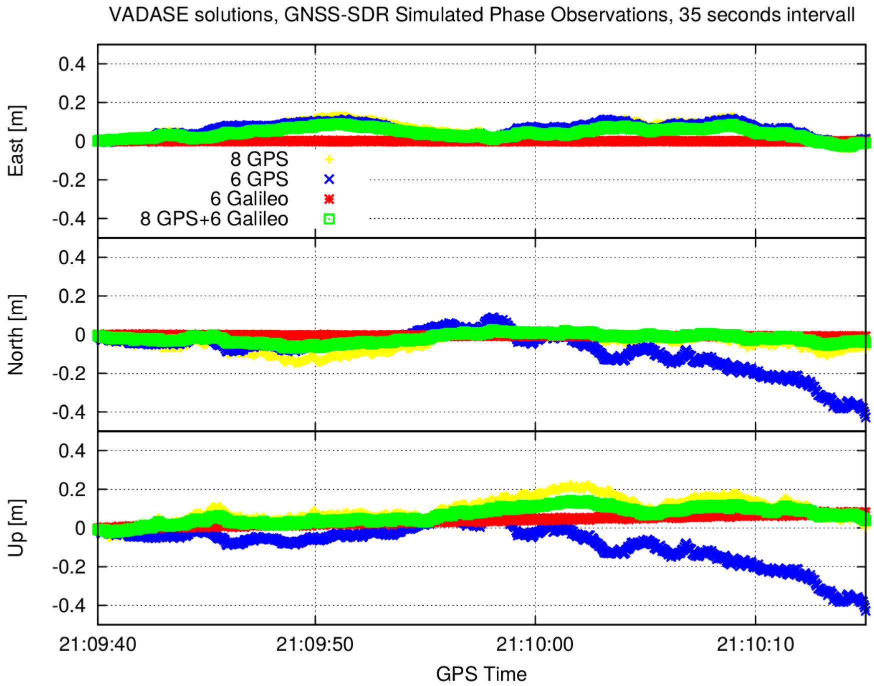


Figure 4: Experiments with simulated signal: zoom over the first 35 seconds of VADASE solutions with GNSS-SDR phase observations.

Galileo				
	East [m]	North [m]	Up [m]	
Average	-0.005	-0.028	0.27	**6 Sat GDOP=2.44**
St. dev	0.004	0.015	0.159	
RMSE	**0.007**	**0.032**	**0.313**	

Table 3: Experiments with simulated signal: statistic over a 3 minutes time interval of VADASE solutions with GNSS-SDR Galileo phase observations.

4.2 Experiments with real life signal

Here the capability of GNSS-SDR for high precision applications is effectively investigated, through the application of the variometric approach to phase observations retrieved from real life signal. In the experiment only GPS observations were collected, thus the GPS-Galileo interoperability with real data is not investigated, but it is addressed to future studies.

The satellite geometry during the data collection is reported in Figure 5. As the case of simulated signal, GPS solutions phase observations have a reduced period of reliability. Hence, accuracy was assessed over the time interval from 10:17:15 to 10:17:50. Statistics are reported in Table 4 and solution are graphically represented in Figure 6.

The RMSEs of the solutions are 0.15 m, 0.20 m, 0.12 m in the East, North and Up components respectively. As expected, the performances with real life signals are slightly worse that the ones obtained with simulated data, due to the presence of unmodelled sources of error both from GNSS-SDR and VADASE. Moreover, the satellite geometry is not in favour of a good final accuracy

(GDOP values are variable between 9.1 and 11.5). However the obtained results are really encouraging and represent the first step towards the use of SDRs to high accuracy scientific applications.

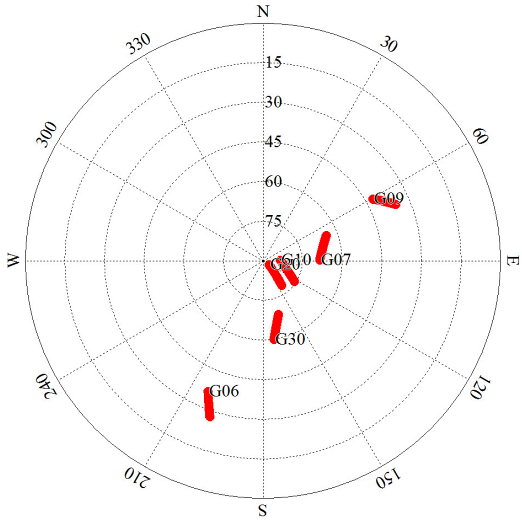


Figure 5: Visible GPS satellites in the experiment with real signal.

GPS				
	East [m]	North [m]	Up [m]	
Average	0.131	-0.154	0.054	**5 Sat 9.1<GDOP<11.5**
St. dev	0.067	0.123	0.109	
RMSE	**0.147**	**0.197**	**0.121**	

Table 4: Experiments with real signal: statistics over 35 seconds time interval of VADASE solutions with GNSS-SDR phase observations (GPS constellation).

5 Conclusion

In this work the reliability of GPS L1 and Galileo E1 code and phase observations produced by the open source GNSS-SDR is investigated. The receiver was used in its hybrid configuration, in order to produce GPS and Galileo observations referred to a unique system time, allowing to work with a unique "expanded" constellation. Observations were processed with an innovative methodology of GNSS data processing (implemented in the VADASE software), which is capable to work also with single-frequency phase observations. In particular, GPS only, Galileo only and GPS-Galileo solutions were computed.

In a first stage the implementation of GPS and Galileo observable generation was validated over simulated signal. The results demonstrated that observations are correctly retrieved (for the considered time interval) and that the opportunity to set the receiver in a hybrid configuration is a considerable added value: the user can effectively work with an expanded constellation, improving the satellite geometry and consequently the final accuracy.

The effective accuracy was assessed over real life GPS phase observations. An accuracy at the decimeter level in the three component was found with five

satellite in view.
The performed experimentations showed the potentiality of the integration between GNSS-SDR and VADASE and the achieved results suggest that SDRs will have an important role in the future scientific (or industrial) high accuracy applications.

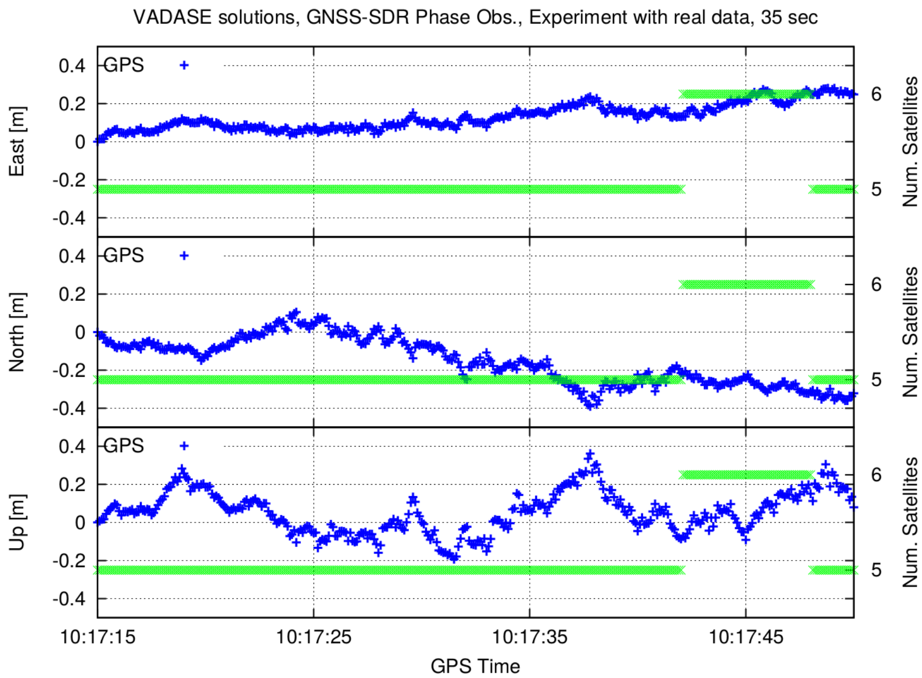


Figure 6: Experiments with real signal: VADASE solutions with GNSS-SDR phase observations over the entire acquisition interval (GPS constellation).

References

- ✔ Arribas J., Branzanti M., Fernández–Prades C., and Closas P. (2014). Fastening GPS and Galileo tight with a software receiver. In Proc. Of the ION GNSS 2014 Conference, Tampa, Florida), Nov. 2014.
- ✔ Arribas J. (2012). GNSS Array-based Acquisition: Theory and implementation (Doctoral Dissertation)
- ✔ Benedetti E., Branzanti M., Colosimo G., Mazzoni A., Crespi M. (2013) VADASE: state of the art and new developments of a third way to GNSS seismology. In Association of Geodesy Symposia Series - VIII Hotine Marussi Symposium on Mathematical Geodesy, Rome, June 2013.
- ✔ Benedetti E., Branzanti M., Biagi L., Colosimo G., Mazzoni A., Crespi M. (2012). Global navigation satellite systems seismology for the 2012 mw 6.1 Emilia earthquake: Exploiting the VADASE algorithm. Seismological Research Letters, 85(3):649–656, 2014.
- ✔ Branzanti M., Colosimo G., Crespi M., Mazzoni A. (2013). GPS near real-time coseismic displacements for the great Tohoku-oki earthquake. Geoscience and Remote Sensing Letters, IEEE, 10(2):372–376, 2013
- ✔ Branzanti M., Benedetti E., Colosimo G., Mazzoni A., Crespi M. (2014). Real-time monitoring of fast displacements with VADASE: new applications and challenges with Galileo. In ENC-GNSS Proceedings, Rotterdam, April 2014.

- ✔ Colosimo G., Crespi M., Mazzoni A. (2011). Real-time GPS seismology with a stand-alone receiver: A preliminary feasibility demonstration. Journal of Geophysical Research: Solid Earth (1978–2012), 116(B11), 2011.
- ✔ Colosimo G. (2013) VADASE: Variometric approach for displacement analysis stand-alone engine (Doctoral Dissertation) Lambert Academic Publishing AG & Co KG
- ✔ Branzanti M. (2015). Variometric approach for real-time GNSS data processing: a new strategy to Earth monitoring and navigation (Doctoral Dissertation)
- ✔ European Union (2010). European GNSS (Galileo) Open Service. Signal In Space Interface Control Document. Ref: OS SIS ICD, Issue 1.1
- ✔ GPS World staff (2014). Galileo position fix with open source software receiver achieved. GPS World, 25(1):29, Jan. 2014.
- ✔ Hoffman-Wellenhof B., Lichtenegger H., and Wasle E.(2008). GNSS Global Navigation Satellite Systems. SpringerWienNewYork, 2008.
- ✔ Fernández–Prades C., Arribas J., Esteve L., Pubill and Closas P. (2012). An open source Galileo E1 software receiver. In Proc. of the 6th ESA Workshop on Satellite Navigation Technologies (NAVITEC 2012), ESTEC, Noordwijk (The Netherlands), Dec. 2012.
- ✔ Fernández–Prades C., Arribas J., Closas P., Avilés C. and Esteve L. (2011). GNSS-SDR: An open source tool for researchers and developers. In Proc. of the ION GNSS 2011 Conference, Portland, Oregon, Sept. 2011.
- ✔ Teunissen P. J. G., Keusberg A. (1996). GPS for Geodesy. Springer

The MEP project: Map for Easy Paths

Ludovico Biagi[1], Sara Comai[2], Florin Catalin Grec[1], Raffaella Mangiarotti[3], Matteo Matteucci[2], Marco Negretti[1], Cecil Ugur[3], Maria Grazia Visconti[1]

[1] Politecnico di Milano - DICA

[2] Politecnico di Milano - DEIB

[3] Politecnico di Milano - DESIGN

Abstract

The aim of MEP (Map for Easy Path) project is to provide a tool based on a mobile device to collect and deliver data concerning obstacles and accessible routes, in order to help physically impaired people to plan their journey.
The mapping obstacles activities will be held with low-cost receiver and the main problem of the first project phase is the accuracy in GNSS positioning of obstacles and the estimation of the optimal placing of them on maps.

Introduction to the project

Get around the city for physically impaired people sometimes can be very easy: wide sidewalks and absence of any kind of obstacle, but sometimes is a problem: many kinds of obstacles are present and the path from one point to another can be very difficult and also can be dangerous.
For physically impaired people it would be very important to know the obstacles before journey start.
The main goal of MEP project (Map for Easy Path) is to provide a tool to map obstacles and accessible routes and delivery it, in order to help physically impaired people to plan their journey and avoid any kind of contingencies.
MEP solution is based on a mobile device with an app to collect and deliver data: this data would be collected in a implicit way, like position data from GNSS satellites and in a explicit way, when the user maps a particular point, gives a description, takes a photo and so on.
We not only collect data, but also the data will be processed and the result will be available, not only for end users, but also for for associations and public agencies that can use it to improve mobility: we want not only map obstacles, but also help to remove them.
MEP project won the "Polisocial Award 2014", the academic social responsibility programme launched by Politecnico di Milano; it is financed by this award and it started in October 2014.
Three departments of Politecnico di Milano are involved:

- DICA, environmental and civil engeneering department to manage GNSS data,
- DEIB, the department of electronic, computer science and bioengineering, to manage the harvesting of explicit data
- Design department, to define how the end users can interact with this

project

In the project are involved also some associations and Cernobbio municipality (see fig. 1. "MEP organization schema")

Figure 1: MEP organization schema.

The problem of correct positioning

When we are mapping obstacles, POI or any other kind of object, the problem is not only to find the correct point coordinates, but it is to find the correct coordinates respect to the map that is used as reference.

In the sample in figure 2 we suppose that the coordinates acquired from GPS (red point) contain an error and we can suppose that we are able to correct him (blue point); also the map can have an error, so may be that, when we put the point on the map, the right position of obstacle respect the map (green point) is different from that we calculated.

Therefore, two main goals can be identified:

- GNSS positioning of obstacles estimation;
- optimal placing of them on maps.

The goal of the first project phase is the accuracy assessment of the common GNSS receivers: to do this we identified a test path in a urban context (figure 3) in order to analyse different conditions of satellite visibility: urban canyons, in the old town, areas of intermediate visibility, outside the old town and open areas, mainly near the lake. We are surveying the test path with low-cost receivers and geodetic instrumentation, that gives us a reference to evaluate the results of low-cost receivers.

Up to now we have collected two different kind of results of our survey: the

point coordinates data obtained from receivers like smartphones and tablet and, using more complex receiver but always low-cost, the raw data, in order to elaborate them and try to obtain a better results. For this elaboration we are testing two open source software: rtkLIB[1] (Takasu & Yasuda, 2009) and goGPS[2] (Realini & Reguzzoni, 2013).
As reference map we will use OpenStreetMap[3]: a collaborative project to create a free editable map of the world. Using this map we can download the vector data and apply our algorithm to match the points survey with the corresponding segment. Moreover we can correct data, enhance with other information, upload them and make them available to all.
The project has just started and in the next few month we will have the first results of this activities.

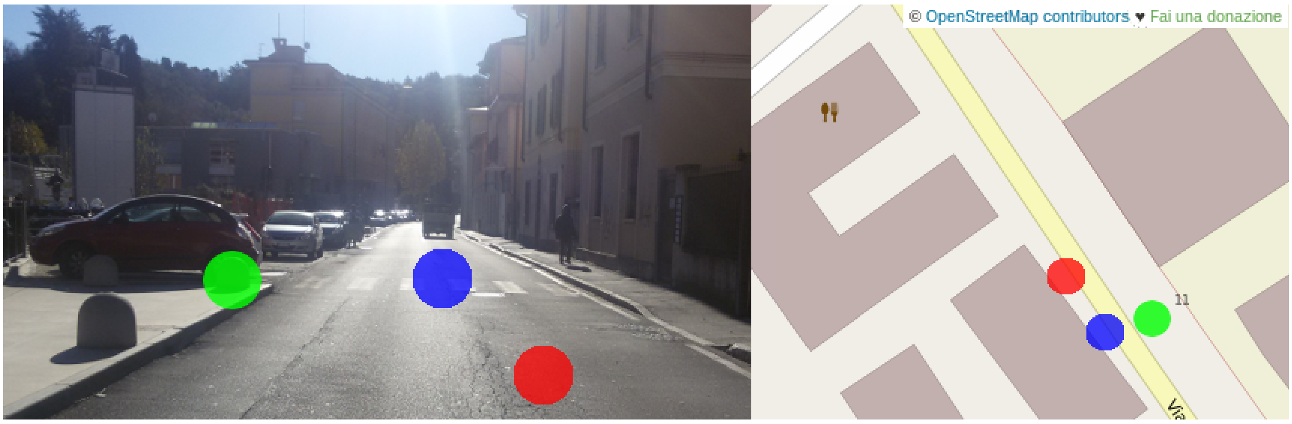


Figure 2: difference between coordinate acquired (red), correct (blue) and the right position respect the map (green).

1 http://www.rtklib.com/
2 http://www.gogps-project.org/
3 http://www.openstreetmap.org/

Figure 3: Test path in a urban context.

References

- ✔ Realini E. & Reguzzoni M. (2013). goGPS: open source software for enhancing the accuracy for enhancing the accuracy of low-cost receivers by single-frequency relative kinematic positioning, *Measurement Science and technology, 24 (11)*, 115010
- ✔ Takasu T. & Yasuda A. (2009). Development of the Low-cost RTK-GPS Receiver with an Open Source Program Package RTKLIB, *Proceedings of International Symposium on GPS/GNSS*.

The first online portal for authoring of indoor mapping data and indoor graphs as IndoorGML standard

Catalin Popa[1], Lucian Brancovean[1], Pietro Elisei[2]

[1] Indsoft Sibiu, Romania

[2] Urbasofia Bucharest, Romania

Abstract

This paper details the first version of the geoportal portal of the EC-funded project i-locate (www.i-locate.eu). The portal will be available to the public for anyone to share indoor mapping data. This has been engineered to act as public portal for indoor open GI and it can be regarded as the indoor counterpart of OpenStreetMap. This will allow easy discoverability, access and sharing through the Internet of open GI related to publicly accessible indoor spaces.

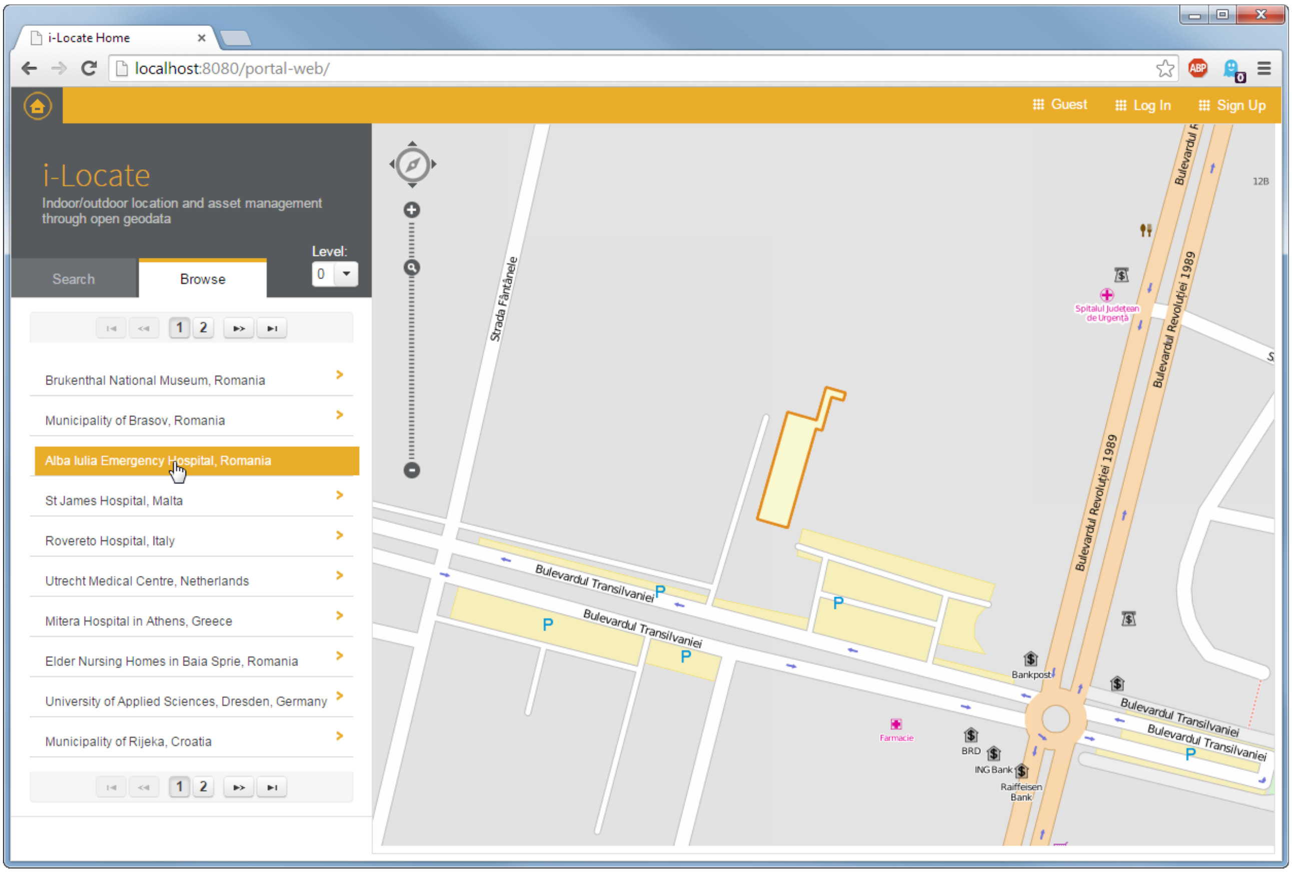


Figure 1: an overview of the i-locate portal.

Keywords

i-locate, indoor, geoportal, mapping, routing, OpenStreetMap

1 System Architecture

The overall system architecture has been designed as a multi-tier Service Oriented Architecture (SOA). At the lower logical end there is the data layer. This provides extensive support for standards. Most notably provides the implementation of a data model ensuring a partial binding to OpenStreetMap – OSM Indoor, the first implementation of IndoorGML standard and its further extension for asset management in line with ISO 55000 standard. The data layer was at the hearth of both the i-locate toolkit, providing support of all the aforementioned standards, and of the i-locate portal.

2 GI Portal

The paper will illustrate the main functionalities of the portal showing how this can be used to create indoor graphs, an essential component of indoor LBS.
The portal is one of the most important outcomes of i-locate. This has been engineered as a “hub” that facilitates the use of open (freely available) geographic data regarding indoor spaces. The open GI Portal allows free publishing and download of mapping indoor geospatial data of public spaces.
The portal functions in a similar manner to an indoor version of OpenStreetMap, allowing access to indoor mapping information regarding publicly accessible spaces (e.g. hospitals’ public areas) which can be available as open data via standard web services for geographical information. To facilitate accessibility, the datasets within the portal may be divided in sections, e.g.: hospital, stations, airports etc.

3 The toolkit

On the other hand, the Portal allows also creation of new indoor mapping data starting from existing files (e.g. from floor plans) through a number of web-based tools that will be available as a result of the project. This will be a valuable tool to facilitate creation and sharing of new data to the public.

4 Operating modes

The portal operates in two modes, staging and preview mode. A log-in mechanism allows admin users to authenticate to the system using a username and password. After a successful login, the admin user will be taken to the site status and operations screen.
The site status and operations screen display the maps currently available for the site administered by the logged in user. Within this part of the portal there is a list with the maps in staging, and a list with the maps in production. A command allows pushing all the maps currently in staging in production mode, replacing those already in production. The interface allows also to remove maps from staging, and to upload a new map or a new version of an already existing map.

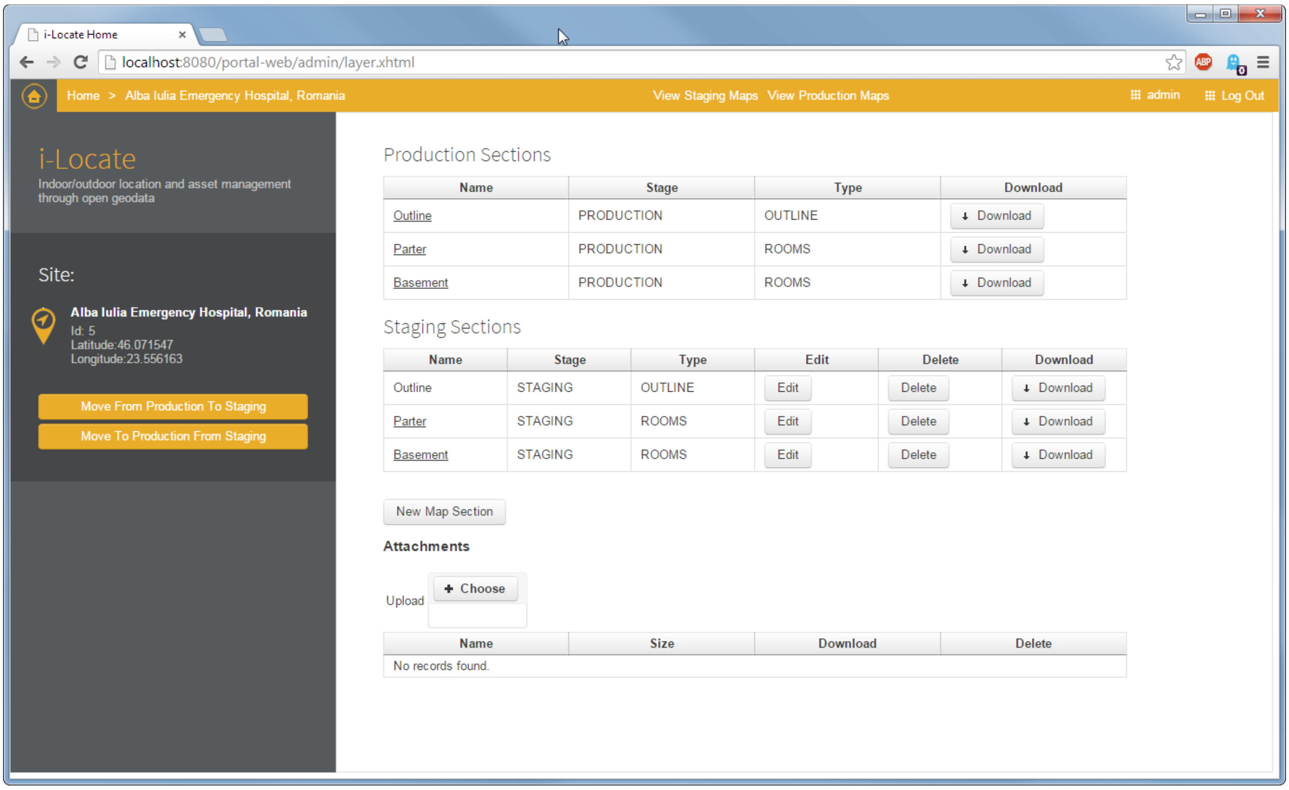


Figure 2: the staging interface.

The admin user can see a preview of the maps, in staging mode, as they will appear in production once promoted. This screen shows the staging maps for the current site, with the navigation graph overlaid. It should be noted that at this stage, if the user moves the map to see a different site, he/she will see the production maps. The staging maps are only visible for the user's own site.

5 Navigation graph

It should be noted that this is a navigation graph which is interactively created through the portal and which is encoded, internally, as IndoorGML graph. This is the recent standard by the Open Geospatial Consortium that allows defining graphs used to support indoor navigation (in several modalities) for the creation of indoor Location Based Services.
A specific section of the portal allows fully interactive creation of the indoor graphs, in a very simple manner, by having the user interactively click on indoor maps to create the different nodes and edges in a visual way.

6 Platform and standards used

The implementation of a the portal has been essentially developed through maps control within a HTML page which have been implemented with the OpenLayers JavaScript library, which supports standards such as OGC WMS. This allows combining data from i-Locate, Open Street Maps and other sources.
The map control obtain the map data to display from two sources: the mapping services behind the i-locate service infrastructure.

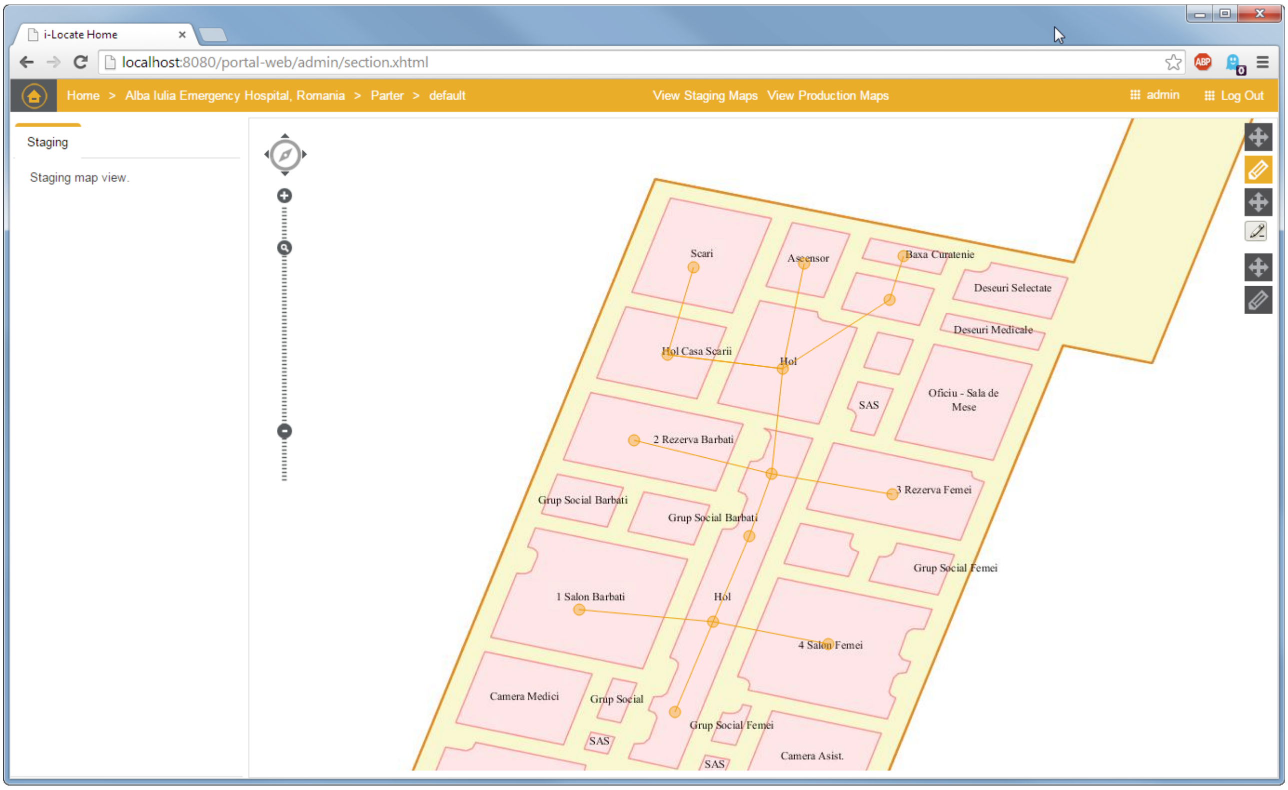


Figure 3: the indoor graph.

i-locate - Indoor/outdoor LOCation and Asset management Through open gEodata

Stefano Piffer, Nicola Dorigatti, Giuseppe Conti

Trilogis srl

Abstract

The main objective of i-locate project (Indoor / outdoor location and Asset Management Through Open Geodata) is to simplify the life of users by helping them to "navigate" inside buildings, using their smartphone, to reach their place of destination. i-locate also allows tracking objects and portable equipment within "indoor" spaces to improve management and maintenance.
"i-locate" will last three years starting from 1 January 2014, and is co-financed by the European Commission through the ICT-PSP programme. The project involves 24 partners from 9 European countries (HR, GB, DE, GR, IT, LU, MT, NL, RO).

Keywords

Indoor, location, asset management, RTLS

1 The project

Recent studies have highlighted that, on average, we spend approximately 90% of our time indoors, often in unfamiliar environments. Nevertheless, Location Based Services (LBS), for both business or personal use, today are essentially limited to outdoor scenarios and are based virtually only on GPS location. Clearly, being able to seamlessly locate people or objects within indoor AND outdoor spaces could enable a number of new Location-Based Services (LBS) of significant economic relevance in domain such as logistics, personal services, mobility, retail, health, to name but a few.
In particular domains, such as logistics or in modern hospitals, the availability of asset management & maintenance platforms with integrated RTLS (Real-Time Location Systems) features becomes an significant service enabler and business driver, thanks to the real time control of where each piece of equipment is located. The derived information can be then used to infer usage patterns, derive indicators essential to assess actual levels of use, to increase safety, to reduce maintenance errors and, ultimately, to lower costs. In some specific domains, such as healthcare, being able to provide real-time efficient and accurate "asset" tracking and management (be this medical staff, patients, visitors, equipment, etc.) is extremely important in economic as well as in social terms.
All market analysts agree that market of indoor LBS is set to grow in the next 4 years to a multi-billion market. Players such as Google, Apple, Nokia, as well as a multitude of SMEs, are investing significant resources to create indoor mapping databases and services. Technology, from a hardware point of view, has already reached full maturity and has consumer-grade costs. In fact, over the past few years, increasingly accurate indoor localisation technologies,

based on technologies such as Bluetooth, ZigBee, and Wi-Fi, have expanded the scope of LBS to include indoor spaces.
Nevertheless, the development of true seamless indoor/outdoor location services still requires removing three major barriers:

1. Lack of Geographic Information (GI) indoor spaces. While outdoor data can be easily accessed as Open Data (OD), a notable example being OpenStreetMap (OSM), the availability of Geographical Information of indoor spaces is not available on a large scale as Open Data. In case of publicly accessible buildings, such as hospitals, stations, airports, shopping malls, and public offices, having access to geographical data of indoor spaces, and particularly as Open Data, could allow new business activities and bring a number of social benefits.
2. Lack of technological ecosystems. Technological fragmentation essentially bind location based services to specific native technologies, limited to specific functionalities. No ecosystem allow leveraging on multiple (concurrent) location technologies for indoor and outdoor scenarios to create innovative Location Based Services based on routing and asset management.
3. Lack of support to indoor/outdoor LBS by current GI standards. This limitation does not allow mixing technologies or solutions from different vendors. i-locate tries to overcome the aforementioned three barriers so as to enable the creation of innovative businesses based on indoor and outdoor localisation of people and objects. To this extent, i-locate has been engineered to address very clear market requirements emerging from a core of specialised SMEs, targeting the following objectives:
    - To create a public geoportal, the so-called "virtual hub", that collects, makes discoverable, and provides access to indoor geographical information as Open Data. To this extent, the portal will represent an indoor counterpart to OpenStreetMap.
    - To adopt and extend of open standards in the field of indoor/outdoor LBS.
    - To create an open source "toolkit" (technically a middleware) for software developers, that allows integrated indoor-outdoor LBS based on the aforementioned open standard.
    - To ensure that such a toolkit is built on sound privacy and security policies, for the highest protection of personal/critical data.
    - To develop template of client software for mobile devices (App), addressing real world scenarios, for LBS that use the toolkit's services via the aforementioned open standard protocols.
    - To test -for more than one year- the "virtual hub", the "toolkit," and the final applications in real operational scenarios with real users and stakeholders within 14 pilots sites in 8 EU countries.
    - To stimulate innovation and business activities around indoor Geographic Information through direct engagement of industrial players and SMEs, to ensure that the relevant critical mass is created around the results of i-locate and to foster development of innovative services.

The activities of the project started with a very comprehensive analysis and design phase, the consortium engaged in a very in-depth analysis of the various use cases to be addressed by the pilots and of the requirements

emerging from the real downstream users (e.g. technicians, doctors, public officers, etc.), as well as in a very precise definition of the exact geographical scope (both outdoor and indoor) of each pilots. Such an attentive pilot-driven work, which saw the engagement of both technical and pilot partners, was complemented by a very accurate analysis of regulatory constrains at international, national, and local level, paying particular attention to privacy, ethical, and security issues.
Concurrently, from a more technical description, the consortium has identified the various technical requirements emerging from various technology providers. This phase greatly benefitted from the involvement of few key experts from the industrial and standardisation world, in particular from OGC - Open Geospatial Consortium (the reference standard organisation in the geospatial domain), who were formally engaged as members of the advisory board. As a result, a key decision was made regarding the best standard on top of which later development should be based.
The final choice fell on IndoorGML, at that time a standardisation candidate from OGC specifically designed for indoor Location Based Services, which was scheduled to be officially approved as standard in the course of 2014 (eventually officially voted as standard in September 2014). It soon become clear that the current (modular) core of the standard would have to be extended with an additional module, to be proposed as a result of i-locate, for location based services specifically designed for asset management. It was designed, in turn, that such a module should be designed following the principles of the very recent (2014) ISO 55000 abstract standard specifications on asset management. Further standards selected for some of the core features included, among other, support for ISO/IEC 24730-1:2014 standard on "Information technology — Real-time locating systems (RTLS)". Such a very strong attention to standards was essential to ensure that the result of the project could have a wide industrial update.
This choice, was borne out of real market requirement as perceived by the companies involved in the consortium. This initial phase set the foundation for the definition of a comprehensive system architecture, comprising a number of very complex services, from localisation to geofencing, from generation of analytics to spatial services, etc. At this phase, it become clear that i-locate would have the unique opportunity to become the first reference implementation of IndoorGML, a very new and innovative standard within the promising LBS market.
Additional, earlier activities, also included a detailed survey of the data available to the various pilot sites, their adjustments and loading into the central database that in the meantime was being deployed as part of on-going development of the core "virtual hub" (the geoportal).
The latter has been progressively extended with a number of "connectors" required to provide real-time access to other open data repositories and with a specifically-designed service that allowed geometrical and topological validation of data against the specifications of IndoorGML. The first core of virtual hub was progressively extended with graphical user interface capabilities that allow ingestion of indoor mapping data as well as interactive drawing of indoor graphs encoded as IndoorGML. The first tangible result has been the early public deploy of a first version of the i-locate portal, which can be already used for the pilot preparation activities.
Meanwhile, after the design phase was concluded with the release of the

system architecture, the software development activities started. The result of this phase has been the successful release to the public of the first version of the toolkit that marked the end of the first year, as open source solution.
During the first year the consortium has also engaged in several preparatory activities of the pilots, including definition of detailed deployment and validation plan as well as in the creation of a proper infrastructure for online training (available from the project website) and related interactive training material. This will be particularly important to ensure wide uptake of technologies within the wide community of stakeholders and users.

References

✔ http://www.i-locate.eu

Indoor location and sensor technology to improve quality of life and independency of elderly people

Leonardo Plotegher, Giuseppe Conti

Trilogis srl

Abstract

UNCAP (Ubiquitous iNteroperable Care for Ageing People) will develop an open, scalable and privacy-savvy ICT infrastructure designed to help aging people live independently while maintaining and improving their lifestyle. UNCAP will leverage on commercial biosensors and localisation solutions to monitor –in a non-invasive way- and assist the users.

UNCAP will last 3 years starting from Jan. 2015. The project is co-financed by the EU through the Horizon 2020 programme and it involves 23 partners from 9 countries.

Keywords

Ageing, assisted living, localisation, biosensors

1 The project

UNCAP maximises use of indoor location and geographical standards to create an interoperable open source infrastructure that leverages on location and sensor-based technologies to create radically new paradigms for service care delivery, aiming at:

- Improve effectiveness of the health care processes through more effective evaluation processes during the hospital-hospice recovery.
- Enhance home care treatment and prevention, in order to delay cognitive impairment of elderly and possibly postpone the necessity of recovery at hospitals or hospices.
- Support more independent living and improve quality of life and dignity of cognitively impaired aging users by helping them be more independent and for longer time.

From a technical standpoint the five pillars of UNCAP are:

- Interoperability and use of open standard, through support for a range of open standards from the Geospatial Consortium (OGC) from the European SDOs (CEN, CENELEC and ETSI) for all its key services (e.g. position, sensors, building automation systems, clinical assessment, storage of clinical data etc.) allowing for future extensions in terms of hardware and software.
- Openness, through release of open specifications and open software components.
- Scalability, through use of cloud-centric approaches.
- User friendliness, ensuring compliance with all most common usability standards (e.g. Web Accessibility Initiative - WAI or ISO/TR 16982:2002).

- Privacy and security, through attention to all related privacy and security aspects.

In practice, UNCAP will deliver an open source suite comprising of a low-cost Android-based unit, called the "UNCAP BOX", to be connected to standard TV sets, which collects data coming from different indoor and outdoor localisation technologies (e.g. sensor flooring, camera-based detection system, etc.) as well as form sensors (e.g. measuring vital parameters, environmental temperature etc.), and makes them available - via secure communication channels- to the "UNCAP CLOUD".

The UNCAP BOX will also be used as the interface for end users, caregivers and family members who will be able to communicate (also via video conference), exchange health data (via HL7 standard), perform assessment of the patients conditions (through InterRAI™ assessment tools and methodology) as well as to place emergency calls. Most interestingly, the UNCAP BOX will support interoperable communication, via KNX open protocol, with building automation systems which can be not only controlled by UNCAP services but can also be used as further feedback channel (e.g. by flashing lights to enforce a reminder, or turning lights on to a given room the user wants to reach).

The UNCAP BOX will be also complemented by an App (UNCAP App also referred to as "UNCApp") for smartphones or tablets, to provide a further convenient portable access to UNCAP services as well as to allow for access to selected UNCAP services in online as well as offline mode or from locations that are not instrumented with UNCAP infrastructure.

Both the UNCAP BOX and CLOUD will be designed based on open specifications, will be developed as open source software and will use open standards (i.e. SWE and SensorThings by OGC for sensor and location information, HL7 for health data and KNX for communication with building automation systems) to deliver the following set of value-added services.

The technical solutions will be extensively piloted for a total of 24 months (made of two sets of 12 months), within 14 sites, across six countries, carried on in real operational scenarios (instrumented apartments, nursing homes, elderly care canters and private homes), involving a statistically significant number, almost a thousand, between final users and caregivers, which have been selected for their diversity in order to cover a wide spectrum of practices and care models. In particular, a first 12 month-long assessment of the pilot sites (including economic aspects) and its users' physical and cognitive performances (through InterRAI™ scales), will be carried on prior to the deployment of UNCAP in the pilot sites (to assess "as is" conditions). A second 12 month-long assessment will be then repeated after UNCAP will be installed at the pilot site. This will enable collection of very clear evidence of the impact in terms of improvement of the efficiency of the health and care system and quality of life, based on statistically relevant and scientific grounding.

References

✔ http://www.uncap.eu

Local monitoring by low cost devices and free and open sources softwares

Ludovico Biagi, Florin-Catalin Grec, Marco Negretti, Maria Grazia Visconti

Politecnico di Milano, DICA@ComoCampus

Abstract

The purpose of local monitoring is to closely model displacements, evolution and deformations of civil engineering structures and model local phenomena like landslides. For certain applications, latency of the monitoring results is not a crucial point. In other cases a small latency is fundamental: for example, a suspected Alpine landslide that insists on villages. In these cases, contiunous quasi real time monitoring by permanent stations is required. Here, a typical monitoring configuration is given by a GNSS reference permanent station, outside the monitored area, in a stable site and several permanent stations in the monitored area. Data can be processed and the results can be monitored in a quasi continuous way by automated procedures. In these situations, the use of low cost GNSS receivers could be interesting because it can significantly increase the number of monitored points without notably increasing the network cost. This investigation is the scope of our work and several experiments have been set up and performed to assess them. To reach our goals, geodetic and low cost receivers are teamed up: Leica GR10 as reference station and several single frequency u-blox 7P receivers. Once the aquisition phase is completed data are processed with commercial (Leica GeoOffice), academic (Bernese) and free and open source softwares (goGPS and RTKLIB). The results are compared in the final step of the project, accuracy and reliability of solutions are discussed. Their results will be described in the presentation and the full paper.

Keywords

Low cost receivers, monitoring, deformations, landslides

1 Introduction

Several experiments are planned and will be performed in the next months. At the present, a first experiment is under analysis in order to assess whether or not is it possible to use u-blox receiver for precise geodetic applications. On the roof of the Politecnico of MILANO building a week of u-blox static data has been collected, from the GPS day 074 to GPS day 079 for the year 2014 (from 15 to 20 of March 2014). The place chosen for the experiment is about 60 m distant from Milano GNSS permanent station. The data have been acquired at 1 second sampling rate.

2 Results

Milano PS coordinates have been estimated in ITRF2008. The reference coordinates for u-blox receiver have been calculated with LGO software in post processing with respect to MILANO permanent station. The following processing parameters have been adopted: 10 degrees cut off angle, Klobuchar ionospheric model, Saastamoinen tropospheric model and broadcast ephemerides.
Reference UTM coordinates for u-blox station are the following:

E = 517 880.331 m; N = 5 036 253.753 m; h = 193.380 m

Then, u-blox data have been processed in LGO by hourly static sessions to assess their accuracy. Therefore, 168 results have been obtained. First of all, during the processing 4 blunders, whose results differ more than 5 cm from reference solution, were identified for the whole session. For these solutions, ambiguities have not been resolved and errors' magnitude is of about 20 cm, the maximum reaching 1 m for East component. Table 1 provides a more detail review of the outliers influence in the final 3D position of u-blox station.

Session	dS [cm]
34	37.1
58	12.3
106	129.9
160	38.9

Table 1: Vectors of residuals for outliers.

At a second step, the same procedure was carried in Bernese. Bernese processing parameter are set up according to guidelines for single frequency, short baselines: 10 degrees cut off angle, dry Neill tropospheric model, IGS ionosphere model, precise ephemerides double difference on phase (L1 carrier), Sigma method for ambiguities handling.
At the time of this contribution only preliminary results are ready. The results are quite similar with those from LGO but 5 blunders are present, probably caused by problems in solving of cycle slips: sessions 22, 24, 30, 40 and 142. By averaging the hourly solutions, except those affected by blunders, Bernese gives as final solution the exact reference coordinates. Table 2 reports the results of the nor bludered sessions, both for LGO and Bernese.

	Leica GeoOffice			Bernese		
Statistics	**dEast [m]**	**dNorth [m]**	**dh [m]**	**dEast [m]**	**dNorth [m]**	**dh [m]**
Mean	0.0002	-0.0009	0.0019	0.0009	-0.0015	0.0008
RMS	0.0018	0.0025	0.0044	0.0021	0.0049	0.0045
Max	0.0041	0.0033	0.0145	0.016	0.047	0.0143
Min	-0.004	-0.0089	-0.008	-0.0031	-0.010	-0.010

Table 2: Comparison between statistics of residuals from LGO and Bernese.

Below, the graphical representation of the residuals of LGO in the 3 components is provided without blunders to help visual interpretation of results. They prove that u-blox receivers are suitable for applications with cm-level accuracy: for the 2D position the errors are well below 1 cm in most of the cases. However, for height, as expected, the residuals are slightly higher but still without moving away from the cm-level accuracy (around 1.5 cm the maximum residual).

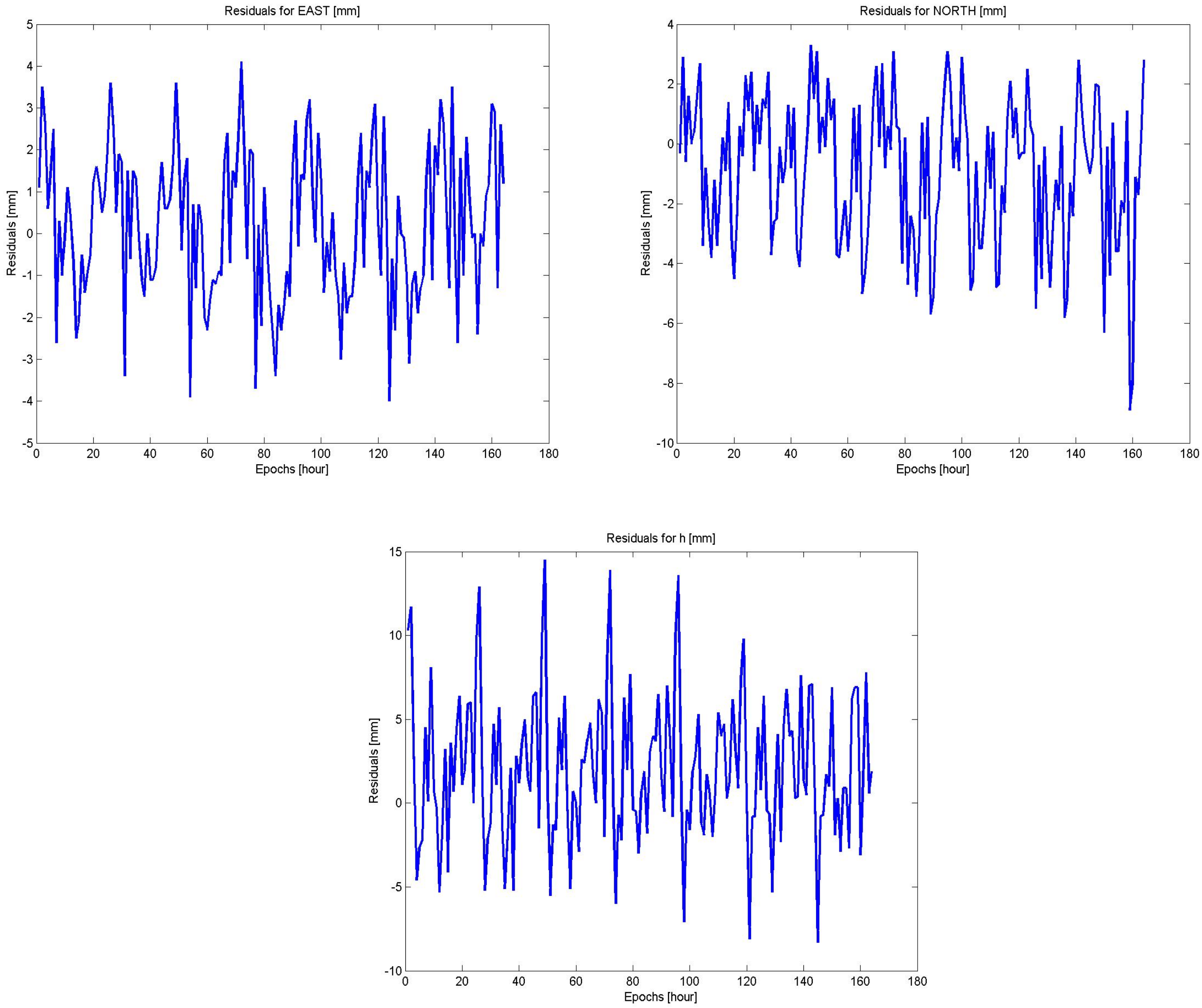


Figure 1: Time series for residuals (LGO processing), static method, 1 hour sessions – East (blue), North (yellow) and h (red).

These results were confirmed by Bernese processing. Again, for the horizontal components the residuals are slightly lower than for the vertical component and the number of blunders is smaller. Float solutions are graphically compared with the fix ones. From doing so, it can be easily seen the impact of ambiguities on the final results. Once this step was fulfilled, in good sessions the residuals shrunk from cm level to mm level. On the contrary, for problematic sessions, fixing worsenes the solution because the residuals increase with respect to float solution. These problems are presented in Table 3.

Session	dN [cm]	dh [cm]
21	-	From -18.1 to -22.4
23	From 37.7 to 39.7	From -15.0 to -15.4
40	-	From -2.47 to -5.9
142	From 9.5 to 10.1	-

Table 3: Change in residuals magnitude after fixing ambiguity for problematic sessions.

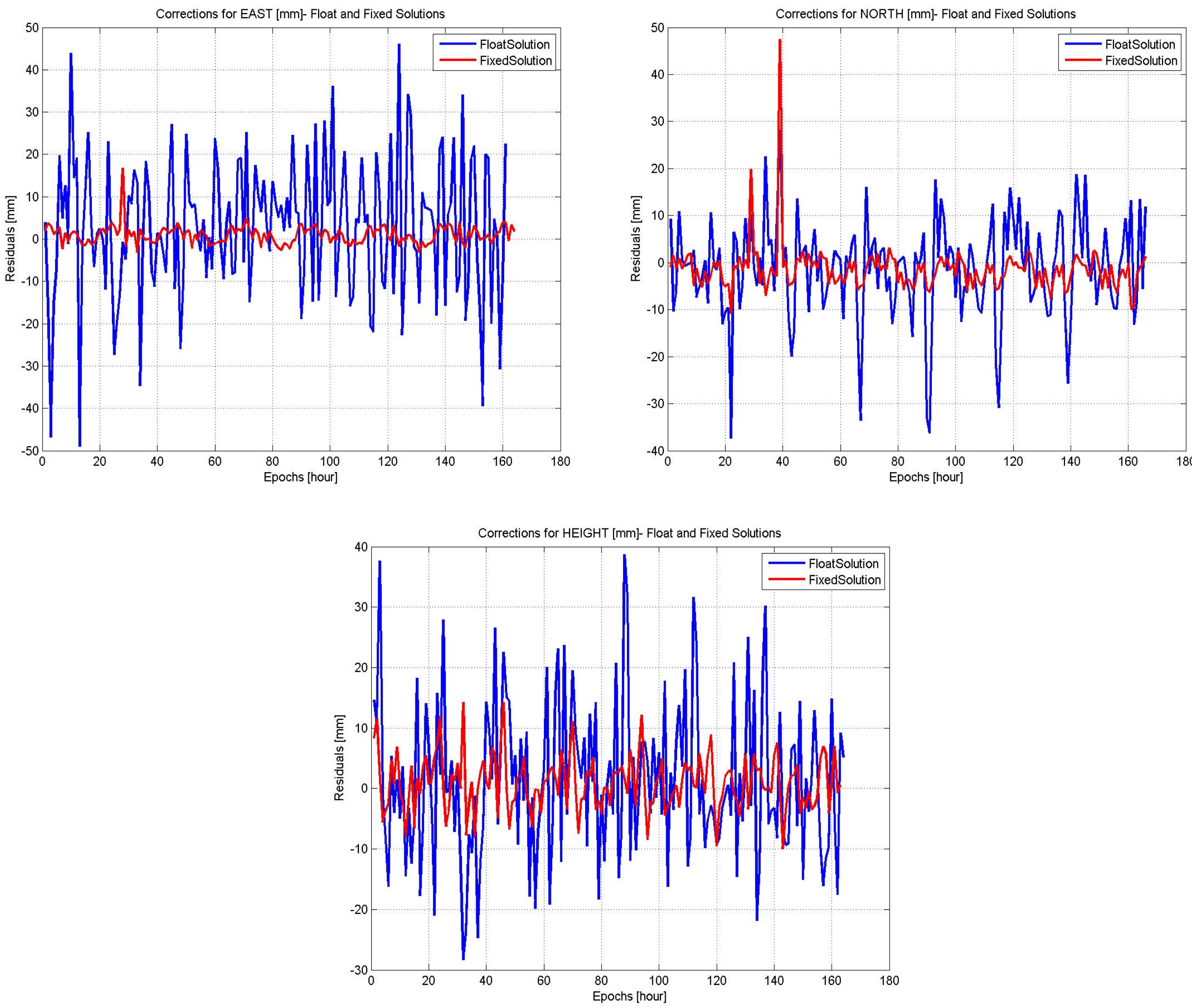


Figure 2: Comparison between residuals for float and fix residuals (Bernese processing), static method, 1hour sessions.

At the moment of writing the abstract, the data are also being analysed with two free and open source softwares, goGPS and RTKLIB, and the results will be compared with LGO reference results. goGPS is a software package developed at the Geomatics Laboratory of Politecnico di Milano designed to improve the positioning accuracy of low-cost (single-frequency) GPS devices by RTK technique. Currently it requires raw data (observations) at 1 Hz in input. RTKLIB is an open source program package for standard and precise positioning with GNSS (global navigation satellite system). It supports standard and precise positioning algorithms with: GPS, GLONASS, GALILEO, QZSS, BEIDOU and SBAS.

It supports various positioning modes with GNSS for both real-time and post-processing: Single, DGPS/DGNSS, Kinematic, Static, Moving-Baseline, Fixed, PPP-Kinematic, PPP-Staticand PPP-Fixed.
The results will be presented in the final paper.

3 Conclusions

In general, low cost receivers like navigation type or chip sets use the phase only to smooth the code observations but some manufactures like u-blox offer the option to extract the phase data.
This first experiment has been performed as a preliminary step just to test the accuracy and reliability of u-blox receivers for monitoring purposes. The investigations reveal that up to now the post-processing of phase observations meet the required accuracy for monitoring of local phenomena. Therefore, in the next months experiments more appropriate to landslides monitoring are going to take place.

References

- ✔ Dach, R., Hugentobler,U., Friedz, P., Meindl, M. (2007, January). *Bernese GPS software.* Astronomical Institute, University of Bern
- ✔ Hongki, J., & Sung-Han, S., & Tatakowski, A., & Spencer, B.F., & Nelson, M. (2012). Feasability of displacement monitoring using low-cost GPS receivers. *Structural Control Healt Monitoring,* doi: 10.1002/stc.1532
- ✔ Realini, E., & Reguzzoni, M. (2013). goGPS: open source software for enhancing the accuracy for of low-cost receivers by single-frequency relative kinematic positioning. *Measurement Science and Technology,* 24 (11), 115010
- ✔ Schwieger, V. (2009, September). High-sensitivity GPS: an availabilty, reliability and accuracy test. *INCT Bulletin des Sciences Geographiques* (No 23), Institut National de Cartographie e de Teledetection (INCT), Algier, Algerien.
- ✔ Takasu, T., & Yasuda, A. (2009). Development of the Low-cost RTK-GPS Receiver with an Open Source Program Package RTKLIB, *Proceedings of International Symposium on GPS/GNSS.*

ADO: a prototype of African Drought Observatory for drought monitoring and forecasting

Luana Valentini[1], Paulo Barbosa[1], Diego Magni[2]

[1] Joint Research Centre of the European Commission, JRC, Ispra, Italy

[2] Arcadia SIT, Vigevano, Italy

Abstract

The African Drought Observatory (ADO) has been developed based on the experience of the European Drought Observatory (EDO) and on several tested drought products. The Map Viewer makes available drought-relevant information provided as maps of drought indicators derived from different data sources (e.g., precipitation data, satellite measurements, meteorological variables and socio-economic indicators). The products derived from these data sources are made freely available through a web-based Map Viewer and are a complementary source of information for stakeholders that would like to use them to monitor and forecast droughts in Africa as well as to understand which areas are more vulnerable to drought.

Keywords

Africa, drought, webGIS, drought monitor, drought forecast

1 Introduction

Detecting and analysing drought conditions and their evolution is a topic of prime interest for any drought management system. Although, due to its complexity, there is no universal criterion to detect drought and describe its severity (Wilhite, 2000a), there are a number of accepted indicators and methods that can be used to monitor and forecast droughts (Heim, 2002, Mishra and Sing, 2010).

Since there are different indicators and methods to monitor drought, a common accepted approach is to make this information available into a single drought information system (Svoboda et al., 2002). Such a system should be the main input for the implementation of a drought monitoring and early warning system to assist end users in making objective decisions for drought risk management, including mitigation and adaptation (Wilhite et al., 2000b). Some examples of such systems worldwide are the European Drought Observatory (EDO[1]), the United States Drought Monitor[2] or the North American Drought Monitor[3] .

Over the African continent there are a few valuable efforts in developing

1 http://edo.jrc.ec.europa.eu

2 http://droughtmonitor.unl.edu/

3 http://www.ncdc.noaa.gov/temp-and-precip/drought/nadm/

drought monitoring systems. Examples are the Famine Early Warning Systems Network (FEWS NET[4]), that combines local socioeconomic information with agricultural production and precipitation information to predict food security conditions (Funk, 2009), and the African Flood and Drought Monitor[5], maintained by the Land Surface Hydrology Group at Princeton University, that provides near real-time drought monitoring for Africa using the variable infiltration capacity (VIC) hydrological model and a long-term retrospective meteorological reanalysis (Sheffield et al., 2008).
A drought African Map Viewer, called ADO[6], was built based on the experience of the EDO, established by the Joint Research Centre (JRC) of the European Commission (EC), and on the EU funded DEWFORA project[7]. This web application includes a number of developed and tested drought products and provides historical and near-real time monitoring information, as well as seasonal forecasts describing meteorological drought. It also includes two different indicators carrying information on drought vulnerability and risk, respectively at country and water basin level. In the future, it should also contain information on hydrological drought forecasting that can be used to complement the above mentioned drought information.
In this paper the main features of the Map Viewer are presented from the point of view of its products development and functions.

2 Methodology

2.1 Products development

The African Map Viewer relies on several variables and indicators that are relevant for drought monitoring, forecasting and decision making. A schematic diagram of the drought monitoring and prediction system incorporated into the Map Viewer is presented in Figure 1.

4 http://www.fews.net/
5 http://stream.princeton.edu/AWCM/WEBPAGE/interface.php
6 http://edo.jrc.ec.europa.eu/ado/
7 http://www.dewfora.net/

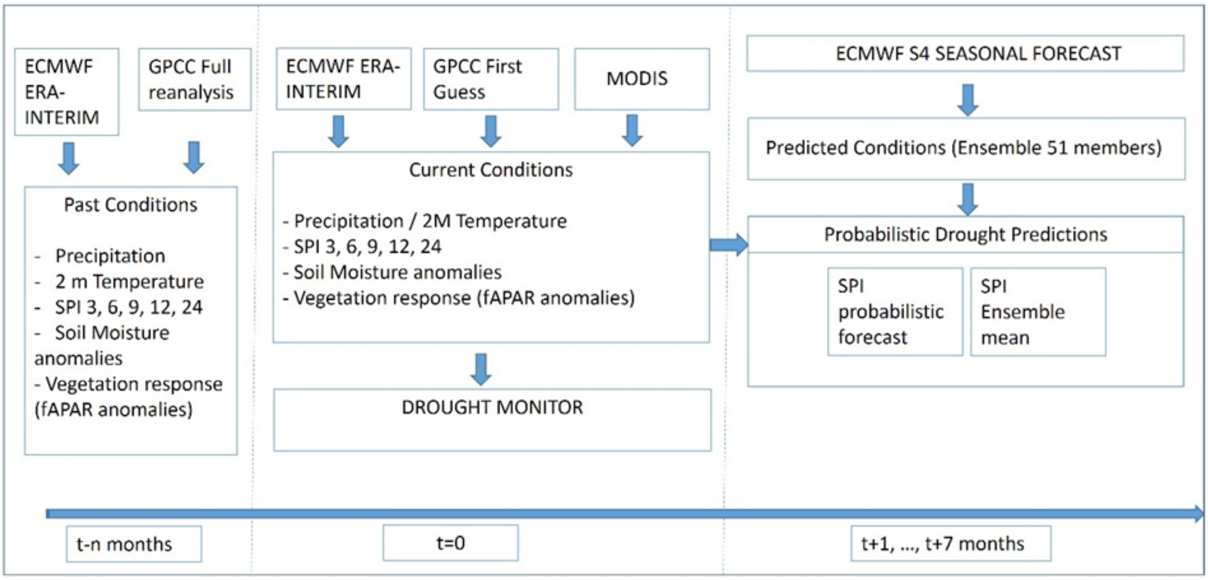


Figure 1: Schematic diagram of the drought monitoring and prediction system.

The monitoring information comes from the Global Precipitation Climatology Centre (GPCC[8]) and the European Centre for Medium-Range Weather Forecasts (ECMWF[9]), while the forecasts from ECMWF only.
The database stores a reference grid composed by 1 decimal degree cells covering the entire world. Each cell has its unique identifier (ID) that permits to assign the variable values to a specific place on the globe by simply combining the spatial data with the stored values.
The main monitoring products are the precipitation and their derived indicators, such as the precipitation anomalies (% of normal) and the Standardized Precipitation Index (SPI). The GPCC and the ECMWF atmospheric reanalysis ERA-Interim precipitation datasets are used as the main sources of precipitation. ERA-Interim monthly temperature and soil moisture including their anomalies are also displayed. The vegetation conditions are assessed through the anomaly of fAPAR (fraction of Absorbed Photosynthetically Active Radiation; Gobron et al., 2007) 10-day time composite derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument. The indicators and datasets were selected after having been tested at regional and continental scale (Dutra et al., 2013 and Naumann et al., 2013a).
Drought vulnerability and risk are assessed at country and basin level following a socio-economic approach (Naumann et al., 2013b). The methodology used to compute the Drought Vulnerability Index (DVI) integrates both quantitative and qualitative characterisations of drought vulnerability at different spatial scales.
The meteorological seasonal forecast relies in the ECMWF dynamical ocean-atmosphere coupled system 4 (S4) (Molteni et al., 2011). The seasonal forecast generates 51 ensemble members in real-time, and 30 years (1981-2010) of back integrations (hindcasts) with 15 ensemble members. The forecasted lead time is 7 months, including the month of issue. The two main outputs are the forecasted SPI following a probabilistic or a deterministic approach. In both

8 http://gpcc.dwd.de/
9 http://www.ecmwf.int/

cases the SPI is computed with GPCC data as monitoring product merged with the S4 forecasts (Dutra et al., 2013 and Dutra et al., 2014).
A number of geographic background data sets are also available such as climate classifications (Aridity Index and Köppen-Geiger Classification; see Spinoni et al. (2013) for further details), land cover and soil maps (Land cover, Agricultural Management Factor, Soil Fertility Index), water resources (Irrigation cropping pattern zones, Surface water bodies and River Basins), and population indicators (Population Density and Infant Mortality Rate). Many of them are loaded as Web Map Services distributed by the Food and Agriculture Organization (FAO) of United Nations (UN) and only visualized in the Map Viewer. The main advantage of the services is that it is not necessary to store and maintain this data in our server.

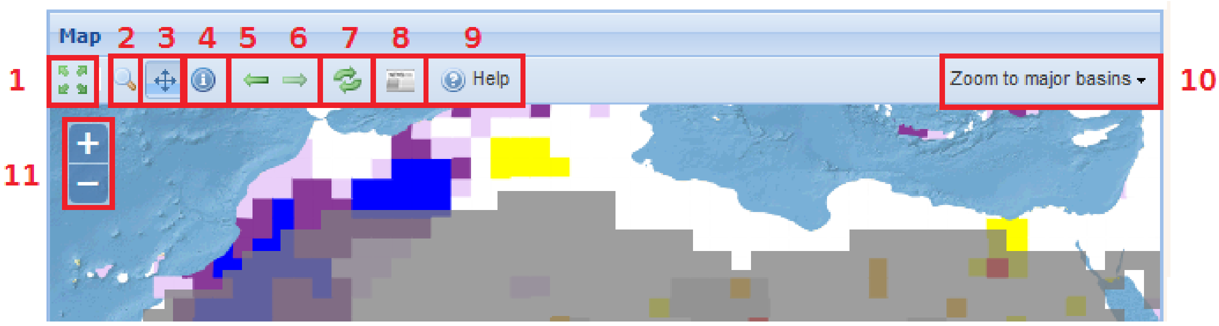
2.2 System architecture and Map Viewer features

The Map Viewer is a web Geographic Information System (GIS) available at http://edo.jrc.ec.europa.eu/ado/ado.html.
The source of geographical and thematic contents used for the maps and charts are different: locally stored raster images and ESRI shapefiles, Open GeoSpatial Consortium (OGC) Web Map Services (WMS) layers, and Oracle spatial tables.
OSGeo MapServer 6.4.1[10] is adopted to provide maps as Web Map Services (WMS).
Layers are added to the map using OpenLayers 2.13.1[11], a pure JavaScript library for displaying map data in web browsers, with no server-side dependencies. OpenLayers implements a JavaScript API for building rich web-based geographic maps and it is free software, developed for and by the Open Source software community.
To build up the Map Viewer, OpenLayers was not enough and GeoExt 1.1[12] was adopted too. GeoExt brings together the geospatial know how of OpenLayers with the user interface knowledge of Ext JS to help to build desktop style GIS apps on the web with JavaScript. Since our data structure and time dependency is quite complex, some function of Ext JS 3.4[13] were adopted too, in order to extend the capabilities of GeoExt to our needs.
Several PHP classes and programs jointly with JavaScript functions were implemented to retrieve data from the database, draw charts of the main drought indicators and get GeoRSS news feeds.

2.2.1 Map Toolbar

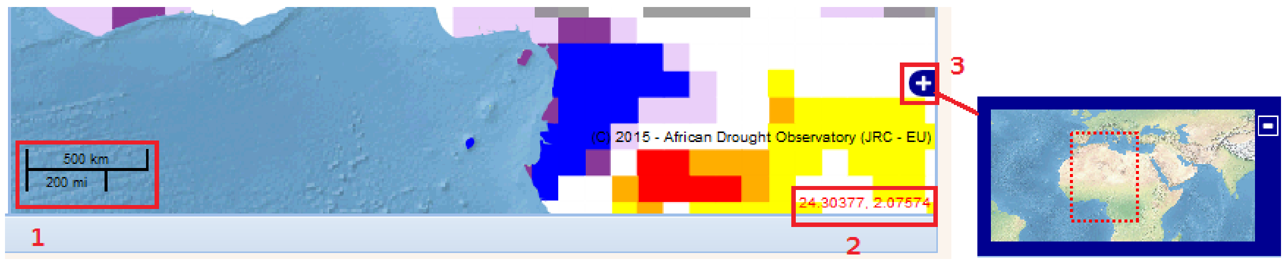
Basic tools are available in the toolbar and in the footer of the map panel (see Figure 2 and Figure 3).

10 http://mapserver.org/
11 http://openlayers.org/
12 http://geoext.org/
13 http://www.sencha.com/products/extjs/



1 - zoom to the maximum extent
2 - zoom box
3 - pan
4 - identify feature
5 - previous zoom in history
6 - next zoom in history
7 - refresh
8 - EMM GeoRSS news
9 - help
10 - zoom to major basins
11 - zoom in / out

Figure 2: Map panel upper toolbar.



1 - current map scale
2 - geographic coordinates of the mouse pointer
3 - overview map

Figure 3: Footer toolbar.

In particular, the "Identify" tool can be very useful for the final user. When acting on a time dependent monitoring layer, in fact, the information regarding the specific time stamp selected for the layer is shown.

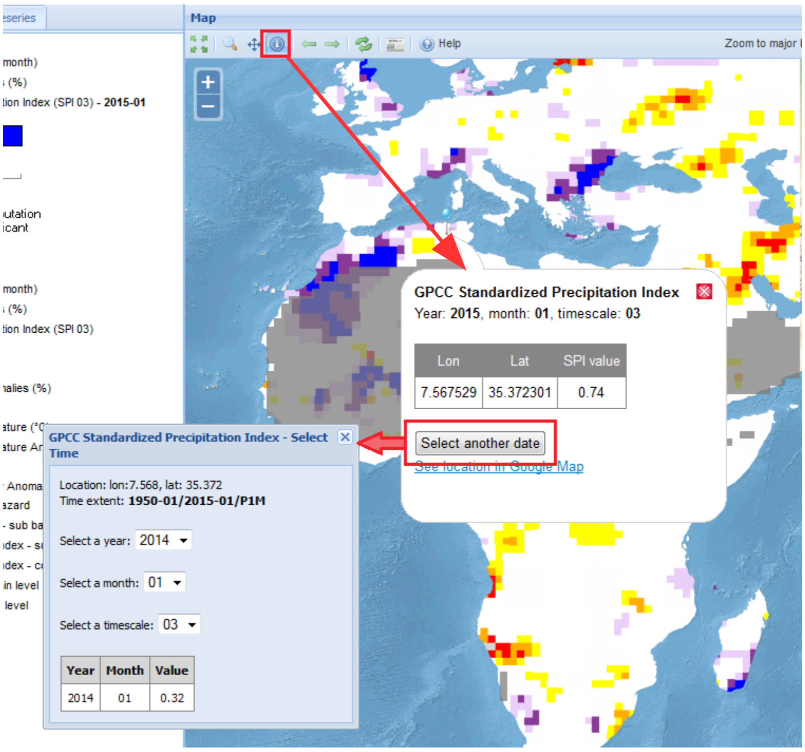


Figure 4: Example of "Identify" on a time dependent monitoring layer. In this case it is possible to compare the value of SPI3 of the same cell for January 2014 and 2015.

In Figure 4 an example is reported for a time dependent layer, such as SPI. The information regarding the specific time stamp selected for the layer is presented. Moreover, it is possible to select another date, to compare the values of the same cell in time, and to check the cell location in Google Maps. When a forecast layer is queried, in addition to the possibility of consulting the value of the cell for the selected month, the user can open a new popup in which all the forecasted month values are reported, eventually selecting the timescale (if the layer is timescale dependent), as in Figure 5.

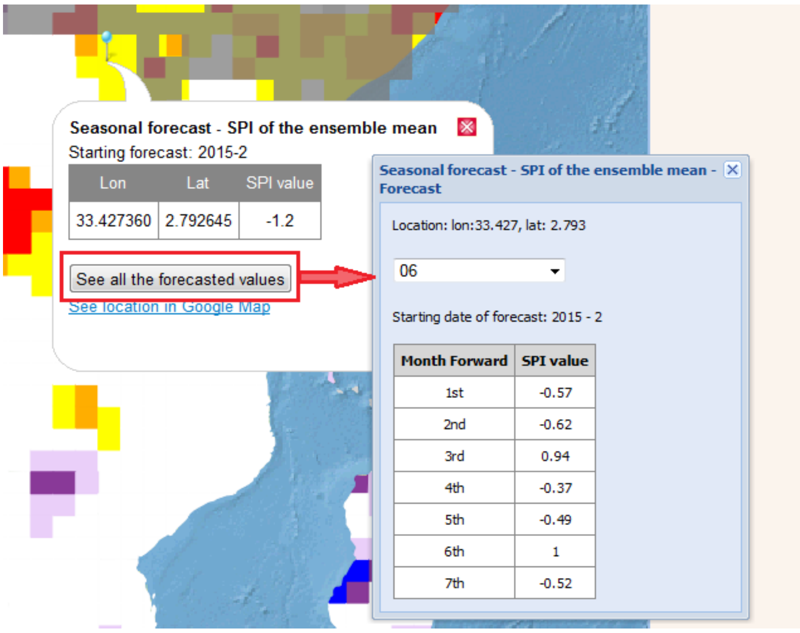


Figure 5: Example of "Identify" on a forecast layer. In this case it is possible to consult for the same cell all the forecasted values for the 7 months available.

Another feature of the toolbar is the possibility to consult on the map news related to droughts in Africa. The European Media Monitor (EMM[14]) reads and analyses around 40.000 new news items per day from around 1000 sites worldwide and provides a GeoRSS feed that can be filtered in order to show on the map only news related to drought in English.
The result is a new layer on the map that indicates with an icon the presence of news. When clicking on it, a popup with a short description and the external link is visualized, as in Figure 6.

14 http://emm.newsbrief.eu/NewsBrief/

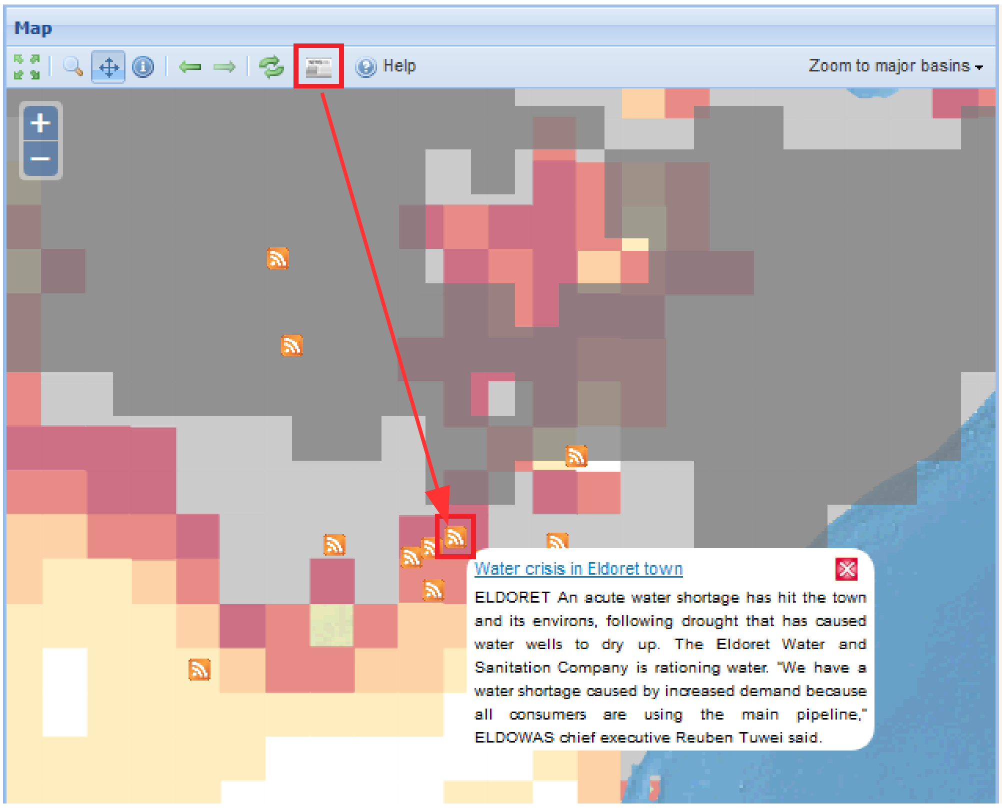


Figure 6: Example of drought news in Eldoret town.

2.2.2 Context Menu

Besides the toolbar, other functions are available at layer level. When turning on a layer in the layer tree, the legend is displayed below the layer name and the last date is presented (if the layer is time dependent). Right clicking on the layer name a context menu is shown with different functions according to the layer type. The menu commands are:

- layer opacity – to change layer transparency;
- select time (if time dependent) – to change the time extent;
- layer info – to retrieve a layer description;
- layer legend – to display the legend;
- zoom to layer extent – to zoom to the layer bounding box (not for all layers).

In the 'Select time' pop-up the user can act on two slide bars to choose the month and year and eventually select the ten-day period or timescale to visualize (see Figure 7). Every selection corresponds to a WMS request to the server with new parameters and the result is displayed in the map panel. In this way the time evolution of the indicator is clearly visible for the end user.

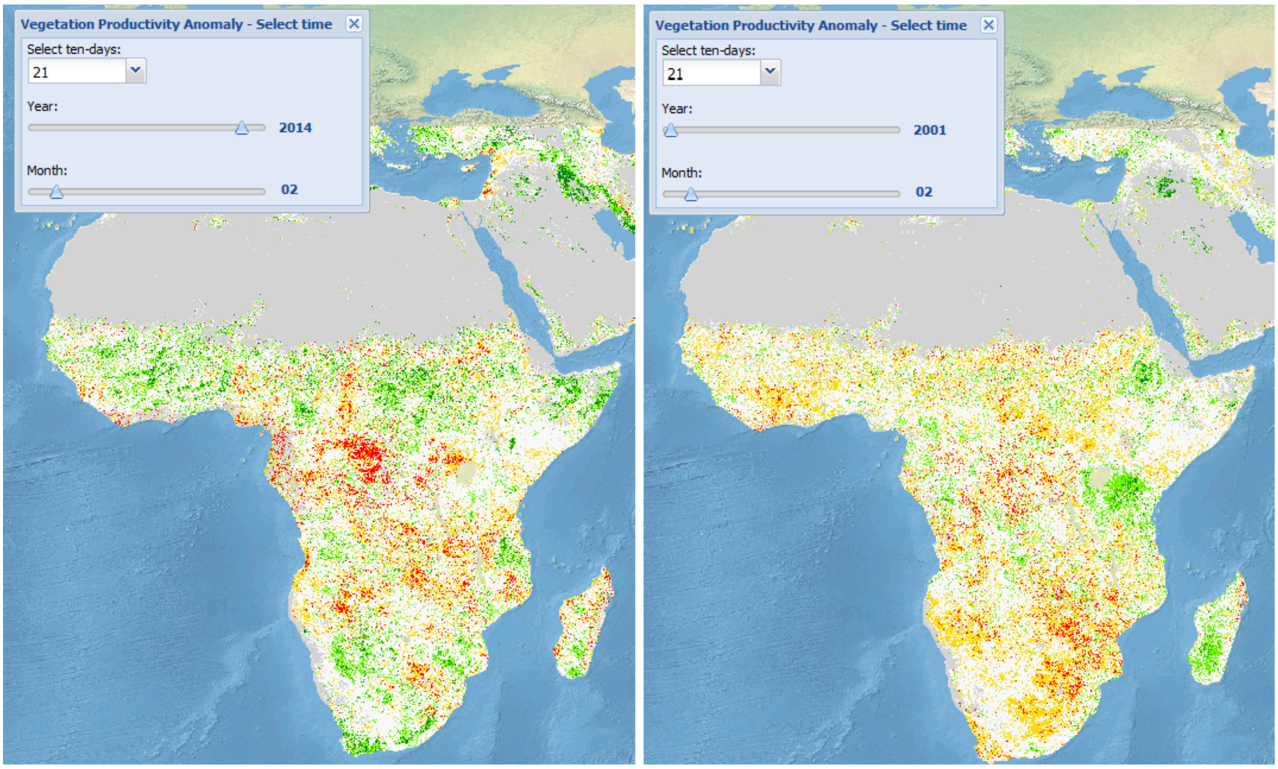


Figure 7: Example of time selection for the fAPAR anomaly layer. Data of the third ten-day period of February 2014 and 2001 are shown.

In the forecast layer tree, the time selection is differently managed. When the layer is turned on, a window is presented reporting the starting date of forecast and a slide bar to select the following seven months forecasted. Each cursor movement corresponds to a WMS request even in this case. In this way the evolution in time of the indicator prediction is clearly laid out.

2.2.3 Timeseries

The last tab in the left side of the panel, called 'Timeseries' is more data oriented. Once a location is selected, by typing the geographical coordinates or clicking on the map, it is possible to pick up one of the indicators stored in the database and the time extent that the user wants to analyze. For instance, all the values of Top Soil Moisture for all the months of January between 1990 and 2015, as in Figure 8:

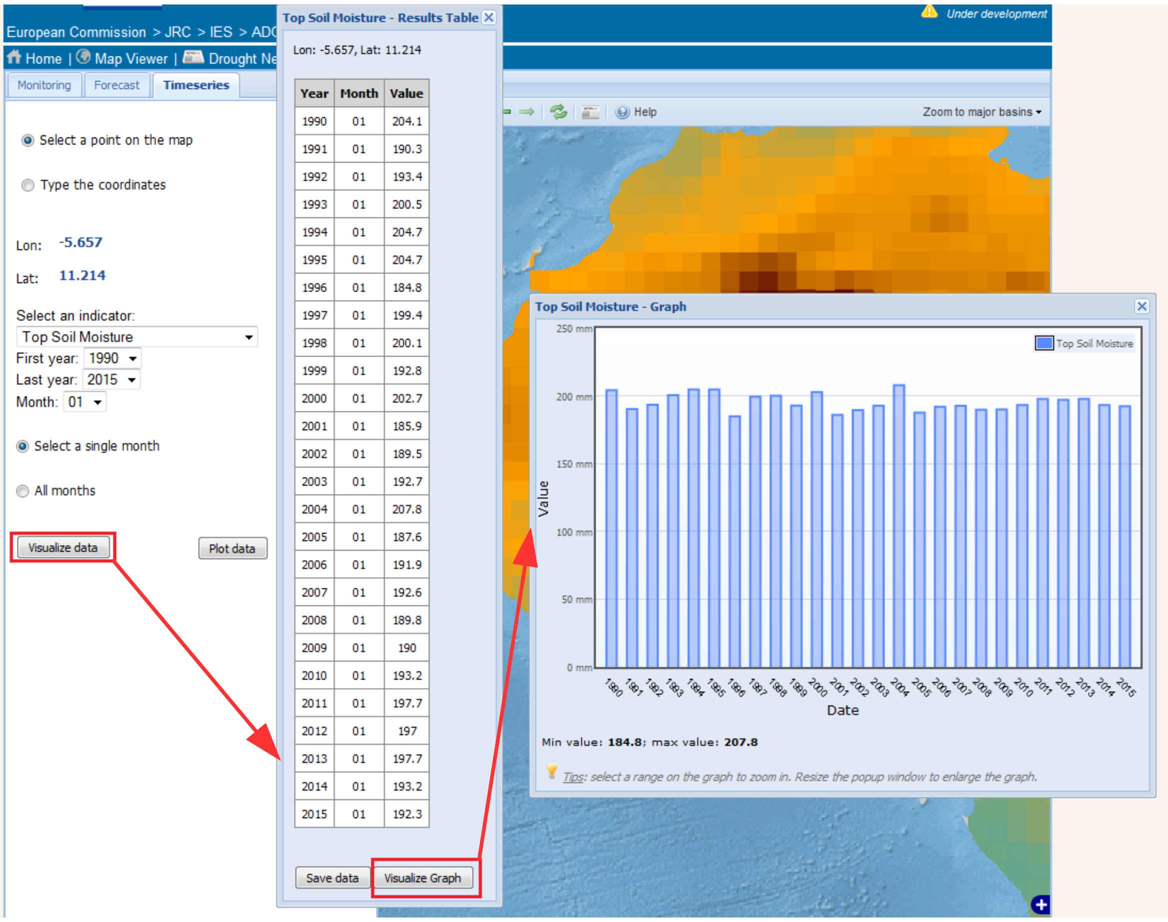


Figure 8: Data visualization for the Top Soil Moisture indicator for month 01 from 1990 to 2015.

Another possibility is to plot all the monthly values for the chosen years, as in Figure 9:

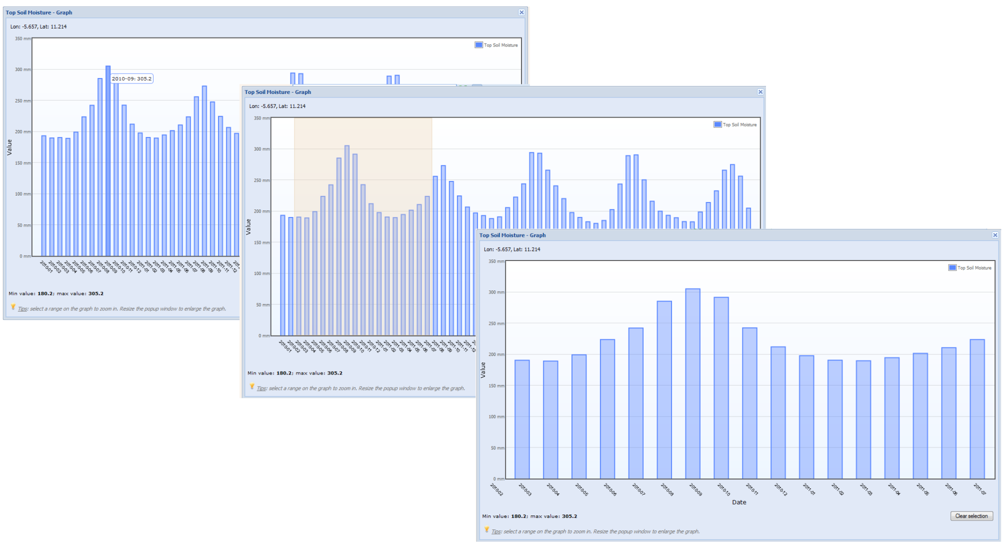


Figure 9: Chart example for the years 2010 - 2014. A zoom around the maximum peak is performed.

In the chart it is possible to highlight a selection of the graph and zoom on it to better consult the indicator values.

2.3 Drought News web page

An enhanced version of the drought news RSS feed visualization is implemented at http://edo.jrc.ec.europa.eu/ado/news.html.
Here news are not geolocated but listed according to the publication date. The news title, source, date and eventually location are listed, in order to have an overlook of the news content. A flag of the corresponding country allows the reader to quickly understand the location.
An enhancement with respect to the map layer is the possibility to change the language among English (default), French, German, Spanish and Portuguese.

3 Conclusions

The ADO Map Viewer presented integrates drought monitoring and seasonal forecasting related information in an innovative approach for Africa, based on the experience of Europe through the European Drought Observatory. Such a monitoring and forecasting system provides multiple drought monitoring and forecasting products as well as vulnerability and risk maps derived from socio-economic indicators. The system can be considered as a pre-operational prototype, with a variety of tools and drought related products that are accessible to end users and stakeholders interested in drought information in Africa. The use of free and open source software makes it transferable and extendible with possible African partners and collaborators.

References

- Dutra, E., Di Giuseppe, F., Wetterhall, F., & Pappenberger, F. (2013). Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index. *Hydrology and Earth System Sciences*, *17*(6), 2359-2373.
- Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J. & Pappenberger, F. (2014). Global meteorological drought–Part 2: Seasonal forecasts. *Hydrology and Earth System Sciences*, *18*(7), 2669-2678.
- Funk, C. (2009). New satellite observations and rainfall forecasts help provide earlier warning of African drought. *The Earth Observer*, *21*(1), 23-27.
- Gobron, N., Pinty, B., Mélin, F., Taberner, M., Verstraete, M. M., Robustelli, M., & Widlowski, J. L. (2007). Evaluation of the MERIS/ENVISAT FAPAR product. *Advances in Space Research*, *39*(1), 105-115.
- Heim Jr, R. R. (2002). A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, *83*(8), 1149-1165.
- Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, *391*(1), 202-216.
- Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T. & Vitart, F. (2011). *The new ECMWF seasonal forecast system (System 4)* (p. 49). European Centre for Medium-Range Weather Forecasts.
- Naumann, G., Dutra, E., Barbosa, P., Pappenberger, F., Wetterhall, F., & Vogt, J. (2013). Establishing the dominant source of uncertainty in drought indicators. *Hydrology and Earth System Sciences Discussions*, *10*, 13407-13440.
- Sheffield, J., Wood, E. F., Lettenmaier, D. P., & Lipponen, A. (2008). Experimental drought monitoring for Africa. *GEWEX News*, *18*(3), 4-6.
- Spinoni, J., Vogt, J., Naumann, G., Carrao, H., & Barbosa, P. (2014). Towards identifying areas at climatological risk of desertification using the Köppen–Geiger classification and FAO aridity index. *International Journal of Climatology*.
- Svoboda, M., LeComte, D., Hayes, M., Heim, R., Gleason, K., Angel, J., Rippey, B., Tinker, R., Palecki, M., Stooksbury, D., Miskus, D. & Stephens, S. (2002). The drought monitor. *Bulletin of the American Meteorological Society*, *83*(8), 1181-1190.
- Wilhite, D. A., Sivakumar, M. V. K., & Wood, D. A. (2000, September). Early warning systems for drought preparedness and drought management. In *Proceedings of an Expert Group Meeting held in Lisbon, Portugal* (pp. 5-7).
- Wilhite, D. A. (2000). Drought as a natural hazard: Concepts and definitions. In: *Drought: A Global Assessment* (pp. 3-18). Routledge Publishers, London, England.

Cloud computing application for water resources based on open source software and open standards – a prototype

Blagoj Delipetrev[1], Dimitri Solomatine[2]

[1] University Goce Delcev, Krste Misirkov bb, 2000 Shtip, Republic of Macedonia.

[2] Water Resources Section, Delft University of Technology, Delft, the Netherlands.

Abstract

This short paper presents research and development of a cloud computing application for water resources based on open source software and open standards using hybrid deployment model of public – private cloud, running on two separate virtual machines (VMs). The first one (VM_1) is running on Amazon web services (AWS) and the second one (VM_2) is running on a Xen cloud platform. The cloud application has three web services for 1) data infrastructure (DI), 2) support for water resources modelling (WRM), 3) user management. The presented cloud application is: available all the time, accessible from everywhere, it is scalable, works in a distributed computer environment, it creates a real-time multiuser collaboration platform, the programing languages code and components are interoperable, and it is flexible in including additional components. This research demonstrate the capability to scale and distribute the cloud application between several VMs. The cloud application was tested on the Zletovica river basin case study with concurrent multiple users. The tests confirmed that the presented cloud application works and can be used as a foundation for development of specialized geospatial application in the water domain.

Keywords

Cloud computing, water resources, open source, application development.

1 Introduction

Presently, most of the existing software for water resources is desktop-based, designed to work on a single computer. This represents a major limitation in many ways, starting from limited computer processing, storage power, accessibility, availability, etc. The only feasible solution lies in the web and cloud. There are various examples of web applications (Choi, Engel, & Farnsworth, 2005; Delipetrev, Jonoski, & Solomatine, 2014; Delipetrev, Mihajlov, Delipetrev, & Delipetrov, 2010; Horak, Orlik, & Stromsky, 2008; Reed, Bills, Anderson, Ketchum, & Piasecki, 2010), cloud web services (Bürger, Kollet, Schumacher, & Bösel, 2012; Quiroga, Popescu, Solomatine, & Bociort, 2013) and mobile applications (Jonoski et al., 2012) in the water domain.
The cloud application for water resources is continuation of previous research (Delipetrev et al., 2014) that presents the development of a web application for water resources based on open source software. The cloud application enhancement are the following:

1. The web application is distributed / deployed on two VMs. The VM_1 is running as a micro instance of Amazon web services (AWS) public cloud, and the VM_2 is running on a Xen cloud platform at the University Goce Delcev in the Republic of Macedonia.
2. The web service for support of WRM that runs on VM_1, and the DI web service that runs on VM_2, are communicating with WFS-T (Web Feature Service - Transactional Protocol) XML messages over the internet, demonstrating distributed computer environment. .
3. Hybrid cloud design is presented, where VM_1 is part of the AWS public cloud, and VM_2 is running in the private cloud. The advantage of this distributed computer environment is that the data security and protection can reside in the private cloud (VM_2), while the web services can be in the public cloud (VM_1).

The cloud computing application was tested using data from the Zletovica river basin located in the northeastern part of the Republic of Macedonia. The application url www.delipetrov.com/his/ provides video presentation and explanation of the system components, guides how to use the services etc.

2 Design and implementation

The main enhancement compared to the web application is the scalability. The design and implementation demonstrate that the cloud application can support seamlessly unlimited number of users. The web services that are shown in Figure 1:

1. DI.
2. Support of WRM.
3. User management.

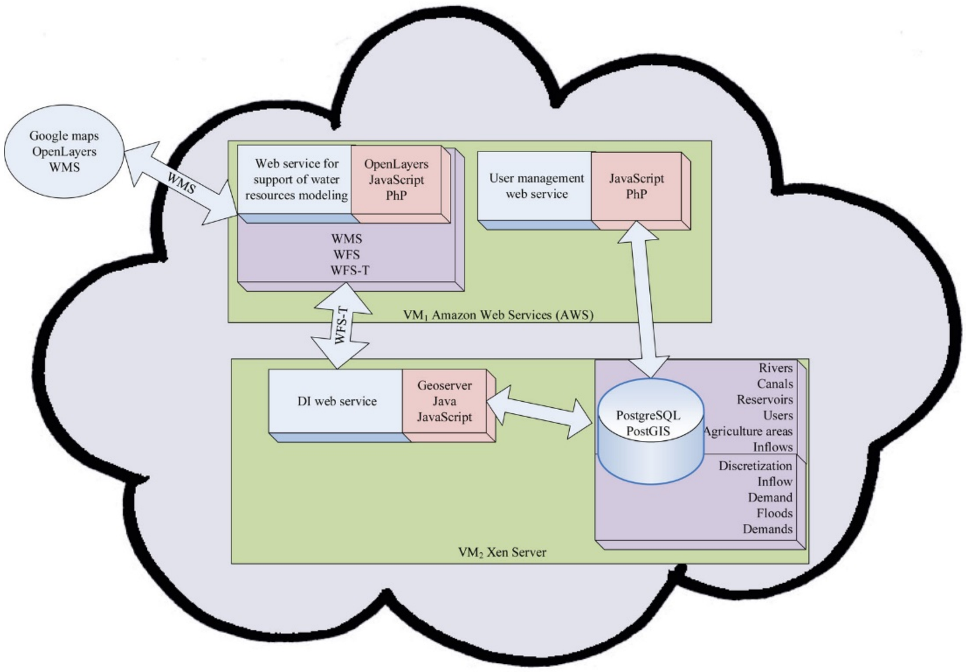


Figure 1: Design of the cloud computing application for water resources.

The DI web service is composed of two components 1) HMak database created in PostgreSQL and PostGIS and 2) GeoServer application. The HMak stores geospatial data, including topographic, hydro-geological, rivers, roads, municipal, etc. The HMak stores six geospatial vector layers: rivers, canals, reservoirs, users, inflows, and agriculture land, and their attribute tables, which are used by the web service for support of WRM.

The web service for support of WRM is intended to provide a web interface for creating and editing geospatial water resources elements just like in classical desktop applications. The main difference between the web service for support of WRM and desktop interfaces is that the web service for WRM is accessible by multiple users simultaneously over the internet using a web browser.

The web service for support of WRM is developed using PHP, Ajax, JavaScript and OpenLayer library that supports OGC standards. The OpenLayer library creates WFS-T communication between the web service for support of WRM user interface running on VM_1, and the geospatial data stored in HMak where the GeoServer acts like a middle tier running on VM_2. The WFS-T communication provides a framework to create, update, and delete geospatial data over the internet.

The web service for users' management is simple with a main purpose to control the cloud computing application access. Each user receives its own login and password to access the cloud application and monitoring the time spent on the system.

Important milestone is the deployment of the cloud application between the two VMs running on separate physical servers. The first VM_1 is a micro instance on the AWS, and the second VM_2 is running on the Xen cloud platform. The VM_1 has 8 GB HDD, 1 GB RAM and Ubuntu 13 as an operating system. The VM_2 has 30 GB HDD, 1 GB RAM and Fedora 16 as an operating system. The VM_2 is running on a physical server IBM x3400 M3 with four-core Intel Xeon E5620 2.40 GHz with 12 MB of cache per processor.

3 Preliminary results and tests

The cloud computing application was tested on river basin Zletovica that is located in the north-east of the Republic of Macedonia. The river basin main challenges are flood and drought protection. Figure 2 shows the hydro system Zletovica model that is created by the web service for WRM by multiple collaborating users. The hydro system contains the reservoir Knezevo, river network, canal network, towns as users and agricultural areas. The towns and reservoir titles are added additionally and are not part from the web service.

Figure 2: Water resource model of the Zletovica river basin.

4 Discussion

Concerning service models, the presented cloud application belong to software as a service (SaaS). Users with a web browser access the cloud application and do not care about underlying cloud infrastructure. The current deployment model is hybrid of public - private cloud because the VM_1 is running in public cloud AWS, and VM_2 is on private Xen cloud platform. The cloud application is scalable, interoperable, and accessible from everywhere and available all the time.

The most valuable cloud application feature is its real time collaboration platform capabilities. Multiple users using only a web browser can jointly work with the web services and collaborate in the same working environment. When a user saves the current work, all other distributed users with just a web browser refresh can see the changes (new/modified rivers, users etc.). All of the data and models are stored in HMak and users do not have to be concerned about hardware and software support infrastructure.

Another important concern in implementing cloud solutions is the data protection and safety. This prototype cloud application makes an elegant solution, residing the web services in the public cloud, while the data is stored in the private cloud. If for instance an attack happens to protect the data, the private cloud VM_2 can be disconnected from the public cloud AWS VM_1. Another key point is that data resides inside the institution, and only the web services are “outsourced.” This concept can be applied in many organizations where the data needs to be stored internally.

5 Conclusions

The presented cloud application is developed using open source software and standards and prototype code. The application is deployed on two VM creating hybrid cloud solution. The cloud application is a state-of-the-art web geospatial collaboration platform for water resources modeling. The presented solution is a prototype and can be used as a foundation for developing specialized web geospatial applications.

References

- ✔ Bürger, C. M., Kollet, S., Schumacher, J., & Bösel, D. (2012). Introduction of a web service for cloud computing with the integrated hydrologic simulation platform ParFlow. *Computers and Geosciences, 48*, 334-336.
- ✔ Choi, J., Engel, B., & Farnsworth, R. (2005). Web-based GIS and spatial decision support system for watershed management. *Journal of Hydroinformatics, 7*, 165-174.
- ✔ Delipetrev, B., Jonoski, A., & Solomatine, D. P. (2014). Development of a web application for water resources based on open source software. *Computers & Geosciences, 62*, 35-42.
- ✔ Delipetrev, B., Mihajlov, D., Delipetrev, M., & Delipetrov, T. (2010). Model of the Hydro-Information System of the Republic of Macedonia. *Journal of Computing and Information Technology, 18*(2), 201-204.
- ✔ Horak, J., Orlik, A., & Stromsky, J. (2008). Web services for distributed and interoperable hydro-information systems. *Hydrology and Earth System Sciences, 12*(2), 635-644.
- ✔ Jonoski, A., Alfonso, L., Almoradie, A., Popescu, I., van Andel, S. J., & Vojinovic, Z. (2012). Mobile phone applications in the water domain. *Environmental Engineering & Management Journal (EEMJ), 11*(5).
- ✔ Quiroga, V. M., Popescu, I., Solomatine, D., & Bociort, L. (2013). Cloud and cluster computing in uncertainty analysis of integrated flood models. *Journal of Hydroinformatics, 15*(1), 55-70.
- ✔ Reed, P. M., Bills, B., Anderson, M., Ketchum, B., & Piasecki, M. (2010). The Mid-Atlantic Watershed Atlas (MAWA): Open access data search & watershed-based community building. *Environmental Modelling & Software, 25*(6), 808-812.

Definition of hydrogeological tectonic blocks into Guarani/Serra Geral Integrated Aquifer System using QGIS

Leônidas Luiz Volcato Descovi Filho[1], Arthur Schmidt Nanni[1], Luiz Fernando Scheibe[1,2]

[1] Universidade Federal de Santa Catarina, UFSC, Brazil

[2] Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brazil

Abstract

The application refers to the Guarani/Serra Geral Integrated Aquifer System (SAIG/SG) in Santa Catarina State, Brazil. Based on Digital Elevation Model SRTM and derivatives, large features reflecting hydrogeological tectonic blocks[1] were defined. The interpretation of diagnostic profiles different scales result hydrogeological tectonic blocks. With hydrogeological database, calculate the hydrogeological tectonic blocks mean surface. Using QGIS Free Geographic Information System in this study was highly feasible.

Keywords

Geomorphostructures, QGIS, Terrain Profile.

1 Introduction

Serra Geral, Botucatu and Pirambóia formations constitute a major aquifer system in Southern Brazil's Paraná Sedimentary Geological Basin, the "Guarani/Serra Geral Integrated Aquifer System" (SAIG/SG). The Serra Geral Formation is a fractured volcanic reservoir that override the Guarani Aquifer System, a porous fine to middle grain sands aquifer type. Cutting both geological units, there are a lot of fractures and normal faults that connect these aquifers, allowing descendant or ascendant flow, with the exchange of different water types (Scheibe and Hirata, 2008).

The major objective of this research is the definition of the hydrogeological tectonic blocks that make up the spatial arrangement of these two aquifers, that should be exploited as a single integrated system, in southern Brazil.

The determination of these blocks is being held through QGIS and its native tools and plugins. These tools combined with a new hydrogeological database, provide support for the determination of the elevation of each tectonic block that define hydrogeological tectonic blocks, allowing to define vertical displacements and enabling to determine edge disconnections between the geologic strata layers.

To improve the technical knowledge of SAIG/SG in Santa Catarina state, this study sought support in geotechnology tools, specially the QGIS, 2.6.1 Brigthon

[1]delimited blocks from erosional features and structures, indicated by simultaneous viewing, using transparency of layers slope, hypsometry, shaded relief and drainage network, with tectonic and hydrogeological significance.

version QGIS (2015), for the definition of the hydrogeological tectonic blocks in the study area.

2 Methodology

All raster bases used in this study are derived from the digital elevation model obtained by radar interferometry by the project SRTM / NASA 2000. The spatial resolution of the images are 90 x 90 meters, originally presenting the geographic coordinate system World Geodetic System 1984 / WGS84, that was converted to plane coordinates system Mercator Universal Transverse, WGS84, zone 22 South of Equator. The following maps were drawn:

- Slope;
- Geomorphostructures at 1:1.000.000, 1:500.000 e 1:250.000 scales;
- Depth of Serra Geral and Guarani Aquifer Systems geological contact, according to well's informations, gathered in internal reports and field work.

This approach aims to define the relative movements and elevation of the hydrogeological tectonic blocks, and was complemented by the use of the plugin Terrain profile.

3 Preliminary Results

The hydrogeological tectonic blocks are interpreted on the slope map at 1:1.000.000 scale (Figure 1). The variation in slope to different areas, is a differential erosive product and suggest a relationship with hanging wall and footwall of hydrogeological tectonic blocks.

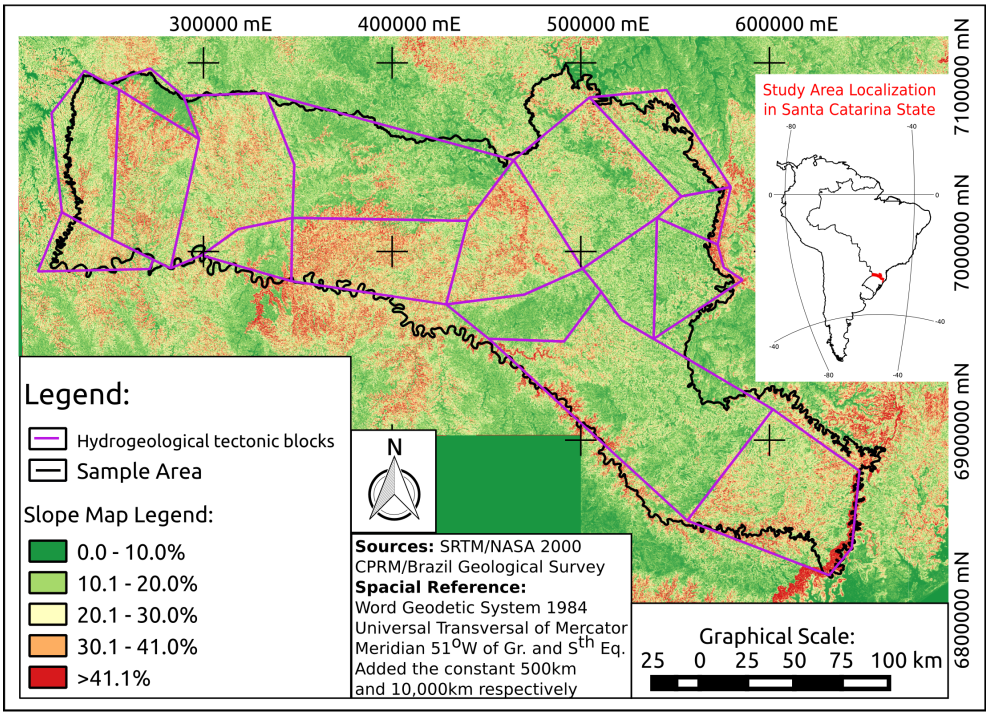


Figure 1: Hydrogeological tectonic blocks interpreted after the slope base map.

The interpretation of the hydrogeological tectonic blocks was carried out using:

- DEM: hypsometry, slope, expeditious profiles for confirmation of the blocks

(using elevation and terrain profile plugins) and;

- conference with the geological points of contact between the Botucatu and Serra Geral formations, coming from a new hydrogeological/stratigraphic database.

The analysis of geomorphological boundaries were proceeded through profiles (Figure 2) and the drainage network contributed to better definition of these limits of the blocks, through detected relief and drainage asymmetries, and indicating the likely direction movement of the blocks (hanging wall/footwall).

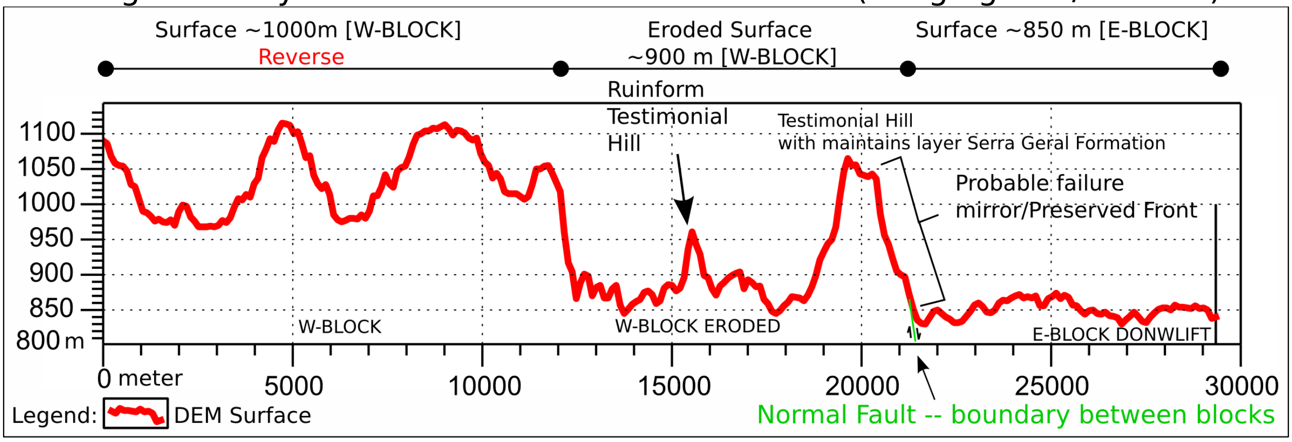


Figure 2: Topographic profile presenting geomorphological and geological details that enable the verification of the likely limits of hydrogeological tectonic blocks.

Figure 3 is the preliminary result of interpretation of the hydrogeological tectonic blocks analyzed in the 1:1,000,000, 1:500,000 and 1:250,000 scales. The numbers in the blocks represent the average altitude (related to sea level), of the contact between the base of the Serra Geral Formation and the top of Botucatu Formation (Guarani Aquifer), obtained from wells drilling data. These stratigraphic data were interpolated using the QGIS raster Interpolation tool, by the Triangular Irregular Network / TIN interpolation method. This step is important to set the hanging wall/footwall relationship of the hydrogeological tectonic blocks.

4 Preliminary Conclusions

Preliminary results corroborate and reinforce the ideas of Scheibe and Hirata (2008) and Nanni et al., (2009), which stressed the existence of discontinuities in the Guarani/Serra Geral Integrated Aquifer (SAIG/SG), eventually responsible for low productivity wells and fluoride excess in their water, and, in some cases, the formation of cells and blocks with distinct water residence time and geochemical characteristics.

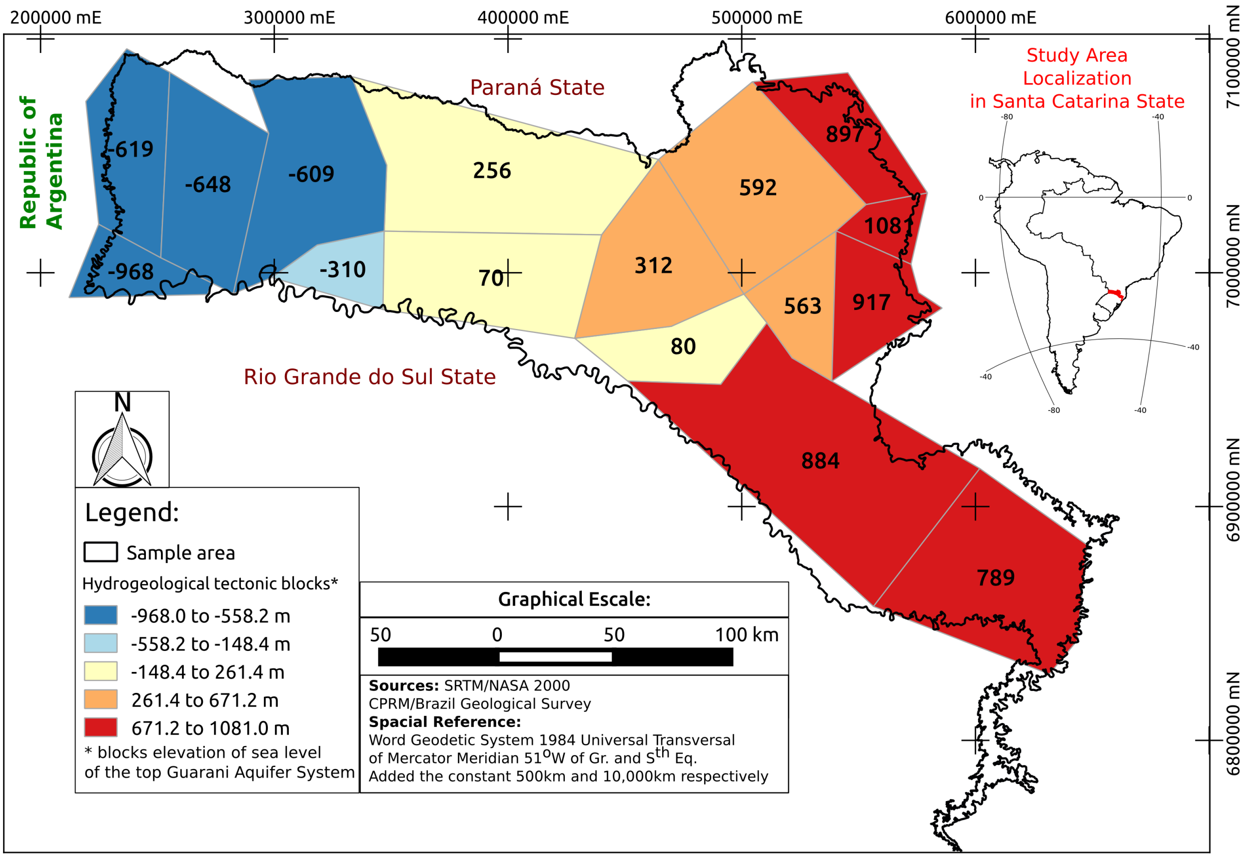


Figure 3: Hydrogeological tectonic blocks with the top altitude of the Guarani Aquifer System.

In this work, we endorse the conclusions of these authors, since significant vertical discontinuities between neighboring blocks are detected and analyzed. These suggest large lateral discontinuities in this aquifer system, which ultimately constitute an integrated system of the two major aquifers in the southeast of South America.

QGIS has been an important tool to develop this research, enabling to elaborate propositions of directions to groundwater management of Guarani/Serra Geral Integrated Aquifer System.

References

- ✓ Nanni A., Roisenberg A., Marimon M. P. and Viero A. P. (2009)“The hydrochemical facies and anomalous fluoride content in the Serra Geral aquifer system, southern Brazil: a GIS approach with tectonic and principal component analysis,” Environmental Geology, 58(6), pp. 1247–1255. DOI 10.1007/s00254-008-1624-2.
- ✓ QGIS Development Team, 2015. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org
- ✓ Scheibe, L. F. & Hirata, R. C. A. (2008) Contexto tectônico dos Sistemas Aquíferos Guarani e Serra Geral em Santa Catarina: uma revisão. In: XV Congresso Brasileiro de Águas Subterrâneas, São Paulo: ABAS, pp. 01-14.

SITGAP2, a decision support system for floods management based on open source and open standards

Massimiliano Cannata, Milan Antonovic

Institute of Earth Sciences, SUPSI, Campus Trevano, 6952 Canobbio, Switzerland

Abstract

The Locarno area (Switzerland, Canton Ticino) is exposed to lake floods with a return period of about 7-8 years. SITGAP2 is the latest implementation of this decision support system; implemented in collaboration with the Civil protection of Locarno and Vallemaggia, SITGAP2 combines a number of open source software, open standards and different data sources. As a result of this orchestration of services and data, SITGAP 2.0 offers features that allows the risk management.

Keywords

Flooding, Decision Support System, Open Standards, istSOS, Lake Maggiore.

1 Introduction to the Lake Maggiore flood risk

The Lake Maggiore basin is characterized by and area of 6958 Km^2, 3229 of which are in Italy and 3369 in Switzerland. The surface of the Lake is about 210 Km^2 and drains four important tributaries (Ticino, Toce, Maggia and Tresa) whose basins cover the 77% of the total lake basin extension. If compared to the watershed size, the limited lake outflow, which is about 2,000 m^3/s, causes recurrent floods in the coastal areas with a return period of about 7 years. Latest important events were registered in 1978, 1993, 2000, 2002 and 2014. The risk is of particular concern because the area is located in a floodplain that registered in the last decades a great increase in settlement and values of the real estates. While the lake floods are generally slow events which do not cause fatalities, this type of floods may result in high socio-economical impacts.

2 The decision support system

2.1 Background and evolution of the system

After yet another relevant event in the 2000, a consortium of stakeholders composed by the Canton Ticino, Swiss Meteorological Office, Civil Protection of Locarno and Vallemaggia (PCi-LV) and the Institute of Earth Science (IST) at SUPSI were set-up to address the issue and develop mitigation measures. As a result, a decision support system named SITGAP and based on geospatial technologies had been developed to enable early warnings and management capabilities. In the following years a series of activities and projects where conducted to enable the data exchange between Swiss and Italian bodies, to bring meteo-hydrological monitoring network data at near-real time availability, to improve hydrological model usability, to orchestrate data gathering from different authoritative sources. In early 2012, considering the ten year old technology which SITGAP was based, the experience gained in more then 10 year of system management and the opportunity to take advantage of all aforementioned efforts, the Pci-LV and the IST decided to renew the decision support system changing paradigm: the information update shall be possible using a Web interfaces by non IT specialized users.

Thanks to impressive technological advances the visionary concept of the Digital Earth is being realizing: geospatial coverages and monitoring systems data are increasingly available on the Web, and more importantly, in a standard format. As a result, today is possible to develop innovative decision support systems which mesh-up several information sources and offers special features for risk scenarios evaluation. In agreement with the exposed view, the authors decided to design a new system, named SITGAP2, fully based on: Service Oriented Architecture pattern; Free and Open Source Software (e.g.: Geoserver, PostGIS, OpenLayers); and Open Standards (e.g.: SOS, WMS, WFS).

2.2 Workflow and technologies

Figure 1 illustrates the system work-flow and the technologies that have been adopted. The hydro-meteorological monitoring network, meteorological forecasts and contextual geospatial data are the system inputs. Sensor data storage, validation and distribution is an essential part of the system that enables the validation and sharing of key measurements used by the hydrological model to provides lake level forecasts. These levels alone cannot be sufficient to take wise decision, in fact only overlaying this information with geospatially referenced data it is possible to assess expected impacts and exposed elements. For this reason a GIS that mesh-up and integrates data coming from several authoritative sources has been set-up. Information is collected from a number of data sources such as the federal register of buildings and dwellings, the cantonal register of residents, the cantonal cadastre, the cantonal hydro-meteorological monitoring observations, the Meteoswiss weather forecasts, and others.

From a technological point of view sensor data acquisition, management and dispatch is performed by means of the istSOS software. The acquisition, re-projection and dispatch of meteorological forecasts (COSMO-7 meteorological model rainfall outputs provided by meteoswiss) according to the hydrological model input format, is performed using PostGIS and pyWPS. Base geospatial layers are made available over the Web thanks to the Geoserver software and the WMS and WFS formats. SITGAP2 features are available trough the combined mesh-up of: RESTful API based on PostgreSQL/PostGIS and Python; OpenLayers web mapping framework; and HTML5 with CSS3 code based on the Sencha Ext JS development toool.

The hydrological model is a SCS-CN-based conceptual model that applies the Nash method. It is activated in case of meteorological alert only. Every model run, after a first phase of calibration based on measured observations collected from the sensors, provides 72 hours lake level forecasts by using meteoswiss rainfall forecasts.

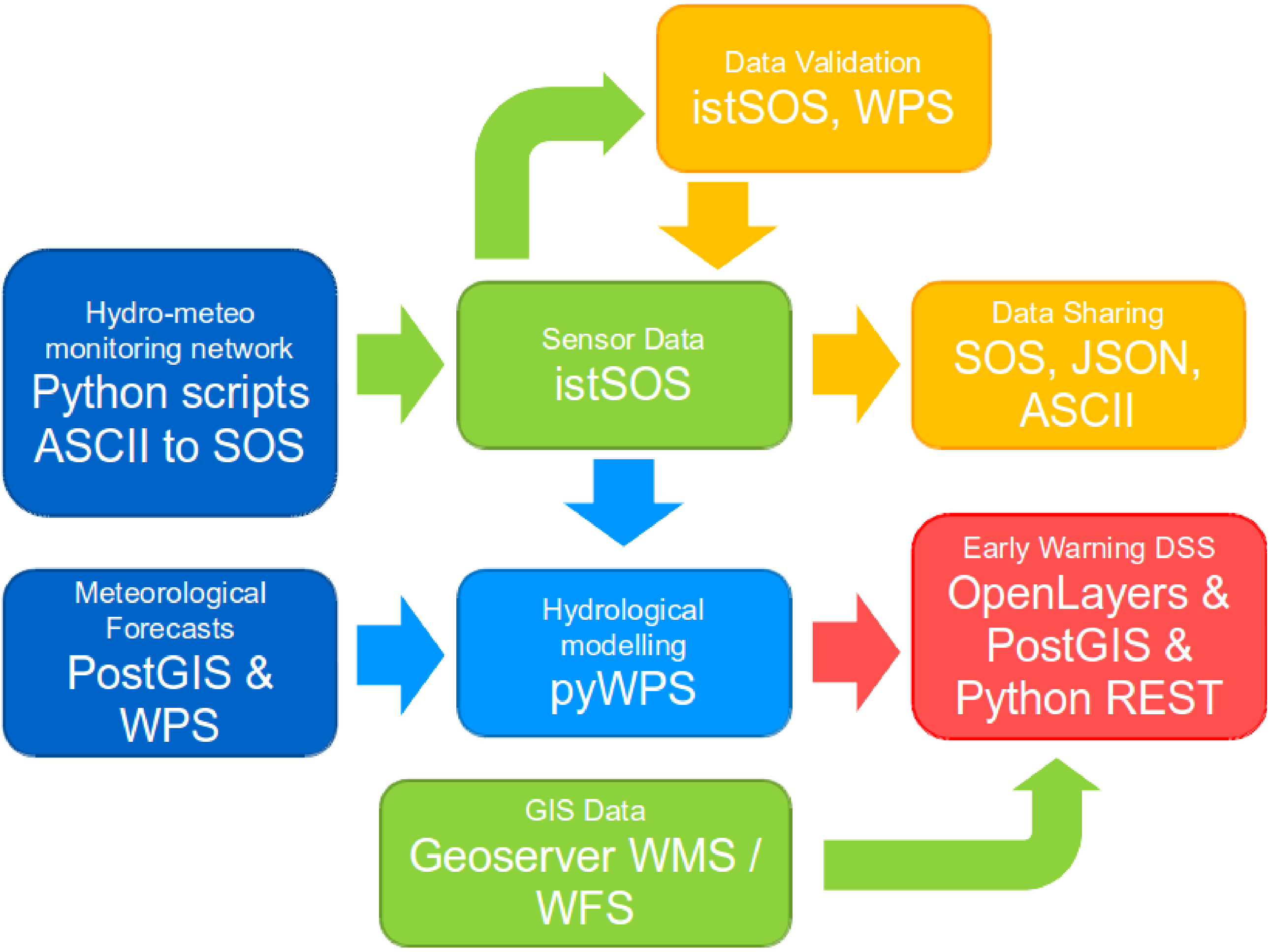


Figure 1: System work-flow and adopted technologies.

2.3 Graphical user Interface

The interface of SITGAP2, as illustrated in Figure 2, is available in Italian only because it is expected to be used by local stakeholders. It is composed by four different components or modes:

- "Regolare" (regular): it provides access to data and functions for resources management like search for parcel, address or person, query and eventually edit elements details.
- "Allarme" (alarm): it gives access to lake level forecasts and elements potentially at risk offering features to detect exposed elements at a given (forecasted if an alarm is active) lake level, visualization of time evolution of lake level forecasts and access to instructions to secure exposed elements.
- "Evacuazione" (evacuation): it provides access to population information used to manage evacuations and relative features: creation of a new evacuation, search and localization of persons, counting and listing of people living in an area, registration of evacuation status for single people.
- "Dighe" (dams): it offers access to dam-break hazard zones and wave arrival times.

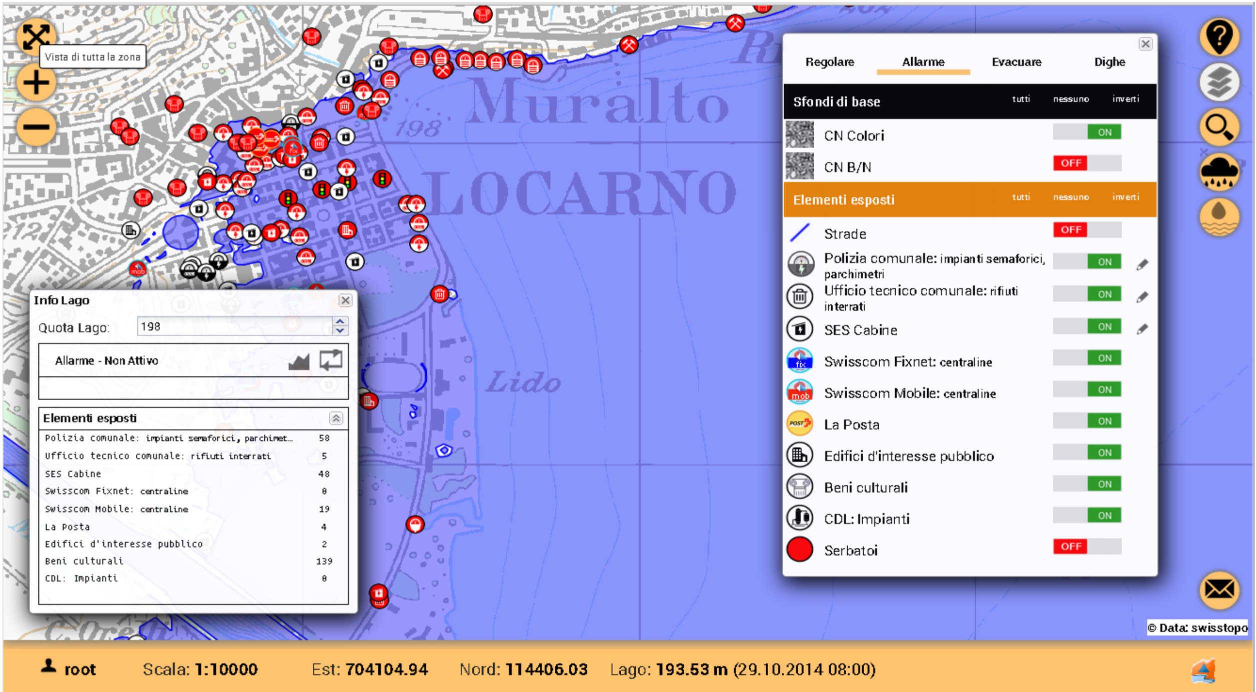


Figure 2: graphical user interface of SITGAP2 in "Allarme"mode (alarm): exposed elements for a given lake level are identified in map. Basemaps (c) swisstopo.

3 Conclusions based on the flooding of November 2014

As reported by the climate bulletin of November by meteoswiss, from the 2nd to the 17th of November 2014 the Canton Ticino registered precipitations characterized by unusual intensity. For instance, Lugano, which is the reference station to the south of the Alps, measured in that period 538 mm.

A higher amount of rain in the period of 16 days was rerecorded only once in the last century: 547 mm were recorded in the October-November of the 1928.

One of the consequences of this intense event was the conspicuous increase of the Verbano lake levels which reached 196.41 meters above sea level; the last registered maximum measure was of 197.57 m in October 2000. An alarm of level 5 was activated by the civil protection in response to the impending flooding hazard and the SITGAP2 platform was for the first time used to manage the emergency. The PCi-LV agents reported a great satisfaction stating that the platform permitted to save time and money in the emergency management, for example indicating when and where catwalks and barriers should be placed.

The system, even not yet in its definitive deployment, already proved its stability and reliability. In fact, during that period no system downtime was recorded and only the 0.004% of the total requests resulted in an error.

During that period, daily average statistics shown: 215'484 requests, 50 unique visitors and 3.4 GB data transmitted, out of which 2.57 GB was images generated by the WMS and 0.83 GB was data transferred by the implemented RESTful API. The requests per seconds ranged between 2.5 and 30.

Information richness and reach in hydrological web applications

Arjen Vrielink

Nelen&Schuurmans consultants, PO Box 1219, 3500 BE Utrecht, The Netherlands

Abstract

This paper demonstrates two applications in water management that are disrupting the traditional reach and richness trade off in the economics of information: 3Di, a real time, web based flood modeling tool and Lizard, a water management information system, which is able to integrate and aggregate the ever growing amounts of (water) data into information for water managers, operators and policy makers.

Keywords

flood modeling, water management, big data, cloud computing, open source

1 Introduction

On the Internet, time moves fast, ideas move fast, data moves fast. Todays hero's are tomorrows losers. Web 2.0 was happening 10 years ago, todays buzzwords are 'the cloud', 'big data', 'open data' and 'data mining'. Most of the buzz surrounding those themes is snake oil. What still holds is that the Internet fundamentally changed the economics of information; traditionally, information was embedded in physical modes of delivery. Be it linked to paper (the news), a person (consultant) or to a hydrological model which can only be run by experts on specialized hardware.

That meant that there was a trade off between richness and reach. You either had a wide reach and low richness of information or a narrow reach with much richness. Reach means the number of people exchanging information. Richness is defined by three aspects of the information itself:

- Bandwidth
- Degree of customization
- Interactivity

The Internet changed that trade off completely. You can now reach many people with highly customizable, interactive, high bandwidth web applications. Thereby deconstructing the traditional information value chain and highly reducing the ability to control the monopoly on information (Evans & Wurster, 1997).

2 3Di: real time flood modeling via the web

For example, in the old days, if you wanted to know the flood risk for a densely populated area in case of the breach of a levee, you had to hire a consultant who would model your area, run a model, write a report and return the results

to you in a few weeks or months. The whole information chain was monopolized by the consultant or his company. If you wanted to run a different scenario, or the same scenario with different parameters or improved input data, you were dependent on that consultant and his hydrological model.
Current advances in hydrological modeling make it possible to accurately calculate flood risks using high resolution input data, in real time. By disclosing such models through the Internet with an interactive web application, the model gets disconnected from its physical mode of delivery (the consultant or the report). 3Di is such an application: a web based versatile water management instrument that supports operational water management, calamity management and spatial planning design. Running a scenario takes about 10 seconds. The 3Di instrument is based on detailed hydraulic computations. The computations are extremely fast and therefore allow for interactive modeling on a touch-table and iPad. The model run can be completely controlled by the user from his browser. 3Di shows how flood model information moved from the consultant or model expert to the user. 3Di is a collaborative effort between Nelen&Schuurmans, Deltares, TU Delft EWI / Virtual Reality and Guus Stelling and completely built with open source tools.

3 Lizard: integrating and aggregating data for intuitive water management

Closely related to 3Di is Lizard. Lizard is an open source information platform for the water world. Lizard consistently combines measurements, (3Di) model results and expert knowledge in one system. By aggregating and integrating data from different sources, Lizard provides information useful for the water manager, operator or policy maker.
Lizard supports all major data types common for water management: time series (eg. sensor data), vector data (e.g. pump station, culvert and canal) and raster (eg. rain, elevation, land use). Lizard has a processing engine based on Apache Spark, HBASE and PostGIS which automatically aggregates and integrates data over time and space into hydrological meaningful information. Information is presented in real time to the user for any scale in space and time. That unique property gives water managers, operators and policy makers immediate insight in complex water systems. Custom, expert designed tools provide fast and intuitive interaction with vast amounts of data. Lizard is developed by Nelen&Schuurmans. The author is lead developer of Lizard.

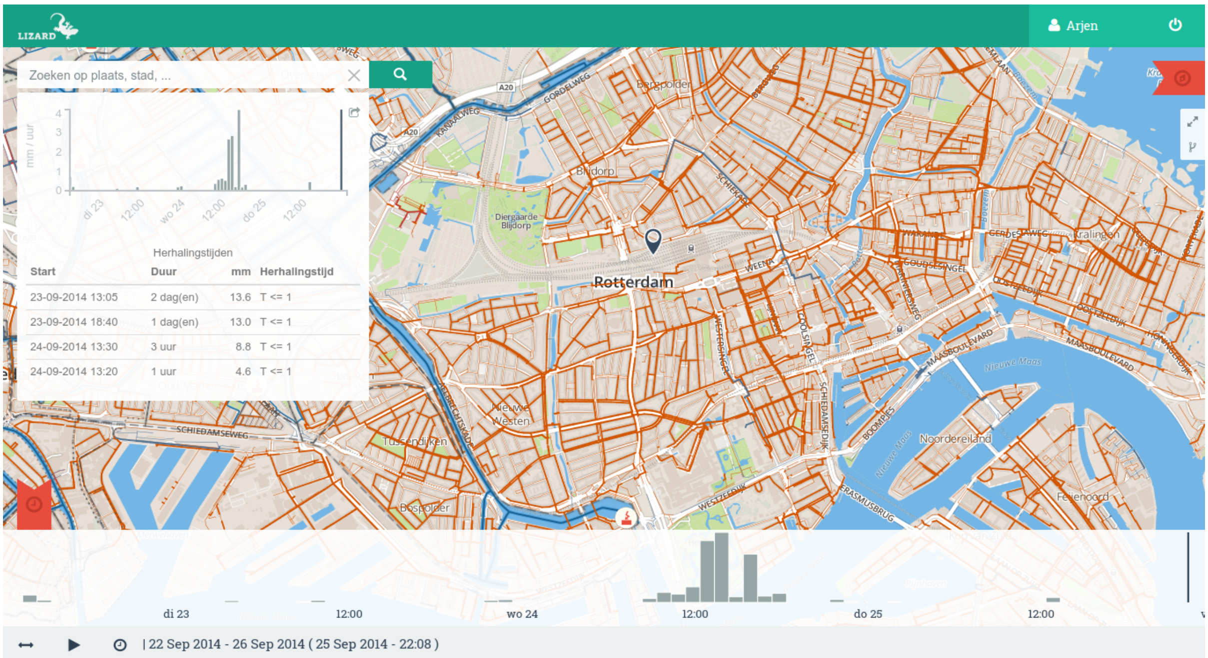


Figure 1: Lizard, sewerage of city of Rotterdam with a rain shower around 12:00. Data is both spatially and temporally aggregated and visually shown

With developing the Lizard platform and 3Di we tackled interesting challenges like:

- Authorization on entities within map layers; Current standard protocols like OGC's WMS don't support authentication and / or authorization of web map layers. Lizard implements an authorization model that enables fine grained control of data access.
- Handling large amounts of sensor data; distributed NoSQL storage (HBase, Hadoop, Spark) allows Lizard to handle all sensors in all the dikes of The Netherlands.
- Fast display and querying of very big (dynamic) rasters (billions of pixels) for seamless integration and aggregation.
- Build the time dimension into the core of the application; in addition to the three spatial dimensions, geo information is becoming more and more time aware. Lizard has built in support for handling time aware data. Lizard can animate time aware raster data (like rain, air quality) and vector data.
- Vector tiles and client side rendering of (time aware) geo data (AngularJS, D3js); to enable fast, intuitive interaction in the browser, you have to cannot rely on server side rendering of map layers anymore.
- Real time, interactive hydrological modeling with 3Di; state of the art hydrological models like 3Di are fast enough to calculate flooding models in real time. New standards like web sockets make sure you can use your browser as an interface to 3Di.

Conclusion

3Di and Lizard are examples of how the Internet helps democratize (Benkler, 2006) heretofore scattered and opaque (water) management information and previously specialist-only tools. Though promising and already successful, it is doubtful that they already show the full potential of networked, interactive, user directed, information systems in the water world. It is exciting to realize we are only at the start of this revolution. The future is only limited by our imagination.

References

- Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press.
- Evans, P.B., Wurster, T.S. (1997). Strategy and the new economics of information. *Harvard Business Review*, 75(5), 70-82.
- 3Di site: http://www.3di.nu/en/ (English), http://www.3di.nu/ (Dutch)
- Lizard: http://www.lizard.net/ (Dutch)

Disaster Risk Management: Urban Flooding and heatstress

Mark Verlaat[1], Leon van der Meulen[1], Govert Schoof[1], Floris Boogaard[2], Jeroen Kluck[3]

[1] Geodienst, Research and Innovation Support, University of Groningen

[2] Hanze University of Applied Sciences

[3] TAUW engineering

Abstract

Natural disasters are a growing concern around the globe. In the Netherlands, water has always played an important role as both friend and enemy. To quickly analyze and visualise possible disaster outcomes has been really difficult.
In collaboration with engineering company Tauw we improved this modelling with an interdisciplinary team of GIS experts, High performance computing and real time visualisation. In a pilot for the city center of Groningen we developed a 3D version of flooding landscape maps (RUG, 2014) after modelling extreme rainfall. With a flooding landscape map you can see at a glance where water is going and where problem areas arise in case of extreme rainfall. Any municipality or county can thus quickly determine which measures are to be taken to prevent for example disruption to traffic or flooding damage to buildings.

Keywords

Water, Flooding, Modelling, Disaster Risk Management, Resilience, 3D visualisation

1 Modelling for climate adaptation

For the pilot we used scenarios of the Royal Dutch Meteorological institute to set up situations which we could both model and visualize. We chose for extreme rainfall and heat stress. This pilot we did for the beautiful city of Groningen. The basis of the model was first developed by engineering company TAUW (Klok, 2012), but hadn't been used on this scale before. We enhanced the model, increasing performance and making it suitable for cluster computing. The results were visualized in 2d maps and 3d models to give detailed insight into the outcomes of the disaster scenario.

2 Floodrisk

For the urban flooding scenario we used several datasets. The most important one is the AHN2, the Dutch elevation model (which is open data). This raster dataset is extremely accurate, 9 points per square meter. Therefore we know the height for every 50 by 50cm. On top of that we used the TOP10, the open data topographic data for the roads and waterways. We also used the BAG, the open dataset with all buildings (2d) of the Netherland. Next Groningen was cut up into 25 pieces and 60mm of rain in one hour was simulated. The sewage system was taken into account by removing the first 20mm of water and the rest flowed through the city. The output of this computation is a map which shows the flow of the water and the water that remains at certain places.

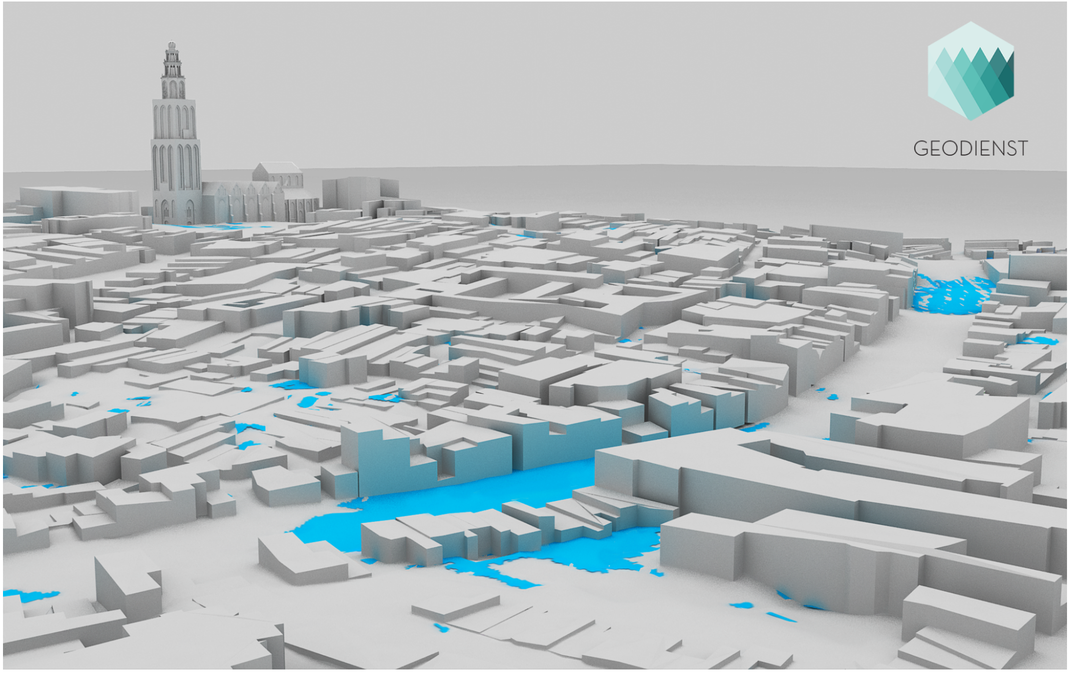


Figure 1: Results visualized in a 3d model.

Figure 2: Map showing water flow and accumulation.

Combining the elevation model, the dataset with buildings and aerial photographs we can make a 3d model of the city and get a better overview of the outcomes of the model.
The model was shown in the 3D virtual reality theater on a cylindrical screen using 6 HD projectors to project an image with a resolution of roughly 5000x1800. To display this model on such a big screen a special 3D viewer, based on the open source OpenSceneGraph 3D toolkit, was used. The software was running simultaneously on 7 PC's, one master PC for the control of the model and 6 slave PC's to drive the projectors.

3 Heatstress

The same model as the floodrisk model is applicable to heat stress and drought. In this model we also included an aerial photo with 10cm resolution. And then used a script to the work and make the sun shine above the city.

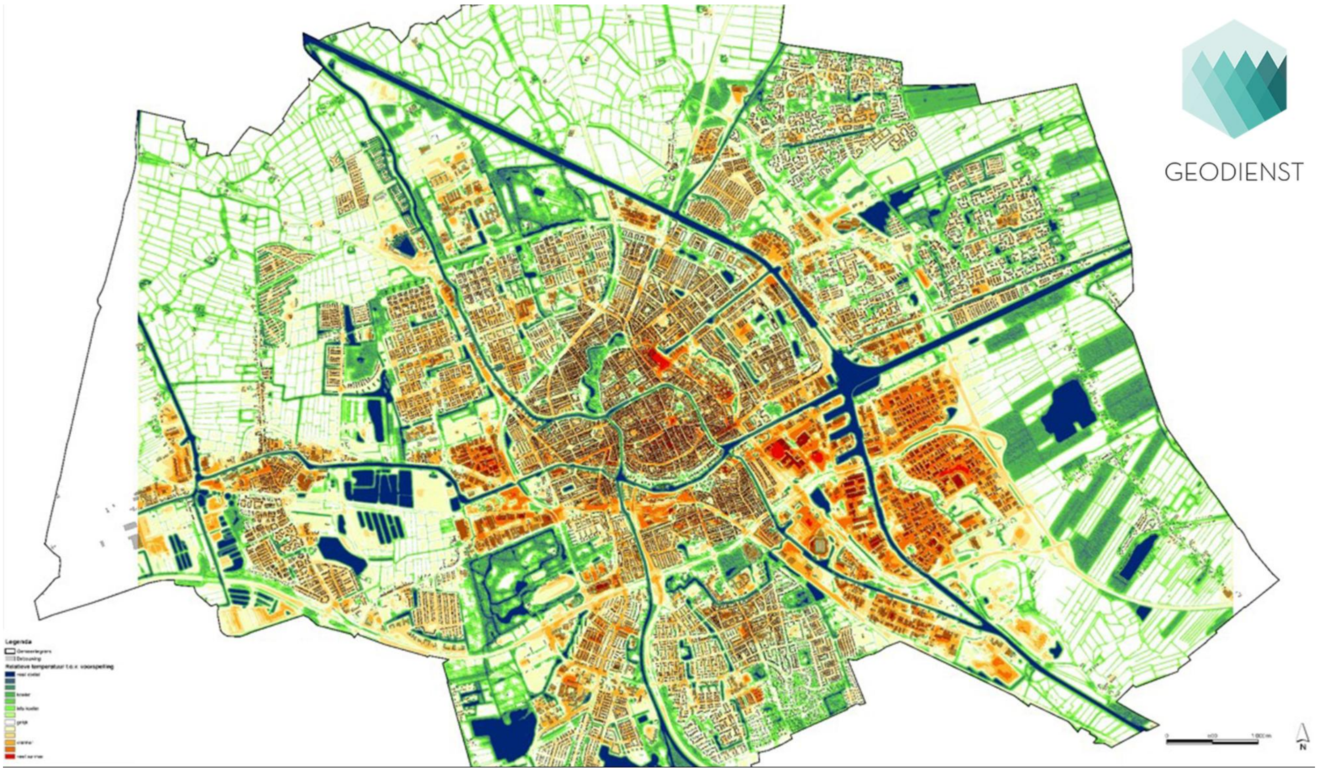


Figure 3: Heat stress map.

4 Challenges and further development

Datasets are getting bigger and bigger, customers get more demanding and want fast and good results. The AHN3 is coming, with an even better resolution, so the model will improve. We can solve the bigger and bigger data problem by using our High Performance computing facilities and 3D visualisations to keep the overview. Currently used other models lack the ability to perform on a large scale.
Currently the high performance and visualisation part is mostly open source based, but the model from TAUW is not. We are working on improvements and building our own parts of the model. By doing this we are working towards an open source model and combining the best tools we can to get the best results

by running the model on the high performance clusters.

References

- ✔ Klok, T.M. (2012). *Modelling of stormwater overland flow in urban areas: Assessment of WOLK as an overland flow modelling tool* (Master Thesis) Retrieved from http://essay.utwente.nl/61450/
- ✔ RUG. (2014). *Disaster Risk Management - urban flooding, heatstress & earthquakes* [video]. Retrieved from https://www.youtube.com/watch?v=mzFahS9RJPA

A FOSS approach to Integrated Water Resource Management: the case study of Red-Thai Binh rivers system in Vietnam

Carolina Arias Munoz[1], Maria Antonia Brovelli[1], Simone Corti[1], Marco Micotti[2], Rodolfo Soncini Sessa[2], Enrico Weber[3],

[1] HydroinformaticsLab, Dept. Civil and Environmental Engineering, Politecnico di Milano, Italy

[2]Dept. Electronics, Information, and Bioengineering, Politecnico di Milano, Italy

[3]Fondazione Politecnico di Milano, Italy

Abstract

The paper describes the development of a FOSS Geoportal targeted to manage information related to water resource management projects. It gives important insights on Spatial Data Infrastructures (SDI) using FOSS, demonstrating that are a feasible and effective solution to data and metadata collection, storage, sharing and visualization in water resources management, using open international standards. A first application of the Geoportal is also presented on the case study of Red River in Vietnam.

Keywords

Spatial Data Infrastructures, water resources management, time series data, multi attribute analysis, alternatives evaluation

1 Introduction

Integrated Water Resource Management paradigm (GWP 2003) has been adopted in the last years by several countries and Institution (i.e. Water Framework Directive of European Union, 2000) all over the world. Hence a number of tools, methods and software have been developed to support this paradigm, dealing with system modeling, alternatives design and evaluation (Soncini Sessa et al., 2007). Topic of present research is the development of a FOSS Geoportal specifically targeted to manage information related to water resource management projects. The Geoportal has been developed by the authors inside the Integrated Management of Red River System (IMRR) project and it is actually in a testing stage.

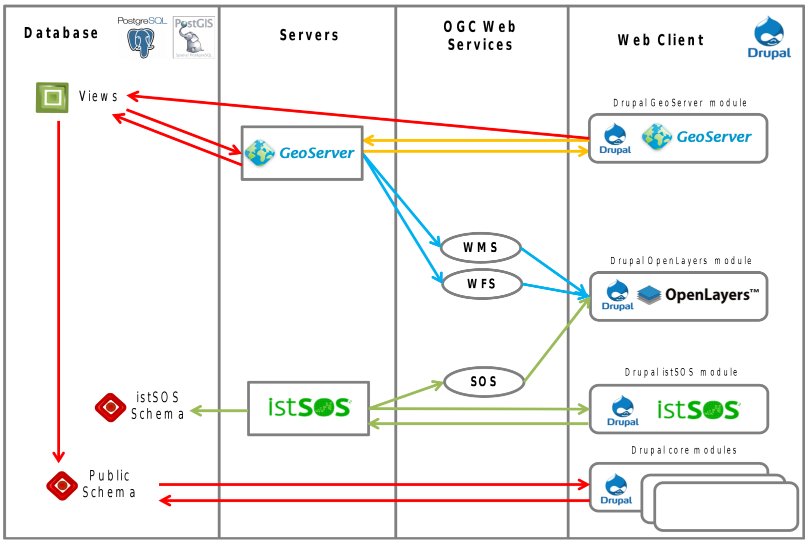
Red river has a cross-boundary basin, mainly divided between China and Vietnam, where it flows through 25 provinces in the North, including Hanoi capital, affecting a population of 26 million persons and feeding a canals network that supplies water for nearly 1.1 million hectares of agricultural lands. In the Vietnamese part of the catchment 4 big multipurpose reservoirs are in operation, targeted to hydro-power production, big floods prevention and water supply support.

The IMRR project aims at developing and promoting strategies for the sustainable management of the Red-Thai Binh Rivers System through coordinated decision-making, supported by modeling and optimization tools and through capacity building of local authorities in the water sector. The Geoportal allows to visualize, analyze and compare different planning alternatives for the management of the Red River water system in Vietnam, as a result of the Participatory Integrated Planning procedure (Soncini Sessa et al., 2007) applied in the IMRR project. The Geoportal characterizes for its effective and not-technical approach to the knowledge base of the project, taking advantage of the component's spatial distribution to improve the use of IMRR project outcomes.

2 The Geoportal Architecture

The Geoportal has been developed like a "light weighted" SDI, which components are reported in Figure 1 and describe hereafter. In the backend, the Database Management Systems PostgreSQL with the spatial extension PostGIS is used to store all data, including time series and geospatial data. Geoserver is used as the GIS Server, creating and managing several types of OGC Web Services such as Web Map Service and Web Feature Service. IstSOS[1] server allows the managing and dispatching observations from monitoring sensors and other points of interest according to the OGC Sensor Observation Service standard. istSOS connects also with the PosgreSQL database using a distinct data schema for each scenario considered in the project.

Figure 1: Geoportal architecture



On the client side we choose to use Drupal[2] CMS (Content Management System), taking advantage of the existing Geoserver and OpenLayers modules (similar to Cartaro[3] distribution) and developing two new modules:

1 https://geoservice.ist.supsi.ch/projects/istsos/index.php/Welcome_to_istSOS_project
2 https://www.drupal.org/
3 http://cartaro.org/

- a general use module[4], to connect istSOS and Drupal and to manage multiple time series and different scenarios inside Drupal;
- a project-customized module, to manage interaction between map, pop-up, time-series charts and search functionalities.

These modules, allow to interact with the different software within the Drupal application, without need to access directly to software such as OpenLayers or istSOS to manage data, because all the main functionalities are available thought the user-friendly interface.
Being a CMS, Drupal offers the possibility to have different user accounts and permissions (i.e. general public, project stakeholders and project owners), gaining additional advantage such as the extendability, namely the possibility to create different plugins, to organize and query contents and information within the website.

3 Geoportal functionalities

The Geoportal have been included several functionalities, addressing specific tasks and involving specific open source tools:

- *Project Hierarchy*: the Evaluation Hierarchy[5], used to describe and identify interests in the system. It is represented through a dynamic hierarchical tree chart and managed as Drupal nodes: every changes in a node is reflected directly in the chart and in the contents organization.
- *Components, Sector and Indicators pages*: Description pages for each element of the evaluation hierarchy, showed as dynamic popups, accessible through the geographic position of each item on the map (see. Fig. 1). In the popup there are also a list of related content (i.e. the list of indicators for each sector) and a dynamic chart, integrated with IstSOS time-series, which shows data related to each indicators (Fig. 2). This functionality is used to analyze historical data, models outcomes or to compare indicators results from different alternatives.
- *Spatial indicators:* Openlayers services and IstSOS data integration. It allows to represent spatially distributed indicators and to visualize performance of different alternatives directly on the map.
- *Special chart pages*: two pages, powered with a two panels chart (*Charts comparison page*) or a customizable number of small charts (*Dashboard page, Fig. 3*). Detailed analysis can be developed in thes pages, comparing Geoportal data (time-series, indicators at different time frequency, ..).
- *Scenarios*: a switcher integrated with IstSOS service to easily manage different scenarios (climatic, socio-economic,..). It is used to browse the different data needed to analyze and manage the water system.
- *Simulation*: a front-end page to execute simulation of the system getting user custom inputs, using Octave as simulation engine[6]. It is used to

4 https://www.drupal.org/sandbox/istsos/2149739
5 Sectors, criteria and indicators as in the classical approach of Multi Attribute Value Theory - MAVT (Keeney and Raiffa, 1976) the methodology used in the PIP procedure
6 Octave runs outside of the geoportal, but a system of automatic uploads provide the exchange of information with

assess water system behavior under different conditions.

- *Notification module*: A custom Drupal content form to get georeferenced feedbacks from registered users.

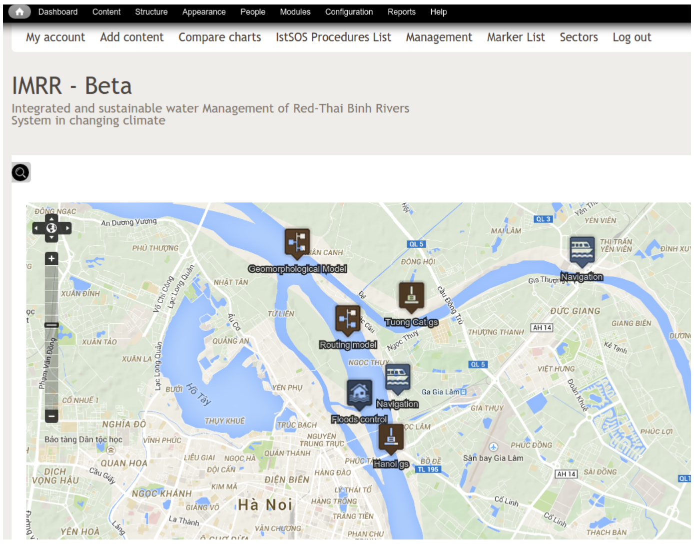


Figure 2: Geoportal home page

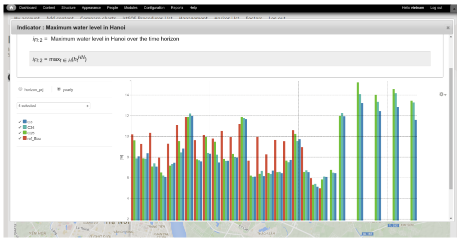


Figure 3: Indicator dynamic chart

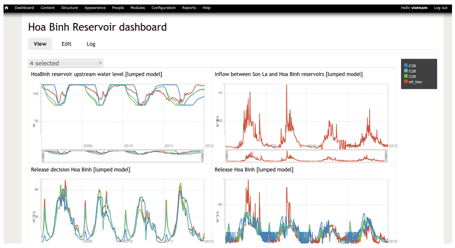


Figure 4: Dashboard

4 Conclusion

Nowadays there are mature and well-tested FOSS for the creation of robust and reliable webGIS services. Our research suggest that FOSS offer features and capabilities comparable, and in some cases superior, to their commercial equivalents for the development of the next generation of Web-based information systems in the water management domain.
Collection, storage, sharing and geographical visualization of data is a crucial task for water resources management in complex contexts. The Geoportal allows to differentiate access permission amongst users, to manage contents related to a project, to visualize, analyze and compare different planning alternatives providing quick and easy access to all its functionalities both for expert and non expert users. The heterogeneous nature of the project data requires the combination of different geospatial data services (Web Map Service WMS, T, Web Feature Service WFS, Sensor Observations Service SOS), servers (Geoserver, istSOS) and interface technologies (OpenLayers, Drupal) enabling interoperability of all complex resources data types.
IMRR project is in the final stage: the Beta version of the Geoportal is still under development and not yet released, but it is managing a significant amount of data, since more than 100 alternatives policies of reservoirs management have been designed over a time horizon of 30 years, with 3 different scenarios (standard conditions, extreme inflows conditions and climate change). Effects of each alternative is described with temporal trajectories of more than 20

the Geoportal.

variables and through 24 indicators, computed at daily, yearly and horizon time scale and related to 5 sectors affected by Red River management: Floods, Hydropower production, Water supply, Environment, Navigation.
The last IMRR Stakeholders meeting, scheduled for September 2015 will be informed with this huge database, accessible through a fully functional Geoportal, able to properly support results analysis and discussion.

References

- ✔ Balbo, S., Boccardo, P., Dalmasso, S., & Pasquali, P. (2014). a Public Platform for Geospatial Data Sharing for Disaster Risk Management. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W3(February), 189–195. doi:10.5194/isprsarchives-XL-5-W3-189-201
- ✔ EU Water Framework Directive (2000). Directive 2000/60/ EC of the European Parliament and the Council of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities.
- ✔ Global Water Partnership (2003). Toolbox Integrated Water Resource Management.
- ✔ Gutierrez, F. V, Manso, M. A., Lang, D. H., Wachowicz, M., Bernabe, M. A., Strauch, W., … Service, I. (2007). Designing web-enabled services to provide damage estimation maps caused by natural hazards, (1990).
- ✔ Keeney, R. L., & Raïffa, H. (1976). Decisions with multiple objectives: preferences and value tradeoffs. Wiley.
- ✔ Pickle, E. (2012). GeoNode – Incorporating Geovisualization in Support of SDI, 125–128.B
- ✔ Soncini-Sessa, R., Weber, E., & Castelletti, A. (2007). Integrated and participatory water resources management-theory (Vol. 1). Elsevier.

GRASS GIS processing to detect thermal anomalies with TABI sensor

Carmine Massarelli, Raffaella Matarrese, Vito Felice Uricchio, Michele Vurro

Water Research Institute - Italian National Research Council (IRSA-CNR), viale De Blasio 5, Bari, Italy

Abstract

In this study, we report the adopted methodology used in GRASS, a free open-source GIS software, which has allowed us to map surface water thermal anomalies and, consequently, to identify and locate coastal inflows in the *Mar Piccolo* sea (Taranto) part of the National Priority List site identified by the National Program of Environmental Remediation and Restoration. An ongoing work, where to apply the same procedure, is being carried out on Bari province coast and effimeral streams.

Keywords

GRASS GIS, thermal anomalies, coastal, effimeral streams

1 Introduction

With the advent of increasingly powerful airborne thermal sensors, it is becoming easier to obtain synoptic information on areas otherwise difficult to monitor or areas that until now have been monitored only with medium/low spatial resolution as Landsat maps. Through the use of these maps, since several years, remote sensing techniques allow to identify (big) thermal anomalies on water cloves, providing useful information to detect, for example, oil spill evolution on sea (Hua, 1991) and to characterize and quantify spatial dynamics of currents in water (Hedger et al., 2007; Matarrese et al, 2004).

The Water Research Institute of the Italian National Research Council (IRSA-CNR), in 2013 and 2014, has been involved in an activity about identification of surface thermal anomalies related to legal or illegal *Mar Piccolo* inflows and to submarine springs (in Italian *Citri*). Successively, a similar activity has started on the coast of Bari province, in a scientific agreement that aims to detect all the illegal discharges in coastal waters.

Nowadays, there are only few cases in which these data have been used to support the action of local Authorities that, usually, base their considerations on image visual interpretation, a very time-consuming, subjective and expensive technique (considering also the cost of specialized software). Furthermore, the big amount of data storage that this images require is usually seen as a limit in their use. Open-source and free softwares can represent one possible solution of all the listed difficulties.

In this paper, we report the partial results obtained using the free and open-source GIS software GRASS v.6.4, applied to the thermal images from TABI-320 sensor providing, as output data, a map with high degree of accuracy of surface water thermal anomalies that, consequently, leading to the identification and location of the inflows, as well as manmade or natural watershed drains or submarine springs.
Furthermore, according to the results obtained, it will be possible for Authority to assess which maps can be useful for the most varied purposes.

2 Materials and methods

2.1 Study area

Two study areas have been considered. The first one, were the analysis has been already performed, is Mar Piccolo sea of Taranto (200km^2). The second study area is the whole coast (84 km x 900m) and the effimeral streams of the province of Bari (~350km x 900m). Both areas are in the Apulia region, South of Italy.

2.2 Data description

In order to have the most accurate thematic map of thermal anomalies, images from ITRES TABI-320 have been processed. The TABI-320 (Thermal Airborne Broadband Imager) is an airborne thermal mapping sensor. Its specifications are described in http://formosatrend.com/Brochure/TABI-320.pdf.
For the 2 cases study, two main campaigns of acquisition were performed: 13 images were collected between March and April 2013 to cover the whole area of Mar Piccolo with 1m^2 of spatial resolution; 29 images of the Bari coast were acquired in November 2014 with pixels of 1.30mx1.30m. The total amount of data was of about 10 Gb.

2.3 Data processing

The following procedure, used for this study, has been developed by the mean of the GRASS GIS (Geographic Resources Analysis Support System), (http://grass.osgeo.org/).
In order to locate thermal anomalies by the airborne images collected, TABI data have been first geo-referenced using the **i.ortho.photo** module (Rocchini et al, 2012).
Then the **i.image.mosaic** module allowed to create the mosaic of the whole zone.
In order to find thermal anomalies along the coast and into the open sea, the whole TABI-320 mosaic image was ‘clumped’. It means that the image was **recategorized** in a raster map by grouping cells that form physically discrete areas into unique categories (Neteler and Mitasova, 2008). This step was necessary to simplify the calculations in order to perform processing focused on small groups of values rather than on the entire original raster map. This was applied only to TABI-320 thermal images.
As second step, the raster data just obtained was transformed into an isolines vector data. This **contour** method, often used to identify temperature and elevation, was able to determine the minimum and the maximum isoline values and the interval (step parameter) for the given raster map: in such application

we used 16°C as minimum value, 28°C as maximum value, and 0.25°C as interval.

Finally, isolines were simplified and smoothed by a **generalization**. The simplification process reduces the complexity of vector features decreasing the number of vertices for each line. The smoothing is a process applied to produce a smoother approximate line than the original. The algorithm used is the *Chaikin*'s one (Chaikin, 1974). This algorithm approximates the given line very well for the purpose of this study because of *Chaikin*'s curve has been shown to be equivalent to a quadratic B-spline curve and is a useful to the analytical definition of B-splines and provides a simple, elegant curve drawing mechanism (Kennet, 1999).

3 Partial results

As result obtained for the *Mar Piccolo* sea, we produced a map of thermal anomalies around the coast surprisingly coincident with the inflows detected during a survey carried out in situ to validate the product (Fig. 1).

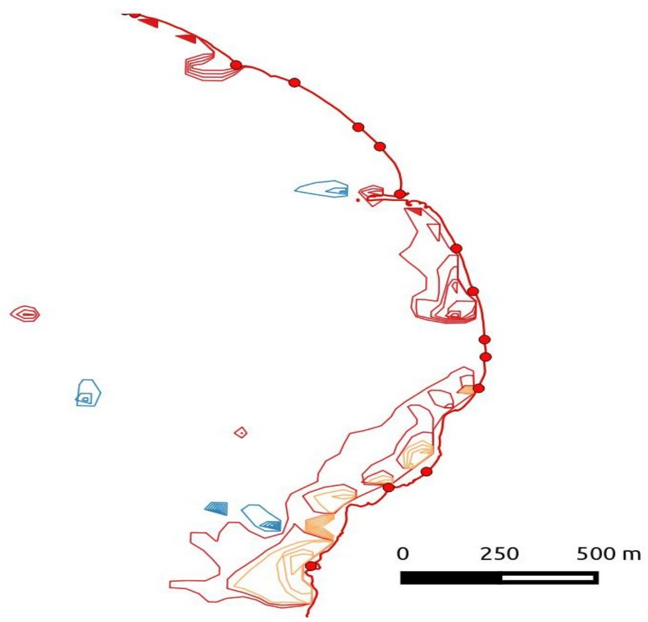


Figure 1: Example of Taranto sea thermal anomalies found by GRASS.

4 Preliminary conclusions

GRASS – GIS allowed to identify most of the thermal anomalies in the Mar Piccolo Sea. Its skills demonstrated that it is possible to perform complex and big data processing with a step by step procedure, considered before an exclusive prerogative of expensive licensed software.

The advantage in using open-source software is due to the possibility for the user to create its own subroutines in order to manage huge amount of data. A non-expert user, since the community that uses this particular software is very active, can easily find help or subroutines already written to satisfy his needs.

Furthermore, the gratuity of GRASS-GIS, is another important added benefit. Working with public administrations, it is relevant to emphasize all the possible advantages deriving from the combination of remote sensed data and open-source softwares, in order to better combine research and common needs.

Acknoledgement
This work was funded by Regional Agency for Environmental Protection and Prevention in the Puglia region (ARPA Puglia) for the *Mar Piccolo* sea and Province of Bari for the effimeral streams and coast of Bari.

References

✔ Chaikin, G., 1974. An algorithm for high speed curve generation. *Computer Graphics and Image Processing*, 3, 346–349
✔ Hedger, R.D., Malthus, T.J., Folkard ,A.M., Atkinson, P.M., (2007). Spatial dynamics of estuarine water surface temperature from airborne remote sensing. *Estuarine, Coastal and Shelf Science*, Volume 71, Issues 3–4, p. 608-615, ISSN 0272-7714, http://dx.doi.org/10.1016/j.ecss.2006.09.009
✔ Hua, H., (1991). Remote Sensing Of Oil Spills In Thermal Infrared - Contour Line Effect. *Geoscience and Remote Sensing Symposium. IGARSS '91. Remote Sensing: Global Monitoring for Earth Management., International* , vol.3, no., pp.1315,1317, doi: 10.1109/IGARSS.1991.579322
✔ Hus, L., 1991. Remote sensing of oil spills in thermal infrared - Contour line effect. *Digest - International Geoscience and Remote Sensing Symposium (IGARSS)*, 3, 1315-1317.

✔ Kennet, I. J., 1999. Chaikin's Algorithms for curves. *On-Line Geometric Modeling Notes*. Computer Science Department, University of California, Davis.
✔ Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T.,Barloon, P.J. and Goetz, A.F.H., 1993. The spectral image-processing system (Sips) – interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44, 145–163
✔ Matarrese R., Vermote E.F., Kearney M., 2004. Impact of Hurricane Isabel on the water properties of the Chesapeake Bay area. *Proc. Of Hurricane Isabel in Perspective Conference,* November 15-17, 2004, Linthicum Heights, Maryland.
✔ Neteler M. and Mitasova H., 2008. *Open Source GIS. A GRASS Gis Approach*. Third Edition. Springer
✔ Spizzico, M. and Tinelli, R., 1986. *Hydrogeology of Galese spring, Mar Piccolo of Taranto (South Italy)*. In Proc. 9th Salt Water Intrusion Meeting Delft, Water Management Group, Delft University of Technology, Delft, pp. 85–97.
✔ U.S. Environmental Protection Agency, 2006, *Guidance on Systematic Planning Using the Data Quality Objectives Process: EPA QA/G-4*. Office of Environmental Information, EPA/240/B-06/001, Feb. 2006.
✔ Weng Q., Lu D. & Schubring, J. 2004. *Estimation of land surface*

temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment, Elsevier Science Inc., Box 882 New York NY 10159 USA, 89, 467-483

✔ Rocchini, D., Metz, M., Frigeri, A., Delucchi, L., Marcantonio, M., Neteler, M. (2012). Robust rectification of aerial photographs in an open source environment. Computers & Geosciences 39 (2): 145-151.

FloodRisk: a QGIS plugin for flood consequences estimation

Leonardo Mancusi[1], Raffaele Albano[2] and Aurelia Sole[3]

[1] Ricerca sul Sistema Energetico SpA, Milano, Italy

[2] DICEM, University of Basilicata, Matera, Italy

[2] School of Engineering, University of Basilicata, Potenza, Italy

Abstract

In times of increasing disaster losses, the reduction (or mitigation) of consequences of natural hazard needs to be effective and efficient. An in-depth understanding of the effects of disasters, is required in order to develop forms of sustainable risk management as well as risk mitigation and adaptation strategies (Directive 2007/60/CE; IPCC, 2012). This is especially true given the limited financial resources available.

In this context, we developed a prototype tool: FloodRisk, a QGIS plugin that provides the assessment of flood consequences, in terms of loss of life and direct economic damages.

The proposed cost assessment tool aims to be used for decision support systems and policy development of risk management.

Keywords

Flood risk, risk map, open source software, desktop GIS, QGIS

1 Introduction

According to the IPCC (2012), one of the most important consequences of climate change will be the increase in the frequency and magnitude of extreme events such as floods.

Flood risk management aims to reduce the impact of floods. The quantification and evaluation of flood consequences is one of the most important factor to be considered in deciding how to reduce flood damages and to evaluate alternative intervention strategies in terms of their relative benefits and costs (Albano et al, 2014). In this context, an open source and free analysis toolbox, part of the open-source geographic information system Quantum GIS, was developed for estimating flood impacts due to flooding and, hence, to help authorities better understand and manage flood risk. These tool are not intended to be business-ready software applications, however it is usable by third parties for evaluation and demonstration purposes.

The proposed prototype software tool, called "FloodRisk", is able to calculate and view the number of people affected and direct damages to properties caused by any flood scenario in any chosen area.

For example, if there has been the need for interventions to ensure the safety of a number of hydraulic structures (eg. dams or levees), which could result in

a hazard downstream, it is necessary to decide how selecting the highest-priority of interventions. In these cases, the assessment of the consequences of any failures or accidents, it is an important element to support the choice of priority interventions. This is especially true given the limited financial resources available.
In this context, furthermore, the spatial distribution of people and property within the flooding area can cause different consequences and therefore a GIS tool is particularly useful to make an analysis that takes into account the actual characteristics of the territorial system downstream the hydraulic structure.
After this brief introduction, sect. 2 focuses on the conceptual framework and the utilized methodology while sect. 3 on the software tool called "FloodRisk".

2 “FloodRisk” tool Conceptual Framework

2.1 General "FloodRisk" Framework for flood risk assessment

The term "*risk*" has a range of meanings and multiple dimensions relating to safety, economic, environmental and social issues.
The terminology of "*risk*" has been developed across a wide range of disciplines and activities, therefore there is potential for misunderstanding in technical terminology associated with risk assessment.
Most important is the distinction that is drawn between the words "*hazard*" and "*risk*". To understand the linkage between hazard and risk it is useful to consider a simple conceptual model: for a risk to arise there must be hazard that consists of a ’source’ or initiator event (i.e. high rainfall); a ’receptor’ (e.g. flood plain properties) and a susceptibility of receptor to be damaged (*vulnerability*), (Gouldby & Samuels, 2005) .
A hazard does not automatically lead to a harmful outcome, but identification of a hazard does mean that there is a possibility of harm occurring. The actual harm depend upon the exposure to the hazard and the characteristics of the receptor. In practice exposure and vulnerability are often captured in the assessment of the consequences; thus risk can be viewed in simple terms as “probability times damage”, and thus describes the expected damage that can occur or will be exceeded with a certain probability in a certain period (Merz et al, 2010).

Risk = (Probability) x (Consequence)

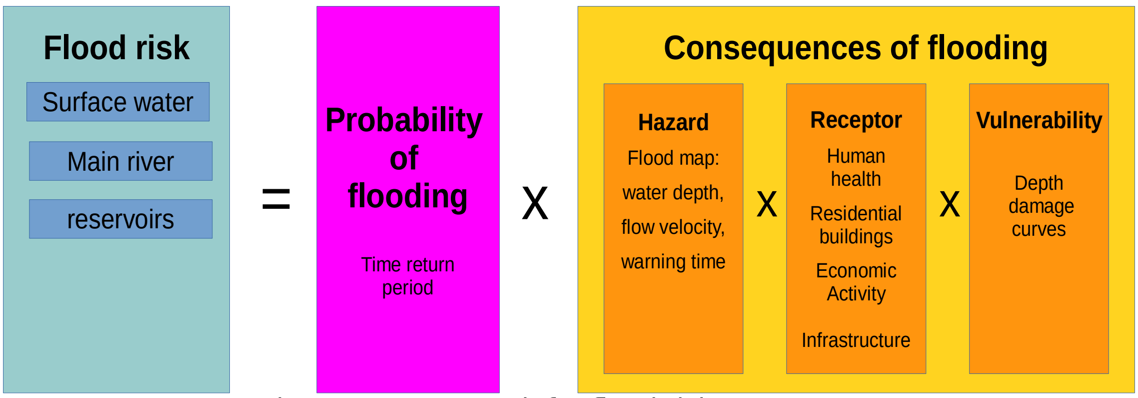


Figure 1 :Framework for flood risk assessment

The conceptual framework for flood risk assessment it is represented in Figure 1. The proposed GIS tool, within this framework, deals with the assessment of the consequences, and therefore it helps to perform the operations described in the last part of the graph. Therefore, "FloodRisk" can support the identification of people and assets at risk, the planning and evaluation of effective flood mitigation and control measures, the creation of flood risk maps for awareness raising.

2.2 Required data for consequences estimation

The tool performs a simple risk assessment in which it consider fixed event scenarios where the probability of each scenario is estimated separately and it calculates the consequences deterministically.
The key data required, described in the following, therefore concern: hazard, receptor and vulnerability.

2.2.1 Hazard

The data needed to quantify the flood hazard are the inundation map detailing the characteristics of the flood. "FloodRisk" tool requires specifically maps of:

- maximum depth values;
- maximum velocity values.

These data are the output from a 2D hydraulic model or integrated 1D-2D model. Outputs generated from a 2D hydraulic model, indeed, can include maximum depth and maximum velocity values for every inundated grid cell in the study area. Using a GIS, depth grids can be multiplied by velocity grids to obtain maximum "Peak Unit Flow rate (DV)" values, that can be calculated also by the ratio of discharge and water depth. These DV values are then categorized into ranges of values which define low, medium, and high severity zones.
Since timely flood warnings can save lives, warning time is a very important data to assess the consequences for people.
Therefore the tool needs as input the information of the zones with different warning time, called here:

- warning time map.

The warning time indicates the amount of time between the reception of a warning and the instant in which the population of each structure could be affected by the flood event, i.e. the amount of time in which the population of each structure can mobilize or adopt mitigation measurements. It may be different from zones even if the alarm should be launched instantly for all people, because there is a propagation time of the flood, and, then, the territories are flooded at different times. Indeed, the warning time depends also by the propagation time of flood wave that can be calculated by using a hydraulic model.

2.2.2 Receptor

The quantification of the receptors, called exposure, refers to people, assets and activities, threatened or potentially threatened by a hazard.
"FloodRisk" tool stores exposure data of the study area in a geo-database. The dataset must consists of the following maps:

- the polygon boundary of the study area;
- census map of population;
- buildings and/or land use map;
- infrastuctures lines maps (eg. roads, railways etc.) .

An essential infromation on the buldings map is the classification by occupancy type (eg. residential, commercial, industrial buildings, etc.). Indeed, the most widely used tool for estimating damage before an event, i.e. the vulnerability curves/functions, (which are described in the next paragraph), relations an hazard parameter, generally the flood depth, with the specific occupancy type of the exposed elements.

2.2.3 Vulnerability

Vulnerability is a characteristic of a system that describes its potential to be harmed.
Catastrophic floods, such as those by dam-break or levee failure, can cause more fatalities: for the quantification of the vulnerability, in this case, it is generally adopted the parameter "*fatality rate*" that is the percentage of population at risk who die. Fatality rates are based on *flood severity*, *warning time*, and *warning quality*.
There are empirical methodologies of literature calibrated on historical cases that provide values of fatality rates as a function of the parameters mentioned above. FloodRisk tool contains some of these tables of values.
A flood can cause, also, many types of economic damages that can be classified in a variety of ways: the tool is able to calculate the tangible direct physical damages, that are, the damages result from the actions of floodwaters on property and structures.
Tangible damages are usually quantified and measured as monetary losses.
Flood damages depend on many variables. These variables might include depth of water, velocity of floodwaters, duration of flood, sediment load, contamination. But flood damage to structure is strongly dependent on the water depth of a flood (Merz et al., 2010): depth of flooding is always expected to be associated with an increase in flood damages; at some point the water will be so deep as to have caused all the damage possible and the damage curve would eventually go horizontal.

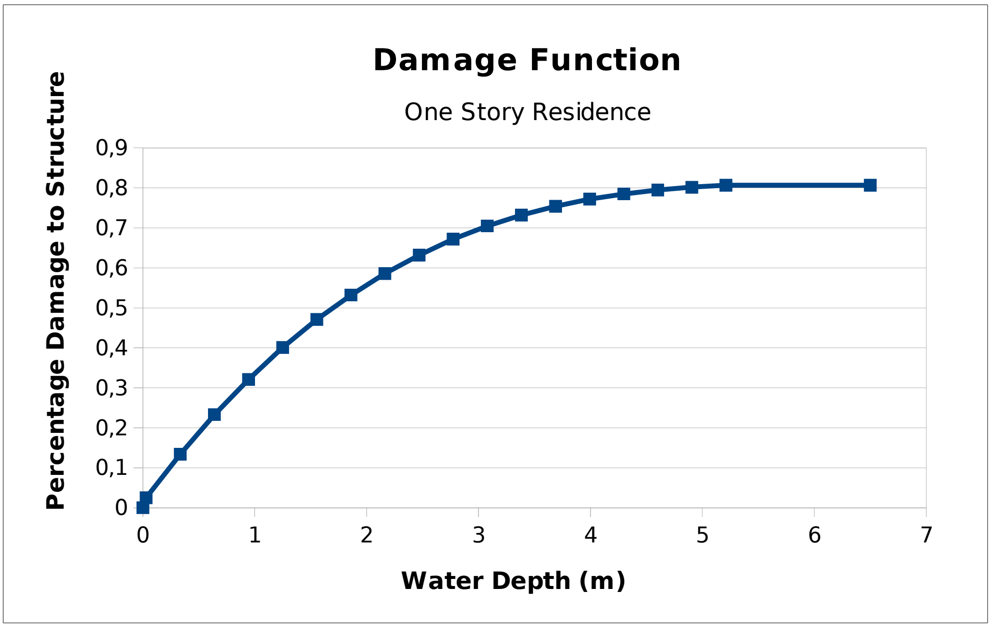


Figure 2 :Example of depth-damage function

Figure 2 above shows a graph of the mean (%) values of damage for single family residential structures. This curve represents structure damages only, content damages are evaluated separately using a depth-percent damage curve for the contents of a house.
To calculate the monetary value of the damage, percentages are multiplied by the maximum damage value of proprieties.
For each class of asset the tool needs to be provided, as input, with the two curves of depth-percent damage for structure and contents. The values of assets are included as properties of the individual elements of the map in unit terms (eg. Euro / sq m).
The tool contains default curves and values derived from the literature. Users can develop, and upload in the tool, their own depth-damage functions deduced following standard methods. However, this can be expensive and time-consuming, and is only recommended for very high value properties that dominate local damage estimates and/or for large, mixed used sites such as hospitals, power stations and water treatment works.

2.3 Utilized methodologies

This section focus on the description of the methods currently available for the direct flood impact estimation through “FloodRisk” tool.
The direct flood impact can subdivide in tangible and intangible (non-market) damage. On one hand, the direct tangible damage covers, for example, damage to structures and contents, destruction of infrastructure such as roads, rail roads. On the other hand, the evaluation of the direct intangible (non-market) damage can be ascribe, for example, to the estimation of loss of life. The methods chosen to evaluate the aforementioned direct impact is described in the following sub-sections.

<u>2.3.1 Direct economic damages estimation</u>

A common method to to assess potential tangible direct physical damages is

the application of depth–damage curves. The damage to residential, commercial and industrial property can be assessed using depth–damage curves, which describe the relationship between levels of inundation and damage incurred.
Surveyed damage estimates for a range of flood levels are essential to their production. It is recommended, where possible, to use depth–damage curves that represent local conditions and the types of buildings present. Where this is not possible because locally specific data does not exist, use available depth–damage curves produced as a result of previous flood damage studies.
There are two different approaches for the application of depth–damage curves. These are the land use based approach and the object based approach. Both methods use land use types as an input variable. The object based approach adding more detail to urban areas; this has been done by replacing the urban land use type of the classical land use based approach through the buildings map. FloodRisk tool can perform the analysis according both approaches, depending on the type of data that the user loads into geo-database. The important thing is that depending on the type variable used for land use/buildings map to be associated with the proper depth-damage curves.
In the GIS-systems, the input variables as land use or buildings map, inundation depth, associated depth-damage curves, maximum damage value, can easily be aggregated to perform an analysis of the amount of damage that would occur, should a flooding happen.

2.3.2 Population at Risk and Loss of Life estimation

Different methods currently existing fo the estimate the loss of life. "FloodRisk" utilized simple and parsimonious methods based on the contributing factors that could be generally available: hazard factors, such as water depth and water velocity, the general preparadness of the society, such as existence of public education on flood risk, warning and communicatation system, coordination between emergency agencies and authorities, time of day, and warning factors, such as warning time. Other more sofisticated models take into account other contributing factors, for example, the physical condition of the people potentially invoved in the calamitous event, such as age, gender and so on, or the water quality and temperature, floating debris, that are of great interest but they are not easy to collect.
In "FloodRisk" the first step for the loss of life estimation is to obtain the map of the population at risk by superimposing the inundation map and the population density map. The first is an input data, the second was created by converting vector files of population census into raster files. In the conversion is assumed, for simplicity, a uniform distribution of the total population for each polygon of census map.
Once calculated the map of the population at risk, the number of potential fatalities is obtained by multiplying population at risk time the fatality rate.
Values of fatality rates are found in several sources in the literature.
The tool allows the user the choice of the fatality rate between two of the available scientific methods:

- US Department of Homeland Security (DHS 2011)

- SUFRI Methodology for pluvial and river flooding risk assessment (Escuder Bueno et al, 2011)

The first concerns the potential risks associated with the failure or disruption of dams, the second for river flooding.
For both methodologies fatality rates depend on warning time, flood severity and on the effectiveness in the evacuation of the population.
Warning time is an input data. ***Flood severity*** is a parameter that assumes the values ***low***, ***medium*** and ***high*** in function of the parameter hydraulic *DV* that is the depth of water multiplied by velocity and it is representative of the general level of destructiveness that would be caused by the flooding.
The third factor is evaluated differently by the two methodologies.
In DHS it is defined as "Flood Severity Understanding" and describe the degree to which the events that are about to unfold are understood.
The flood severity understanding categories are as follows:

- ***Vague*** Understanding of Flood Severity: warning issuers have not yet seen the dam fail or do not comprehend the true magnitude of the flood that is about to ensue.
- ***Precise*** Understanding of Flood Severity: warning issuers have an excellent understanding of the flooding due to observations of the flooding

In SUFRI methodology the third factor is a degree of readiness of population. Ten categories (C1 to C10) are established, dependig on the existence of public education on flood risk, warning and communicatation system, and, coordination between emergency agencies and authorities.

3 Flood Risk tool Interface

The assessment of the consequences of flooding requires the analysis of the extent and distribution of intensity of the hazard (for example, depth of the water and its velocity) and the overlap with the spatial distribution of people and property exposed.
Geographical information systems (GIS) tools are ideal to manage spatial information, providing adequate spatial processing and visualization of results. For this reason, we chosen to adopt a GIS as a basis for the development of a tool for assessing flood risk. Moreover, in this study, a free and open-source GIS solution is utilized as platform to integrate the described tool, called "Floodrisk", taking into account the quality and dissemination nowadays reached by Geographic Free/Open-Source Software (GFOSS) and having it in order to make available the results of the project.
Analyzing different desktop free and open sources GIS projects, we chose Quantum GIS (QGIS): the main advantage of QGIS relies on the easiness and quickness in developing new plug-ins, using Python language. Therefore, this project was developed in QGIS platform and the interface was created in Python.
FloodRisk tool is therefore a plug-in for QGIS, the application groups in a unique toolbar all the procedures needed to produce the assessment of the consequences of the floods.

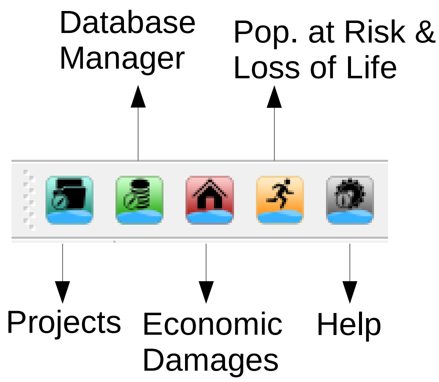


Figure 3 : Application toolbar of FloodRisk plug-in

The toolbar developed for this application is presented in Figure 3. Each button is linked to a window with several options, such as menus, labels, edition windows, combo boxes, and simple buttons, such as Ok, Close, and Help, among others, that help the user to access input and output directories.

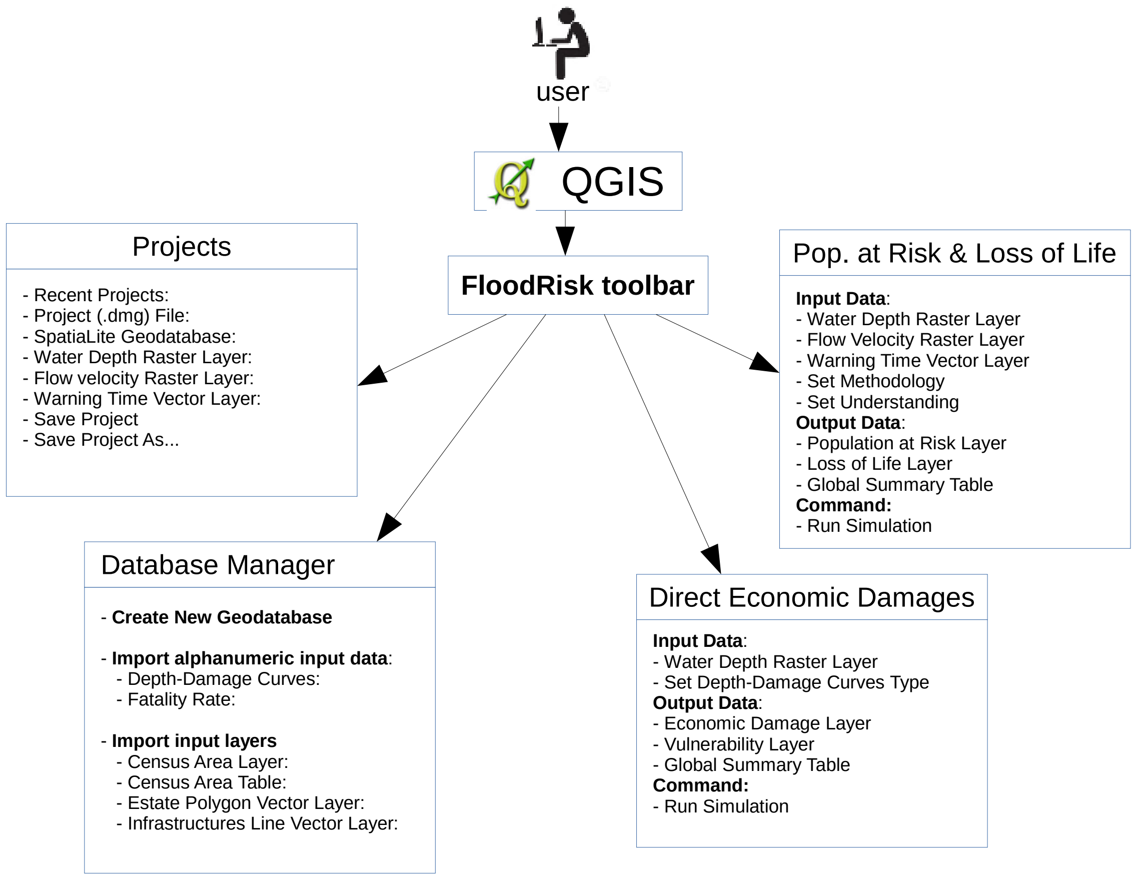


Figure 4 : GUI components

As shown in Figure 4, from the toolbar is possible to activate four windows, two of which for data management and two to run core modules.

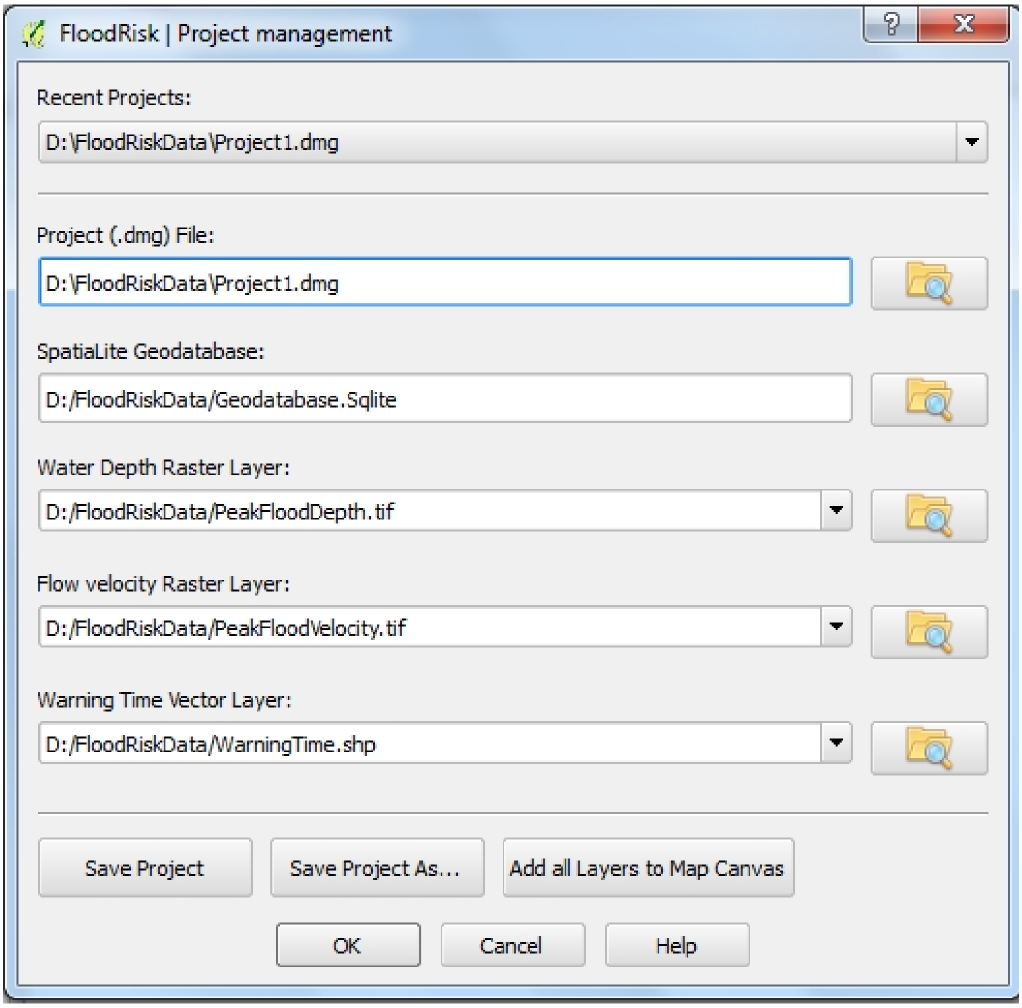


Figure 5 : GUI of project management

Figure 5 shows the first window: it is the interface for the project file management: this is an xml file in which are stored the pathnames of all the files that make up the project. In figure 6 is shown an example of a file.

```
<?xml version="1.0" encoding="UTF-8"?>
<FloodRisk>
  <General>
    <File Button="FilePeakFloodDepth" name="D:/FloodRiskData/PeakFloodDepth.tif"/>
    <File Button="FilePeakFloodVelocity" name="D:/FloodRiskData/PeakFloodVelocity.tif"/>
    <File Button="FileWarningTime" name="D:/FloodRiskData/WarningTime.shp"/>
    <File Button="FileGeodatabase" name="D:/FloodRiskData/Geodatabase.Sqlite"/>
  </General>
</FloodRisk>
```

Figure 6 : Example of project file (.dmg)*

In the example above, the project file contains the names and path of the input files, three of which are files concerning the hazard (flood depth map, flood velocity map and warning time map) and one refers to the geo-database, which contains all the data of receptors and their vulnerability. Hazard maps can be

any file type GDAL Raster Formats, except the warning time map which should be a polygons shapefile.
The data concerning population and assets at risk, are instead organized in a spatialite geo-database. The database manager (figure 7) allows you to create a new geodatabase (a Spatialite DB) having the expected data model and upload all the data into it. The alphanumeric data of the database tables are uploaded from csv file.

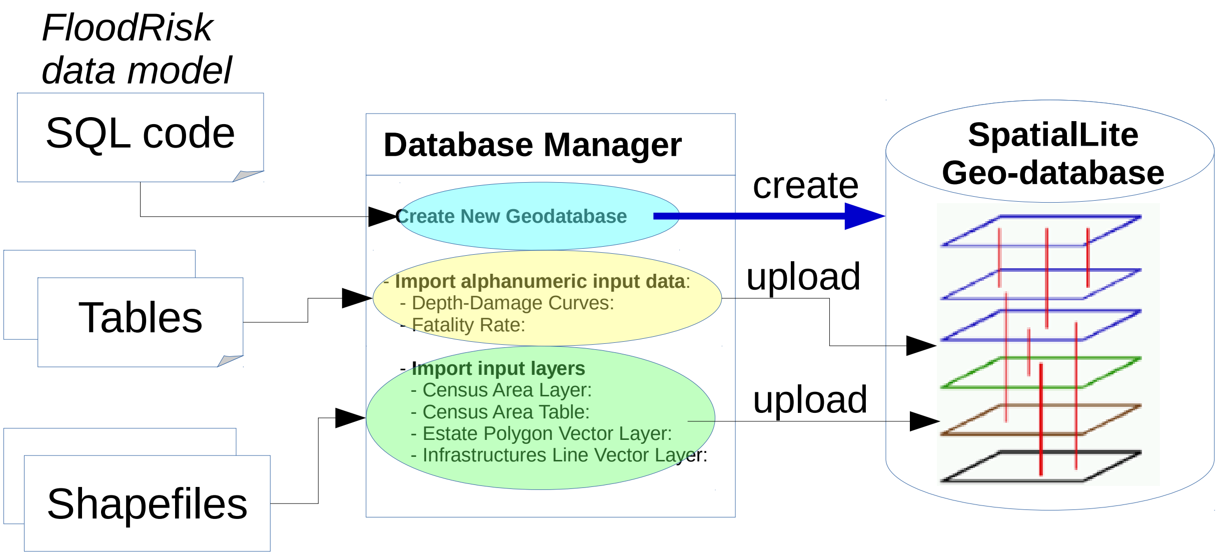


Figure 7 : Database Manager diagram

Geographic data of exposures, which are uploaded from shapefile and stored in the geodatabase, are vector data type: the data detail depends on the scale of the cartography source. The system thus does not have its own scale but is flexible and able to manage different data sources: from maps of land use with a different degree of detail until detailed urban maps.
Data of exposures therefore may refer to maps of large areas or detailed local maps.
The scale of the final result depends not only by exposure but also from the detail of the maps of the hazard.
The type of data that the tool uses to quantify the hazard is of raster data.
The cell size of the raster Hazard define its scale. The vector data of exposure during the processing are also converted in raster having the same cell size.
Data set of exposure in a given area, which are stored in the geodatabase can be used for analysis with hazard data having more resolutions: for each case, vector data of the exposures are automatically converted to a raster with a request resolution.
Data of the vulnerability are stored in database tables and FloodRisk tool, using SQL queries, extract them according to the elaboration required.
Figure 8 shows a diagram of the inputs and outputs of the software components that perform the computation of the consequences for the population and the economic damages.
The software components of computing are python scripts that use open sources libraries, eg. GDAL and numpy, and perform processing using Map

Algebra techniques.

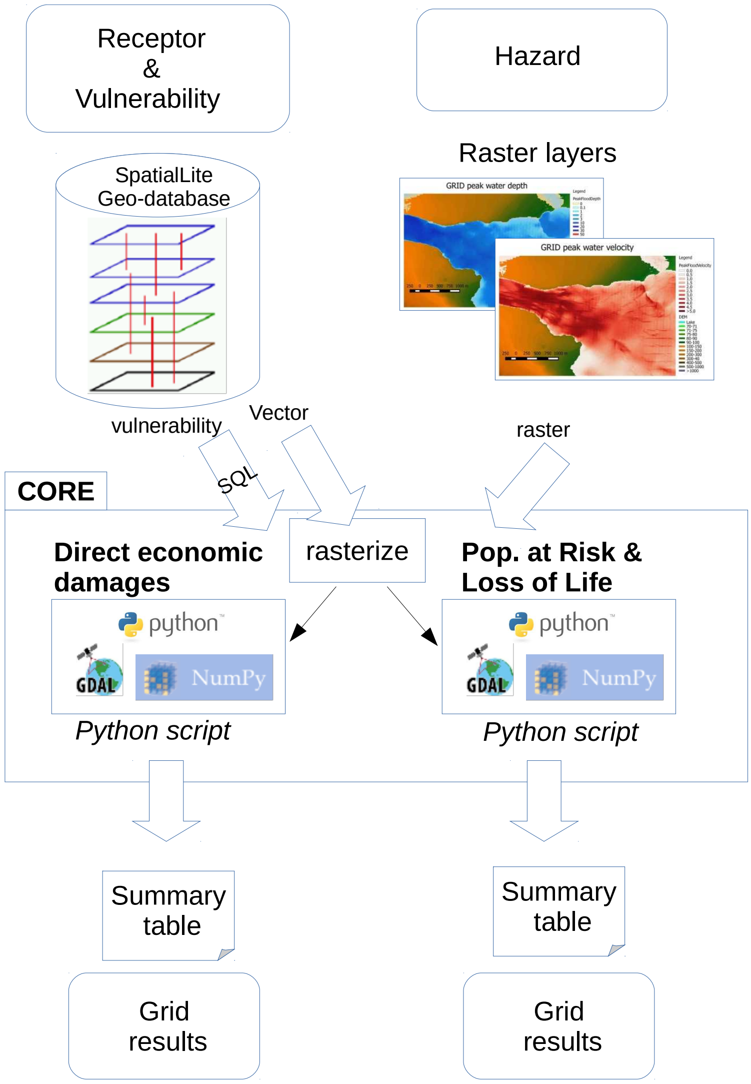


Figure 8 : CORE components & inputs and outputs

These scripts produce the grid maps of the spatial distribution of the population at risk, the potential loss of life and economic damage expected for the structures and for the content of the property at risk.
Two global summary tables of the results are provided, which are also displayed in graphical form (histogram).
Example of loss maps produced using FloodRisk engine and presenting the expected economic losses for a benchmarking study case, proposed by the organizer committee of the 12th ICOLD International Benchmark Workshop on Numerical Analysis of Dams, are presented in Fig. 9: for more details see Mancusi et al, 2013. The results were compared, by the formulators of the Benchmark, with those of five other international participants who have used the same input data but different methodologies and softwares: the results of comparisons are presented in the Proceedings of the Workshop (Zenz & Goldgruber 2013). By using the data set of the workshop and taking into account the results of the comparisons, we are preparing a new data set for a

tutorial of the plug-in that the authors aim to distribute in a repository.

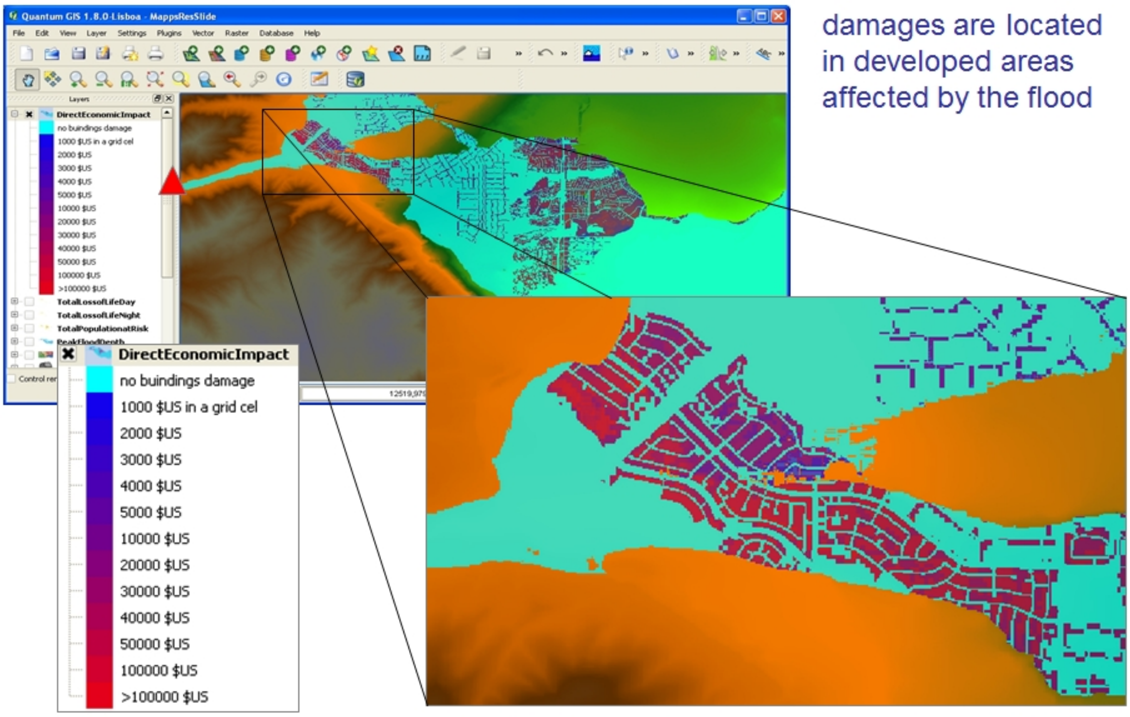


Figure 9 : Example of damages map

4 Conclusion

In this paper a QGIS plugin called "FloodRisk " is described. The tool can be used for the rapid and consistent evaluation of consequences of flood in terms of number of people at risk, number of loss of life and economic damages for residential, commercial, industrial buildings and properties in general. The rapidity is allowed by using aggregate data: maps of land-use, buildings and population census. The consistency is required to ensure comparability between evaluations. This can be obtained if, for the series of scenarios to be compared, the user adopts the same type of data to the same scale. For this purpose the tool is flexible and can process data of different types depending of those that are actually available.
In conclusion the tool can be used to prioritize corrective actions to achieve an informed risk reduction or for the identification of the "optimal" measures of risk mitigation.

Acknowledgements

This work has been financed by the Research Fund for the Italian Electrical System under the Contract Agreement between RSE S.p.A. and the Ministry of Economic Development - General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency in compliance with the Decree of March 8, 2006.

References

✔ Albano, R., Sole, A., Adamowski, J., and Mancusi, L. (2014). A GIS-based model to estimate flood consequences and the degree of accessibility and operability of strategic emergency response structures in urban areas, Nat.

Hazards Earth Syst. Sci., 14, 2847-2865, doi:10.5194/nhess-14-2847-2014, 2014.

- ✔ DIRECTIVE 2007/60/CE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2007 *on the assessment and management of flood risks*
- ✔ IPCC: Summary for Policymakers, in: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, in: A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change, edited by: Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach K. J., Plattner, G.-K., Allen, S. K.,Tignor, M., and Midgley, P. M., Cambridge, UK, and New York, NY, USA: Cambridge University Press, 3–21, 2012
- ✔ Gouldby, B. and Samuels, P. (2005). *Language of Risk - Project definitions*, FLOODSite Project report, T32-04-01, March 2005
- ✔ Merz, B., Kreibich, H., Schwarze, R., and Thieken, A. (2010). *Review article* "*Assessment of economic flood damage*", Nat. Hazards Earth Syst. Sci., 10, 1697-1724, doi:10.5194/nhess-10-1697-2010, 2010.
- ✔ SUFRI *Methodology for pluvial and river flooding risk assessment in urban areas to inform decision-making.* 2nd ERA-NET CRUE Research Funding Initiative Flood resilient communities – managing the consequences of flooding. Report, July 2011
- ✔ US Departement of Homeland Security (DHS). *Dams Sector. Estimating Loss of Life for Dam Failure Scenarios*, September 2011
- ✔ Mancusi, L., Sole, A., Cantisani, A., Giosa, L., Albano, R. *Risk assessment for hypothetical dam break - A method for the rapid and consistent evaluation.* 12th ICOLD International Benchmark Workshop on Numerical Analysis of Dams, Graz (Austria), 2-3 October 2013
- ✔ Zenz, G., Goldgruber, M. *Proceedings of the ICOLD - 12th International Benchmark Workshop on Numerical Analysis of Dams* , Graz (Austria), 2-3 October 2013

Looking through the changes: an analysis of the buried watercourses of Como

Maria Antonia Brovelli, Candan Eylül Kilsedar, Marco Minghini, Daniele Oxoli

Politecnico di Milano, Como Campus, DICA, Via Valleggio 11, 22100 Como, Italy

Abstract

Studying territorial evolutions and investigating their underlying processes is essential to ensure continuity in well-done land management decisions. The case of Como City can be considered as a perfect small-scale example of how human influence acted on natural environment. Several watercourses hidden under the road network of the city represent one of the meaningful consequences. FOSS4G and geospatial data from different epochs of Como City historical development allowed to trace the evolution of the territorial setting and the original position of the watercourses. We quantified the variations in their peak flood discharges as a consequence of watersheds urbanization. A Web viewer was created for an easy access to the outcomes of the study.

Keywords

Buried rivers, FOSS4G, historical maps, watershed urbanization, water.

1 Introduction

The knowledge of the territory is a necessary requirement for any person who wants to operate within it, e.g. an urban planner or simply a dweller interested in his/her life-space. This perception cannot subsist without knowing the physical and environmental features as well as their historical and cultural value (Turri, 2002). The built environment, which constantly interferes with the natural one, can produce countless benefits for resident communities but also unpleasant consequences.

As a relevant example, we present here the temporal and spatial study of the watercourses – partially buried – of Como City. This city is placed in a valley which looks onto the southern part of the occidental branch of Lake Como. The origins of Como City date back up to the old Roman times. As a result, the valley has witnessed many transformations that have profoundly changed its natural landscape, especially regarding its watercourses. The Romans conquered the region in 196 BC. The alluvial plain they decided to settle in was a wetland crossed by many streams which probably occupied the central axis of the valley (Gianoncelli, 1975). For this reason, Romans were forced to divert the riverbeds out of their original positions and these works continued all along the Middle Ages and the Modern Age. With the beginning of the 20th century, a growing need for space for suburbs expansion led to the burial of both the major and minor watercourses within the city centre.

Historical large-scale mapping of Como City allows a detailed investigation of the territorial changes occurred. A huge amount of historical maps is actually preserved in the State Archive of Como (http://www.ascomo.beniculturali.it). The oldest map dates back to the 17th century, while many others belong to different cadastral series from 18th and 19th centuries. Maps from the 20th century, mainly related to Master Plans of the city, are instead available in the RAPu Archive (http://www.rapu.it).

The present work can be considered as a proposal of best practice to be applied to the study of the territory. To guarantee the sustainability, the reproducibility as well as the potential customisation to other contexts, all the analyses were made by taking advantage of the maturity and flexibility of Free and Open Source Software for Geospatial (FOSS4G).

2 GIS tools and data processing

The main purpose of the study was to quantify the human influences on Como watercourses and the related watersheds. To enable GIS capabilities on the aforementioned cartographic dataset, the original paper maps (retrieved as digitally scanned images from the archives) were georeferenced through the QGIS Georeferencer GDAL plugin (see Figure 1A). To show that the original position of the main watercourse (Cosia Torrent) differs from the actual one (Gianoncelli, 1975), flow-accumulation raster from 10m cell-size TINITALY/01 DTM (Tarquini et al., 2012) was created using SAGA GIS (http://www.saga-gis.org). This allowed to extract the original flow-path of the torrent deriving from the pure terrain slope. A qualitative comparison between the original and the actual flow-path, the latter retrieved from Lombardy Region geoportal (http://www.cartografia.regione.lombardia.it/geoportale), visually highlights the large diversion work performed by the Romans (see Figure 1B).

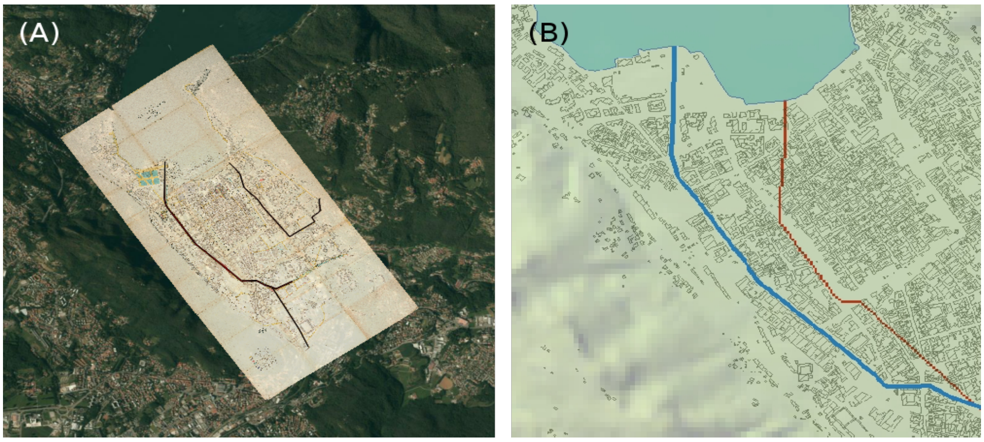


Figure 1: A) A georeferenced historical map of Como City; B) Actual Cosia flow-path (blue line) and DTM-extracted flow-path (red line).

Morphological data of both the watersheds and watercourses of interest were also extracted using SAGA GIS geoprocessing tools. A fast qualitative field-survey was then performed using the Geopaparazzi Android mobile app (http://geopaparazzi.github.io/geopaparazzi). Georeferenced pictures and measurements of several cross-sections along the watercourses (including the channel geometry and the hydraulic roughness) were collected during a number of field campaigns, thus providing a valuable set of input data for the hydraulic modelling (see Section 3).

3 Effects on surface runoff

One of the well-known issues related to urbanization is the widening of impervious areas, which results in an increased surface runoff affecting the watersheds. As a consequence, after intense rainfalls higher peak flood discharges into streams are expected to occur in shorter times.
To quantify this phenomenon on the watercourses under study, Curve Number (CN) maps (USDA, 2004) were created starting from land use D.U.S.A.F. data from different periods (1955, 1999, 2012) and the pedological map of the area. The analysis of CN maps allowed to detect the increase in imperviousness occurred in the area (see Figure 2). This can be easily related to the birth of industrial facilities, new residential areas, etc. which has historically happened within the watersheds under investigation. Moreover, an average weighted CN can be assigned to any sub-area of interest and expected peak runoff can be readily computed involving the Runoff CN method. Project storms with a return period of 2, 20 and 200 years were designed using rainfall records from A.R.P.A. Lombardia (http://arpalombardia.it). Starting from these storms, we computed the expected peak flood discharges for the three land use scenarios at the outlet of the main watercourses (Cosia, Aperto and Valduce). Using the same storms, the differences between the results can be attributed exclusively to the changes in land use brought by the urbanization process (see Figure 2).

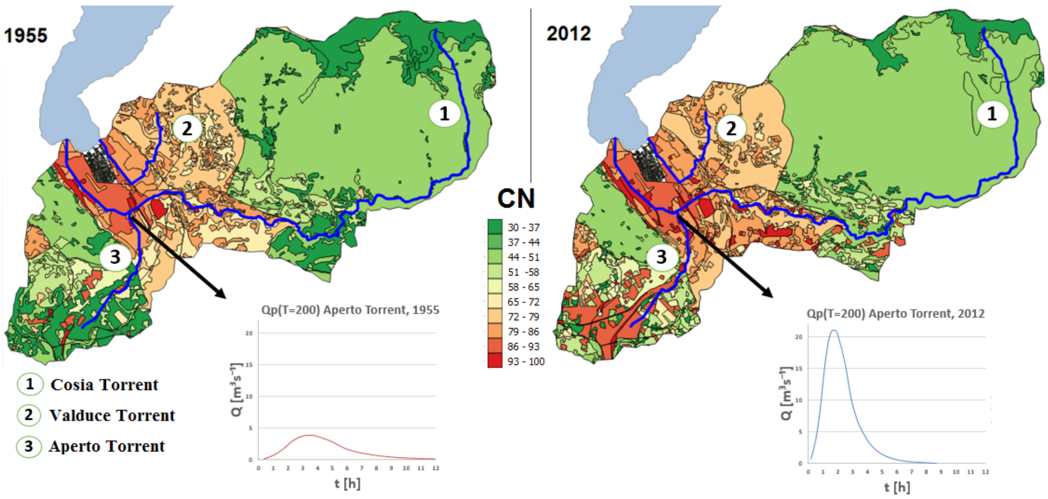


Figure 2: CN maps of Como watersheds in 1955 and 2012 and the change occurred in Aperto Torrent peak flood discharge (200 year return period).

The watercourses were buried more than 60 years ago. Therefore the outcomes of the hydrological modelling can be used to perform updated hydraulic simulations. This is advisable to check the efficiency of the buried channels in conveying the peak flood discharges, whose increase has been clearly proved.

4 Web viewer

To spread the outcomes of this study, a Web viewer based on the Leaflet JavaScript library (http://leafletjs.com) was created to let users access, navigate and compare all the geospatial data (see Figure 3). The data were published as standard WMS layers using GeoServer (http://geoserver.org). To provide users with not only a spatial but also a temporal perception, the inclusion of a time slidebar is planned which will allow to filter the visualization of layers according to the historical period they refer to.

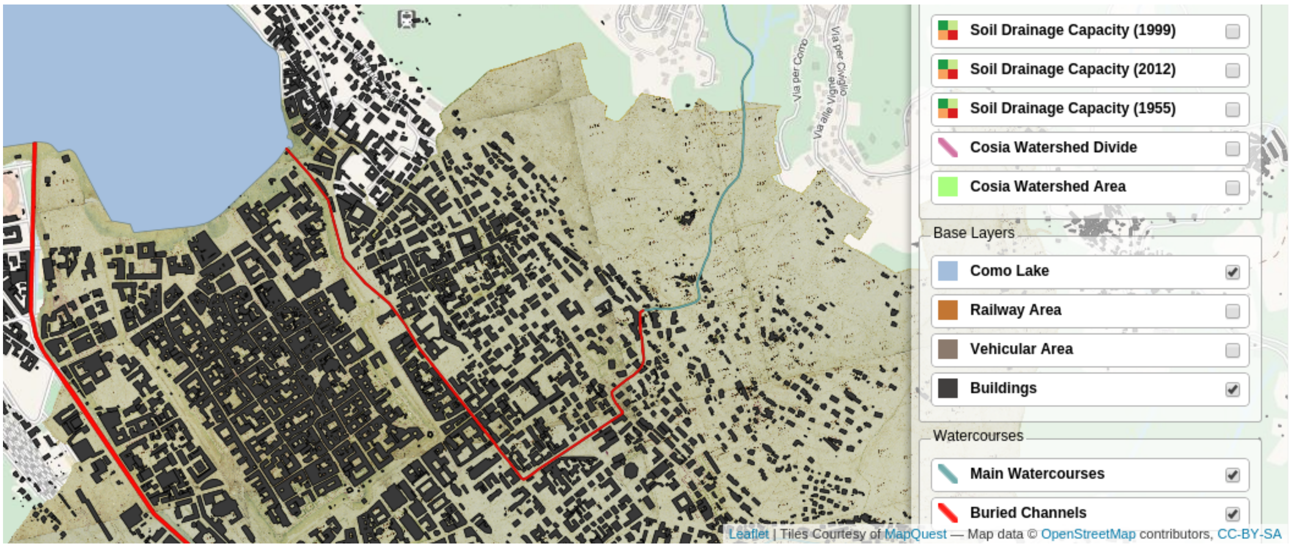


Figure 3: Web viewer showing the geospatial data used in the study.

5 Results and conclusions

The available FOSS4G technologies appeared to be adequate for this analysis. The possibility to retrieve both visual and numerical results makes them suitable for a broad target of users. The achieved outcomes can be considered as a starting point for more rigorous investigations. Besides its simplicity, the presented peak flood discharges estimation proved to be powerful enough to point out the magnitude of changes. Actually, hydraulic simulations based on field-data collected with Geopaparazzi were performed with HEC-RAS software. This choice was due to its simplicity and completeness. Implementation of FOSS4G solutions with equal capabilities is strongly advisable to boost their application in river analysis. Finally, the use of an interactive Web viewer may allow to spread in a very intuitive way the historical memory of changes as well as the awareness about their significant effects on the environment.

References

- ✔ Gianoncelli, M. (1975). *Como e la Sua Convalle: Indagine Storica sull'Origine ed Evoluzione Urbanistica Dei Borghi e Dei Corpi Santi Di Como* (p. 128). Como, IT: New Press.
- ✔ Tarquini, S., Vinci, S., Favalli, M., Doumaz, F., Fornaciai, A., & Nannipieri, L. (2012). Release of a 10-m-resolution DEM for the Italian territory: Comparison with global-coverage DEMs and anaglyph-mode exploration via the web. *Computers & Geosciences*, 38(1), 168–170.
- ✔ Turri, E. (2002). *La conoscenza del territorio. Metodologia per un'analisi storico-geografica* (p. 187). Venezia, IT: Marsilio Editori.
- ✔ U.S. Department of Agriculture. (2004). Part 630: Hydrology National Engineering Handbook. *In National Engineering Handbook.*

The r.inund.fluv tool for flood-prone areas evaluation in GRASS GIS: application to the terminal reach of Magra River

Ilaria Ferrando[1], Bianca Federici[1], Domenico Sguerso[1] & Roberto Marzocchi[2]

[1] DICCA - Department of Civil, Chemical and Environmental Engineering – University of Genoa, Via Montallegro 1, 16145 Genova, Italy

[2] Gter srl Innovazione in Geomatica, GNSS e GIS

Abstract

The present work aims to illustrate the potentiality of the GRASS command *r.inund.fluv*, developed in the Laboratory of Geomatics of DICCA, in the evaluation of potentially flooded areas by means of its application to the terminal reach of Magra River in Italy. Such application underlined the need of a code improvement, in order to better manage the meander-shaped nature of the river through a rectification of river axis. The final results were compared with the official expected flooding map, achieving a satisfying correspondence, that proves the efficiency of the procedure.

Keywords

GRASS GIS, river flooding, flood-prone areas simulation, meanders, Magra River.

1 Introduction to flood-prone areas evaluation with the r.inund.fluv tool

The first step of a risk assessment analysis is the evaluation of flood-prone areas. Its importance is considered for both managing and planning emergency activities. Nowadays, the use of GIS technology for risk assessment analysis is recommended (Peggion et al., 2008). However, it is not widely used for defining inundated areas.

The proposed work uses a GIS module called *r.inund.fluv*, developed in the Laboratory of Geomatics of DICCA (Federici and Sguerso, 2007; Marzocchi et al., 2009), in order to compute perifluvial flood maps. *r.inund.fluv* runs under the version 6.x of GRASS GIS software and it is distributed as an add-on under GNU General Public License terms (*http://grass.osgeo.org/wiki/GRASS_AddOns*).

The tool starts from one-dimensional hydraulic models, and then it takes into account the two-dimensionality of both terrain and flooding event.

The water surface profile along the river axis has to be calculated for a given water discharge through a generic one-dimensional hydraulic model (HEC-RAS, Basement, MIKE 11, etc.). The conformation of the river floodplain has to be described by a high-resolution Digital Terrain Model (DTM).

The coupled integration between the 1D hydrodynamic model and the GIS environment takes the advantage of the simple and well developed 1D hydrodynamic model, able to simulate the river flow in the channel, and of the GIS software to map the flooding extent.

The *r.inund.fluv* tool was already successfully applied to the Tanaro River (Federici and Sguerso, 2007), an Italian river approximately 120 km long, and to Roggia Scairolo (Pozzoni et al., 2009), a stream in the Ticino canton, for the evaluation of flood hazard.

The present study illustrates the application of such tool to the terminal reach of Magra River (Italy), a fluvial weakly meandering river, and a new improvement of the software code to better manage the meanders.

2 Overview to the *r.inund.fluv* procedure

The procedure consists in 5 sequential steps that, starting from simple hypothesis and adding more strict conditions, allow to obtain the map of potentially flooded areas. The steps of procedure can be summarized as follows:

1. <u>First evaluation of potentially flooded areas</u>. The Thiessen polygons technique is used to extend the value of water surface elevation in the river axis to the whole analyzed region. In this way, the value of the water surface elevation in the nearest river point is assigned to every pixel of the DTM. Then this value is compared with the DTM height, so to realize a new map in which all the pixels that have a DTM height greater than water surface elevation are considered not potentially flooded, and conversely for the pixels in which the DTM height is smaller than water surface elevation.
2. <u>Removal of “lakes”</u>. The “lakes”, i.e. the areas considered flooded in the previous step but unreachable by the river flow because surrounded by not floodable terrain (DTM's height greater than water surface elevation in the nearest point of the river axis), have to be “dried”.
3. <u>Hypothesis of orthogonal path from the river axis to the floodplain</u>. The procedure checks which areas between the flooded ones at the end of the step 2 are surely reachable from the water by orthogonal path and which ones are protected, for example by levees, hence have to be “dried”.
4. <u>Search for not orthogonal to river axis paths</u>. The hypothesis introduced is that the water can flow along the maximum slope terrain direction, so it can reach areas previously considered not flooded. From a computational point of view, this step must follow the steps 2 and 3 to restrict the area that can be flooded by alternative paths.
5. <u>Final map</u> is obtained by the union of the maps computed in steps 3 and 4.

The tool has innovative characteristics. In fact, even if it remains substantially one-dimensional, it takes into account the two-dimensionality of floodplain and flooding phenomena, introducing hypotheses that let to correct many typical errors of one-dimensional usually employed procedures. Hence, with respect to the use of one-dimensional model without GIS support, this procedure allows to obtain a more realistic perifluvial flooding map. With respect to a two-dimensional model, it needs a lower computational effort, that allows to apply it to very long river reaches (of the order of 100 km). However, the tool is applicable with reliable results for slowly rising fluvial floods, while it may lead to considerable errors in case of highly dynamic flooding phenomena (mountainous streams, dam or levee breaching, etc.) where a 2D approach is more appropriate (Marzocchi et al., 2014).

3 Application to the terminal reach of Magra River

The Magra River flows at the border between Liguria and Toscana Region, in Italy (figure 1). The whole basin covers about 1700 km^2. The main tributary is Vara River, at 15.7 km from the Magra estuary.

The Magra River estuary is taken as case study because of its remarkable critical situation, both historically and recently proven, as 2009 and 2011 flood events demonstrated. In the present application, the final 6 km of the Magra River, which are the most critical part of the river mainly due to dense population and extensive land use, were analyzed.

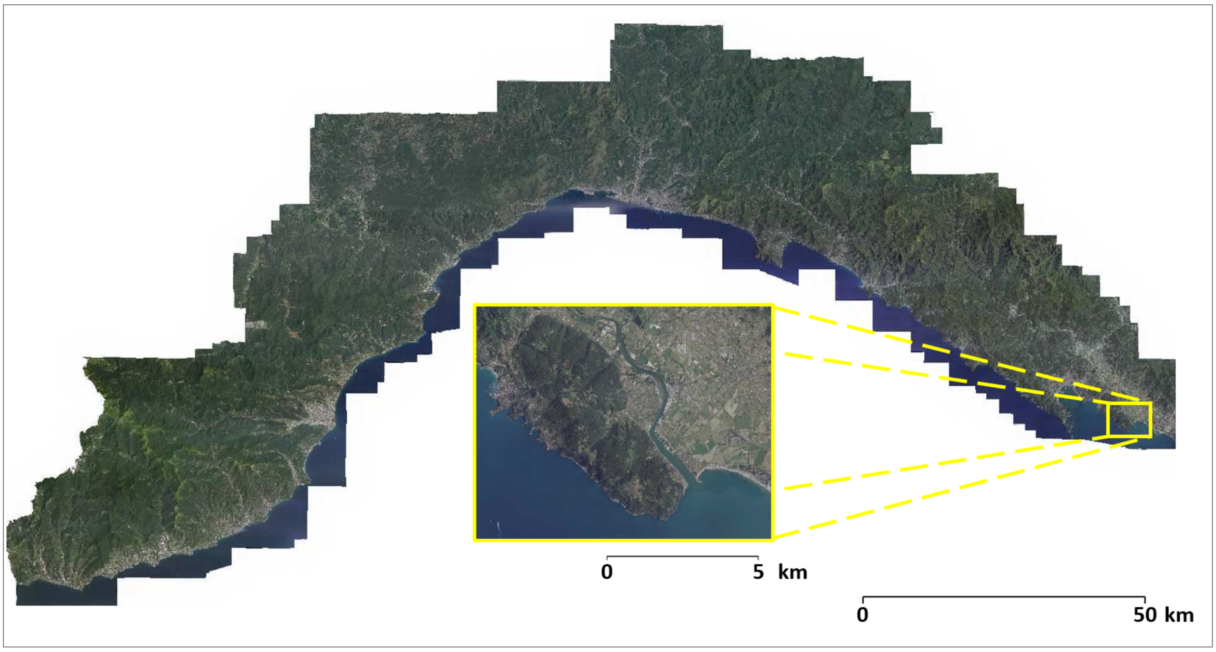


Figure 1: The localization of the Magra River terminal area (2013 Liguria Region ortophoto).

3.1 Input data preparation

The input data used for the application are:

- 1 m resolution DTM, supplied by *Autorità di Bacino Interregionale del Fiume Magra*. It comes from a LIDAR survey carried out for *Ministero dell'Ambiente,* in WGS84 reference system and geographical coordinates;
- 55 cross sections, supplied by *Autorità di Bacino Interregionale del Fiume Magra,* in WGS84-UTM32 cartographic coordinates;
- the official expected flooding areas, supplied by *Autorità di Bacino Interregionale del fiume Magra* in WGS84-UTM32 cartographic coordinates, with respect to the 200 years return period flow rate.

The hydraulic profile was computed in HEC-RAS software knowing the flow rate and the river cross sections, assuming the flow as one-dimensional and stationary. A xyz file, describing the cartographic position of every cross section and the associated water surface elevation, was imported in GRASS. Such file was created by the script hec2grass.sh developed by Roberto Marzocchi (www.gter.it/downolad), combining the cartographic coordinates of the first upstream cross section, the geo-referenced river axis and the HEC-RAS profile file, in which, for every cross-section, the water surface elevation is associated to the progressive distance from the first upstream one (figure 2).

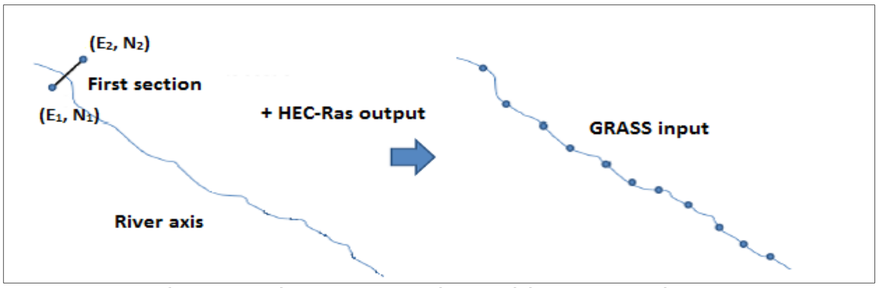


Figure 2: hec2grass.sh working procedure.

3.2 The first results

The procedure was applied to the 200 years return time flow rate, corresponding to a 6400 m^3/s water discharge. The results of the 5 sequential steps of the procedure, described in par. 2, are illustrated in figure 3. They showed a fault in the procedure dealing with curves in the river axis, in the 5th step map. In fact, there are important and unrealistic differences in water surface elevation between neighboring points downstream the meanders (figure 4). Hence a modification of the code was performed.

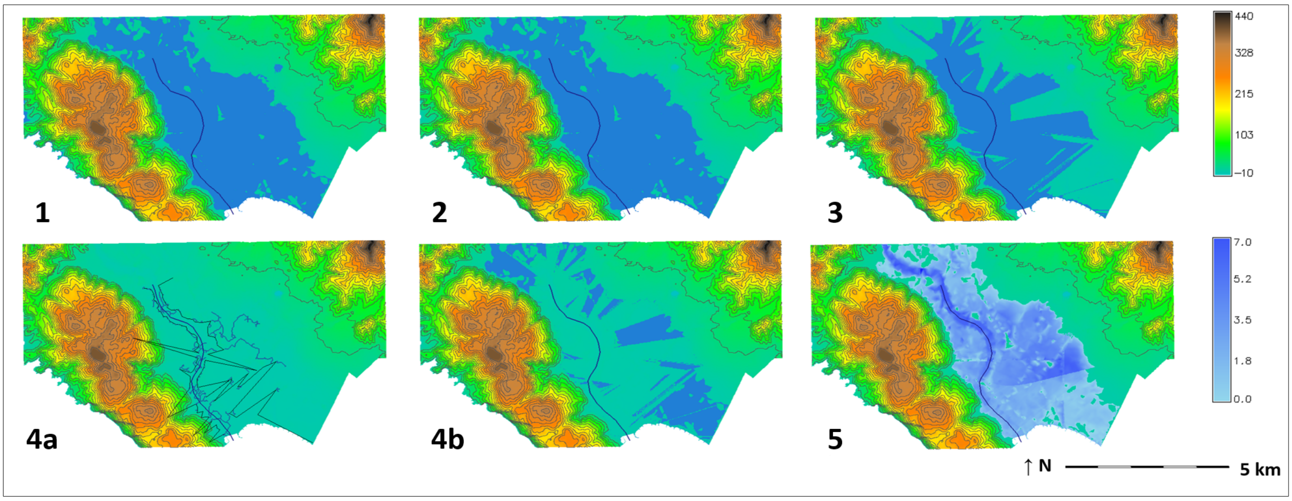


Figure 3: The results of the 5 steps of the r.inund.fluv procedure applied to the terminal reach of Magra River for the 200 years return time discharge. The Figures 3.4a and 3.4b show the identified non-orthogonal to river axis paths and the resulting flooded areas respectively.

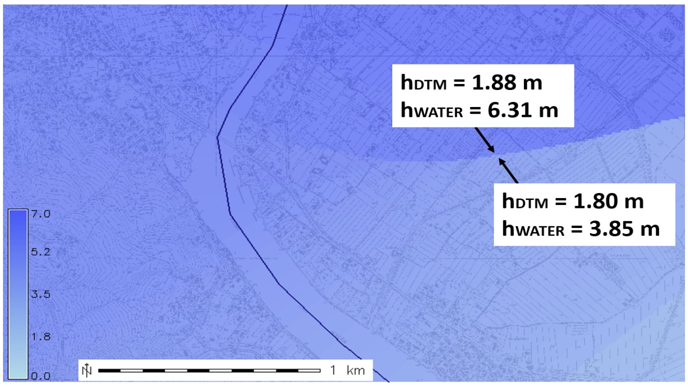


Figure 4: A zoom of Figure 3.5 that highlights an example of unrealistic discontinuity of water surface elevation downstream the meander.

3.3 *r.inund.fluv* code modification

The previously highlighted unrealistic discontinuity in the final map (Figure 4) is due to the Thiessen polygons approach of the first step of the procedure.

The presence of a curve produces a singularity in the center of the curvature (point P in Figure 5a), so that many different values of water surface elevation (H1, H2, …) can be assigned to point P, through a central projection.

The problem was simply solved by using a Matlab code to achieve a river axis rectification, connecting the first and the last point of the reach with a line and orthogonally projecting on it the water surface elevations computed by HEC-RAS along the river axis (as showed in Figure 5b).

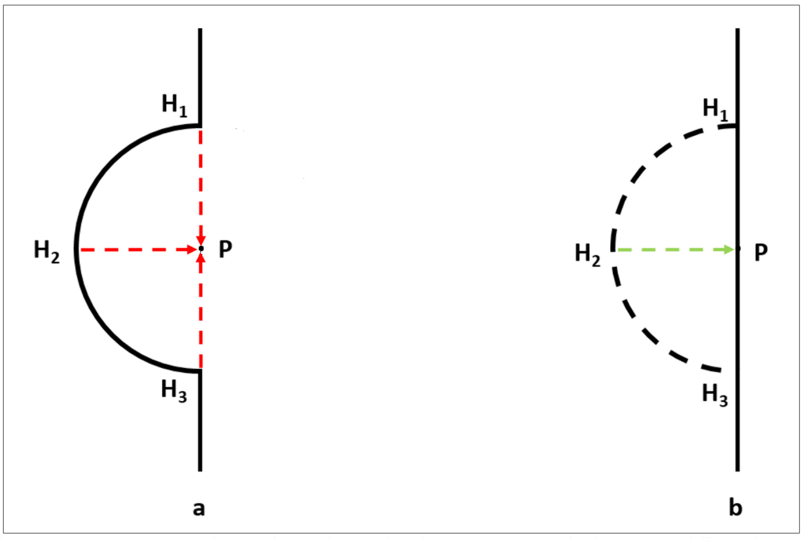


Figure 5: A scheme representing the singularity (5a) and the rectification principle (5b).

The starting and final points for axis rectification have to be chosen by the user, depending on the analyzed river configuration. In the present case, the first and the last point of the river reach were used. Moreover, note that, in case of more marked meandering rivers, the procedure might need a meander rectification through a polyline instead of a simple line.

After the modification, the results of the first step of the procedure is a sequence of strips orthogonal to the rectified axis, to each of which corresponds the water surface elevation of the closest point of the rectified axis. The rectified profile was used only in the first phase of procedure, while the following ones refer to the real river axis. The new final flooding map is illustrated in Figure 6, where the elimination of discontinuities downstream the meanders can be appreciated.

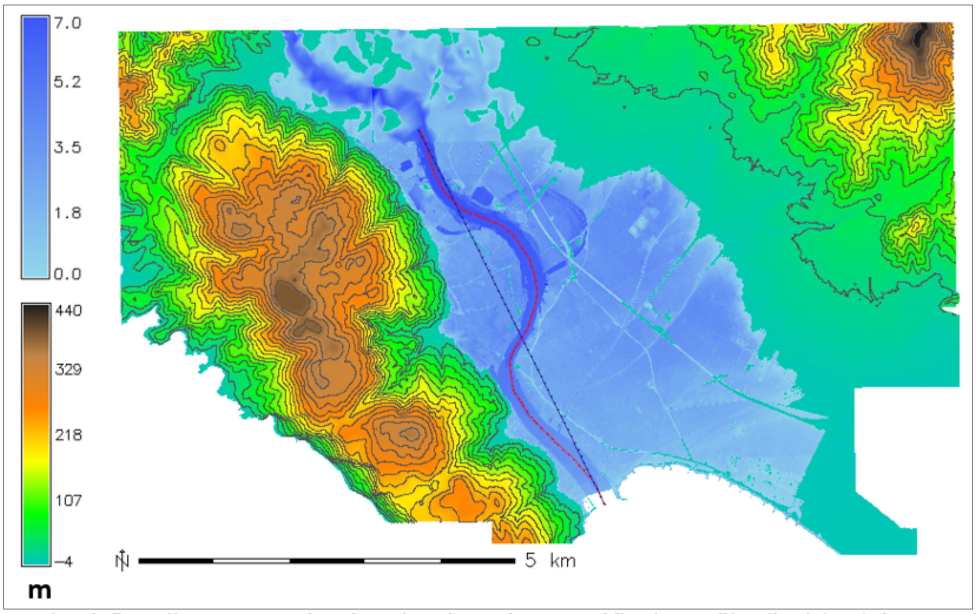


Figure 6: Final flooding map obtained using the rectified profile (in blue) instead of the original one (in red) in the first phase of the r.inund.fluv procedure.

3.4 Comparison with the official expected flooding map

A comparison between the calculated final map and the official expected flooding map for the 200 years return time, supplied by *Autorità di Bacino Interregionale del fiume Magra,* was performed.

The qualitative agreement is illustrated in Figure 7. The blue areas are the perfect correspondence areas, the red ones are areas flooded in the official map and not considered flooded by *r.inund.fluv* procedure, conversely the yellow ones. It is possible to notice a satisfying correspondence from a qualitative point of view, that confirms the good quality of the procedure.

From a quantitative point of view, a performance index (Bates et al., 2000; Di Baldassarre et al., 2006) was computed as follows:

$$PI = \frac{Area_F \cap Area_I}{Area_F \cup Area_I}$$

where $Area_F$ is the flooded area in the official map and $Area_I$ is the flooded area computed with *r.inund.fluv.*
The value assumed by the performance index is 72%, which means that the two maps have a correspondence for the 72% of their extent (the blue area in Figure 7).

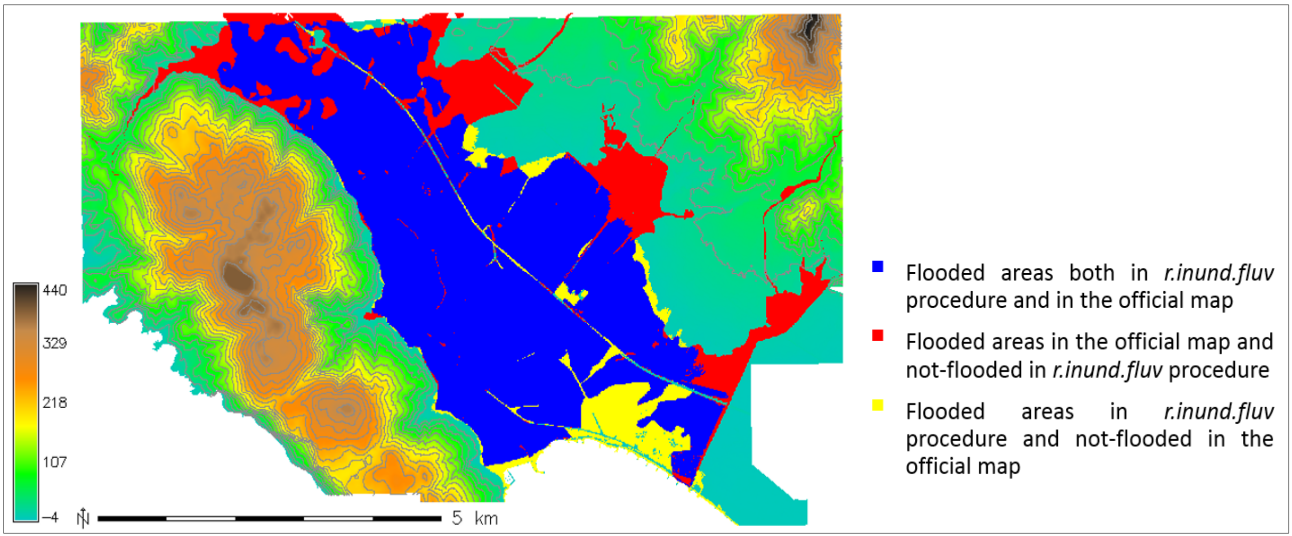


Figure 7: Comparison between the calculated final map and the official expected flooding one.

The main differences are due to the fact that the *r.inund.fluv* procedure does not consider the tributaries, which play a fundamental role in the official expected flooding map, as Figure 7 shows.

4 Conclusions

The presented work consists in the realization of potentially flooded areas maps by means of an automatic GIS procedure using the *r.inund.fluv* tool, starting from a water surface elevation profile, computed through HEC-RAS along the river axis, and a high resolution DTM.

Despite the procedure is based on a mono-dimensional hydraulic code to obtain the water surface profile, it can reproduce the bi-dimensionality of the terrain and of the flooding phenomena, keeping the simplicity and the computational speed of a 1D model. The whole procedure is automatic, but it is subdivided in 5 consecutive phases that, starting from very simple hypothesis, allow to refine the results and obtain the map of potentially flooded areas.

The application of the *r.inund.fluv* tool on the terminal reach of Magra River led to a modification of the code to take into account the meanders of the river, introducing a rectified profile to be used only in the first step of the procedure.

The resulting map for the 200 years return period flow rate was compared with the official expected flooding map, showing a good correspondence, both from qualitative and quantitative point of view.

A deeper application of *r.inund.fluv* to the terminal reach of Magra River, analyzing different water discharges and constructive solutions for the reduction of flood-prone areas, will be object of a future publication. Moreover, an improvement of the code to take into account the contribute of tributaries in flooding event will be studied.

Acknowledgment

We wish to thank the *Autorità di Bacino Interregionale del fiume Magra* for the cartographic material provided.

References

- ✔ Bates, P.D., & De Roo, A.P.J. (2000). A simple raster-based model for floodplain inundation. *Journal of Hydrology*, 236, 54-77.
- ✔ Di Baldassarre, G., Castellarin, A., Brath, A., Horritt, & M., Bates, P. (2006). Modellistica idraulica monodimensionale: alcune considerazioni applicative sul grado di dettaglio ottimale della descrizione topografica. *XXX Convegno di Idraulica e Costruzioni Idrauliche*.
- ✔ Federici, B., & Sguerso, D. (2007). Procedura automatica per la creazione di mappe di potenziale inondazione fluviale. *Bollettino SIFET,* 4, 25–42.
- ✔ GRASS Development Team (2013). Geographic resources analysis support system (GRASS) software, version 6.4.3. http://grass.osgeo.org
- ✔ Marzocchi, R., Federici, B., & Sguerso, D. (2009). Procedura automatica per la creazione di mappe di potenziale inondazione fluviale in GRASS: il modulo r.inund.fluv. *Atti del IX Meeting degli Utenti Italiani di GRASS-GFOSS* 21-22 febbraio 2008, 161–178.
- ✔ Marzocchi, R., Federici, B., Cannata, M., Cosso, T., & Syriou, A. (2014). Comparison of one-dimensional and two-dimensional GRASS GIS models for flood mapping, *Applied Geomatics*, 6(4), 245-254. doi: 10.1007/s12518-014-0140-1, ISSN: 1866-9298, Publisher: Springer Berlin Heidelberg.
- ✔ Peggion, M., Bernardini, A., & Masera, M. (2008). Geographic information systems and risk assessment. *Scientific and Technical Research* series EUR 23058-EN.
- ✔ Pozzoni, M., Marzocchi, R., & Graf, A. (2009). Roggia Scairolo – Zonazione della pericolosità per alluvionamento. *Technical report of Institute of Earth Sciences.*

A RUSLE model application with GRASS GIS: an evaluation study in the Rio Centonara catchment

Elena Mezzini[1], Francesca Ventura[1], Livia Vittori Antisari[1], Federico Magnani[1]

[1] DipSA – Dipartimento di Scienze Agrarie- Università di Bologna: Viale Fanin, 44 Bologna, elena.mezzini2@unibo.it, livia.vittori@unibo.it, francesca.ventura@unibo.it, federico.magnani@unibo.it

Abstract

The RUSLE (Revised Universal Soil Loss Equation) (Renard, Foster, Weesies, & Porter, 1991) is an empirical model estimating the annual loss of soil due to surface erosion processes (sheet, interril and rill). This work calculates the annual amount of erosion in the little hilly catchment of Rio Centonara, in Bologna Province, Italy, based on catchment properties and measured precipitation rates, and compares resulting estimates with long-term field measurements. Both field measurements and GIS-based routine for the period 2006-2009 give a mean soil erosion rate of 35 t ha^{-1} y^{-1}. Although the study has been originally developed as a practical exercise for teaching purposes, considering the good estimate of the erosion rate given by GRASS GIS 7.0, it could be considered as a good example of a GRASS GIS routine applied in environmental and hydrological studies.

Keywords

RUSLE, GRASS GIS 7.0, soil erosion, Centonara catchment.

1 Introduction

Recently, the RUSLE (Revised Universal Soil Loss Equation) (Renard et al., 1991) empirical model has been increasingly applied to estimate the annual loss of soil due to surface erosion processes (sheet, interril and rill). This model is formulated as a simple equation:

$$A = R * K * LS * C * P \quad [t\ ha^{-1}\ y^{-1}]$$

where A is the estimated average soil loss in t ha^{-1} y^{-1}, R is the rainfall-runoff erosivity factor, K is the soil erodibility factor, L is the slope length factor, S is the slope steepness factor, C is the cover-management factor and P is the support practice factor.

In this work this model has been applied to the small catchment of the Centonara stream, in the hills above Ozzano (Bologna, Italy). The original aim of this study was to teach students how to apply the RUSLE model with a practical exercise. The spatial characteristic of several factor of the RUSLE equation suggest to apply this equation at a landscape scale, by means of a suitable GIS software. We chose to use the FOSS4G software GRASS GIS 7.0, that in its new version implements very useful hydrological commands.

2 Materials and methods

2.1 Study area

The study was carried out in the small semi-agricultural Centonara catchment, in the Northern Italian Apennines, about 20 km from Bologna. It covers an area of 273 ha, with an altitude ranging between 84 and 350 m a.s.l. The watershed is characterized by a Mediterranean climate (Köppen-Geiger classification) (Kottek, Grieser, Beck, Rudolf, & Rubel, 2006), where the maximum and minimum temperatures occur in July-August and December- January respectively, while rainfall is concentrated in spring and autumn. The mean air temperature is 15°C and the total annual rainfall is 749 mm y^{-1} (mean of the period 2000-2010).

The hydrological response of the Centonara stream is highly seasonal, with discontinuous discharge values over the year. The watershed is heterogeneous in terms of geology (Farabegoli, Rossi Pisa, Costantini, & Gardi, 1994), slope, vegetation and land cover. About 20% of the total area is represented by badlands (gullies, “calanchi”). These areas, characterized by bare soil and slope over 40%, are affected by rapid morphogenesis mainly due to rill and gully erosion, associated to pipe development and shallow and deep mass movements, such as slides, slumps, mudflows and slope collapses.

The rest of the catchment is covered by arable land (25%), spontaneous vegetation (herbaceous, shrubs and trees) and riparian strips (55%).

Slope values over the entire catchment range from less than 10% (more than 13% of area) to more than 50% (almost 14% of area), with the intermediate classes, 20-30% and 30-35%, as predominant. As a consequence of the lithological variability, the soil texture varies with the area: it is loam in the flat area, while in the hilly part and the badlands areas the textural classes are silty-clay-loam and silty-clay.

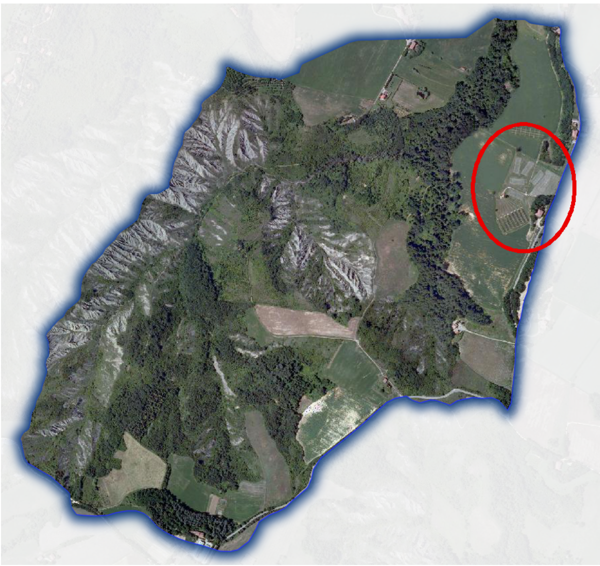


Figure 1: Aerial photography (Servizio Sviluppo dell’Amministrazione digitale e Sistemi Informativi Geografici Regione Emilia-Romagna, 2011) of the catchment area. The red ellipse highlights the measurement area. QGIS 2.6 output.

2.2 Soil loss measurement

Direct monitoring of soil loss was carried out since 1992 by means of an automatic hydrological station (ISCO, 6700), installed to measure the water discharge of the Centonara stream and collect water samples to determine soil loss. Its data logger has a control for collecting both in time and volume: the sampler collects one sample every 24h and one every 3000 m^3 of flow, to avoid losing large events that could occur within a 24h-interval, which are known to transport the majority of sediments. To determine the sediment rate, runoff samples were oven dried at 105 °C for 24h and then weighed. Due to the sampling method, in the case of intense events we had more sediment concentration measures for each event, while for small events the single sediment concentration was assumed to be uniform throughout the runoff period.

The catchment was also equipped with an automatic meteorological station, which records hourly air temperature, relative humidity, wind speed and direction, and solar radiation. Precipitation is recorded every 10 minutes in order to have more details on its characteristics (quantity, intensity, etc.).

Part of these data had been utilized also in other erosion research studies (Pieri et al., 2007, 2013).

2.3 Spatial application of RUSLE

As described in the introduction, the RUSLE model consists of a simple equation. Each factor of the formula could assume different values depending on the position in the catchment. Because of the factors spatial variability, the use of GIS software is strongly suggested. In particular GRASS GIS 7.0 has new hydrological tools, based on raster maps. Thus, the first step consisted in rasterizing all vector layers with the same spatial resolution as the best Digital Elevation Model (DEM) available (10 m). DEM had previously been obtained from the interpolation of contour lines and elevation points from CTR 1:5.000 (Servizio statistica e informazione geografica - Regione Emilia-Romagna, 2013). Subsequently, to estimate soil losses, we had to calculate the spatial distribution of each factor and finally apply the RUSLE equation to estimate the erosion rate in every single pixel of the catchment.

2.3.1 R factor

R is the rainfall-runoff erosivity factor. A large amount of meteorological data are needed to calculate this factor, which are usually difficult to obtain. Moreover very different formulas have been developed to calculate the R factor, but each formula is fitting only to particular climatic conditions. Therefore, since the required continuous meteorological data are available for the study area, and the climate is Mediterranean, we applied for R the equation (Diodato, 2004):

$$R = EI_{30} \qquad [\mathrm{MJ\ mm\ ha^{-1}\ h^{-1}}]$$

where EI_{30} (empirical erosivity annual index) is estimated from the total annual precipitation, the maximum daily precipitation, and the maximum hourly precipitation. Through the application of the WRF (Weather Research and

Forecasting) software, based on WRF model, these calculations are simpler and automatic.
Moreover because of the small dimensions of the study area, for each year all the catchment could be represented by a single value. A substantial interannual variability in the R factor was however apparent, as we applied the RUSLE model from 2006 to 2009.

	2006	2007	2008	2009
Annual Prec (mm)	466.6	554.4	712.2	782.2
R=a*EI30	1446.8	1453.8	1186.4	1436.2

Table 1: Centonara catchment: annual precipitation and R factor values (years 2006-2009).

If the study area had been larger and with more than one meteorological station, GRASS GIS would have been the best choice to estimate the R factor. In fact from version 7.0 it contains specific RUSLE commands, including one specifically designed to estimate R (r.usler), which also gives the option to choose the calculating formula.

2.3.2 K factor

K is the soil erodibility factor; it depends mainly on the content of soil organic matter, soil texture, soil structure and soil permeability, as shown in the widely cited nomograph of Wischmeier & Smith (1978). The same relationship is captured by the equation:

$$K=\frac{2{,}1*10^{-4}(12-OM)M^{1{,}14}+3{,}25(S-2)+(P-3)}{7{,}59*100} \qquad [\text{t ha h ha}^{-1}\text{MJ}^{-1}\text{mm}^{-1}]$$

where OM represents the percentage soil organic matter, M is a structural factor defined as (%silt + %fine sand)*(100-%clay), S is the soil structure and P is the soil permeability factor.
Because of a temporary bug (in GRASS GIS 7.0beta3), quickly but partially solved in the new GRASS GIS 7.0 release, we applied this equation through the r.recode and r.mapcalc commands. However, GRASS GIS 7.0 also includes the command r.uslek, which estimates the K factor only from soil texture values. Nevertheless, soil characteristics are needed to apply both the equation and the r.uslek command. Thanks to previous studies (Pieri et al., 2013), we had a pedological geodatabase of soils in the Centonara catchment, so that we only had to calculate the average percentage soil texture, weighted by the thickness of each soil horizon, with the exception of soil organic matter which was calculated only in the first 25 cm of soil. Soil structure and permeability had been derived from soil taxonomy (Soil Survey Staff, 1975).

2.3.3 LS factor

LS is the slope factor. As demonstrated in previous works (Desmet & Govers, 1996; Moore & Wilson, 1992; Neteler & Mitasova, 2008)3D LS factor can be simplified using the unit contributing area instead of the slope length.

$$LS_r=(m+1)\left(\frac{A_r}{A_0}\right)^m\left(\frac{\sin(S_r)}{S_0}\right)^n$$

In the equation r is a point in the slope, m and n are empirical constants (m=

0,4 and n=1,3 as in Neteler & Mitasova (2008) and Vianello, Lorito, Bigi, & Pavanelli (2004)), A is the unit contributing area (also known as flow accumulation grid), S is the slope in degrees, A_0, S_0 represent USLE standard dimensions (A_0= slope length = 22.13 m; S_0= slope= 0.0896= 8.96 % = 5.15°). To calculate the unit contributing area (Figure 2), GRASS GIS includes the r.flow command, which only needs a Digital Elevation Model (Servizio statistica e informazione geografica - Regione Emilia-Romagna, 2013).
The same raster map is needed to calculate the slope map, by the r.slope.aspect command. Using these two maps as an input for the LS equation in the command r.mapcalc, we were able to derive the LS raster map, although with an error, since the streams were represented with the higher LS values, but they don't contribute to RUSLE calculation. Thus, as a final step, we assigned a NULL value to the highest value pixels (streams, in this study values from 25 to 100).

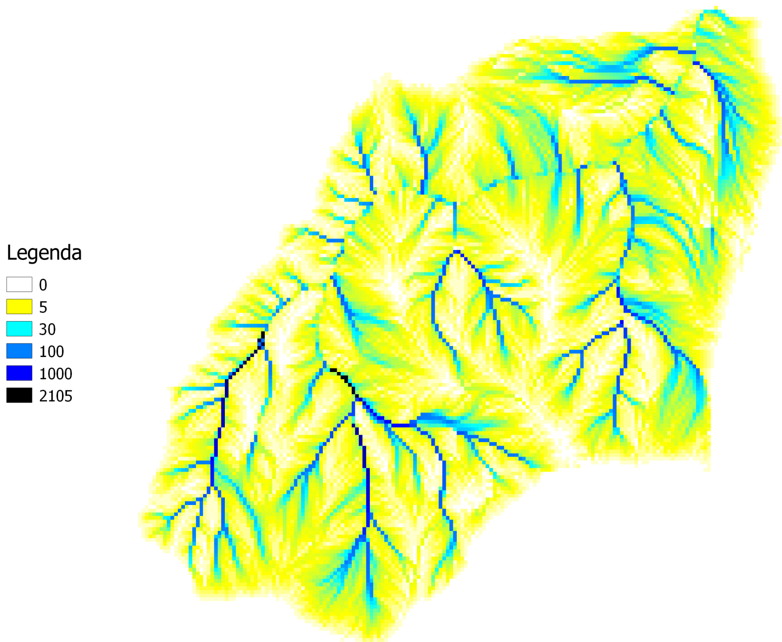


Figure 2: Unit contributing area, also known as flow accumulation grid.

2.3.4 C factor

C is the cover-management factor. It is mainly estimated from the land cover, and the land cover map of the area was obtained through a cross-checking of field observations and regional land cover map.
C is function of previous use of soil, land cover, and roughness of soil surface. This factor is a reduction factor, comprised between 0 and 1. Several studies have been presented in the literature, which assign specific C values to land cover classification types (Confluence Inc., 2009; USDA NRCS, 2000; Vianello et al., 2004; Wischmeier & Smith, 1978; Zampoli, 2007); in the present study, we assigned these values to the different land cover classes through the r.recode GRASS GIS command.

codNLCD_92	Landuse	area (ha)	%	C
21	Low Intensity Residential	1.8	0.94	0.003
31	Bare Rock/Sand/Clay	14.09	7.33	0.36
41	Decidous Forest	47.93	24.92	0.003
42	Evergreen Forest	1.56	0.81	0.003
51	Shrubland	58.33	30.33	0.046
61	Orchards/Vinyards/Other	1.1	0.57	0.45
71	Grasslands/Herbaceous	20.34	10.58	0.02
81	Pasture/Hay	25.99	13.51	0.02
82	Row Crops	0.48	0.25	0.2
83	Small Grains	20.72	10.77	0.24

Table 2: C factor for different NLCD (National Land Cover Dataset, 2001, USGS) classes.

2.3.5 P factor

P is the support practice factor, it is the expression of agricultural practice for soil conservation, to reduce the amount and rate of water runoff. The most relevant practices are basically: cover crops, strip-cropping, contour farming, permanent and annual crops in rotation. As in the case of the C factor, P is a reduction factor, with value between 0 and 1; we utilized r.recode to assign values to P.

codNLCD_92	Landuse	P
21	Low Intensity Residential	1.00
31	Bare Rock/Sand/Clay	1.00
41	Decidous Forest	1.00
42	Evergreen Forest	1.00
51	Shrubland	1.00
61	Orchards/Vinyards/Other	1.00
71	Grasslands/Herbaceous	1.00
81	Pasture/Hay	0.85
82	Row Crops	0.85
83	Small Grains	0.85

Table 3: P factor for different NLCD (National Land Cover Dataset, 2001, USGS) classes.

2.3.6 RUSLE

We applied the RUSLE equation by means of the r.mapcalc GRASS GIS function for each year in the period 2006-2009 by multiplying all the factors described above with an annual time step.

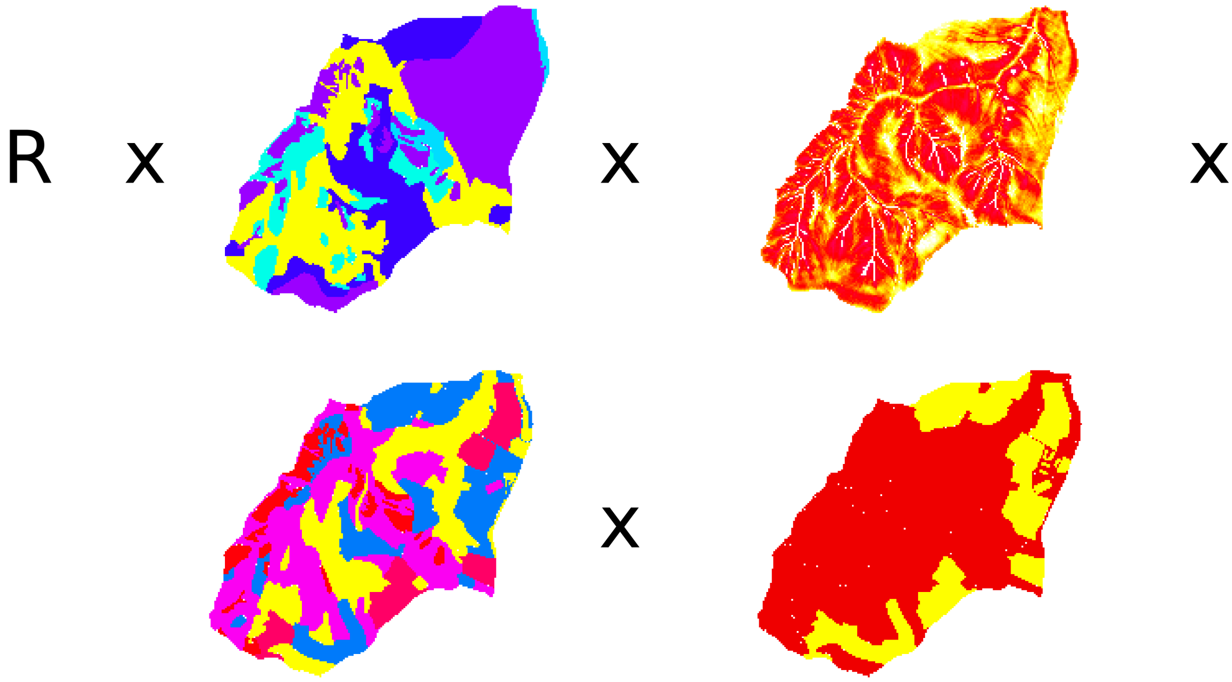


*Figure 3: A = R * K * LS * C * P*

2.3.7 Summary GRASS GIS 7.0 utilized commands

R factor:

- r.usler: to estimate the R factor when there are more than one meteorological stations

K factor:

- r.uslek: to estimate the K factor only from soil texture
- r.recode: to classify soil maps by sand/silt/clay/om content, permeability and structure
- r.mapcalc: to apply the K factor equation

LS factor:

- r.flow: to calculate flow accumulation or unit contributing area
- r.slope.aspect: to calculate slope and aspect from a DEM
- r.mapcalc: to apply the equation

C and P factors:

- r.recode: to classify the landuse map by cover-management and support practice

Statistics:

- r.univar: to calculate general statistics for the whole dataset
- r.report: to calculate general statistics dividing the data in categories previously defined by the user through r.recode

3 Results and Discussion

3.1 K factor

Concerning the K factor, we decided to utilize the mathematical method instead of the GRASS GIS r.uslek command, since even after the bug fix (in GRASS GIS 7.0), in some areas the GIS solution overestimated the K factor by one order of magnitude and didn't calculate it at all in others (Figure 4).

We are confident that the K values calculated by the equation are correct because of some K reference values for Italy in literature (Zampoli, 2007).

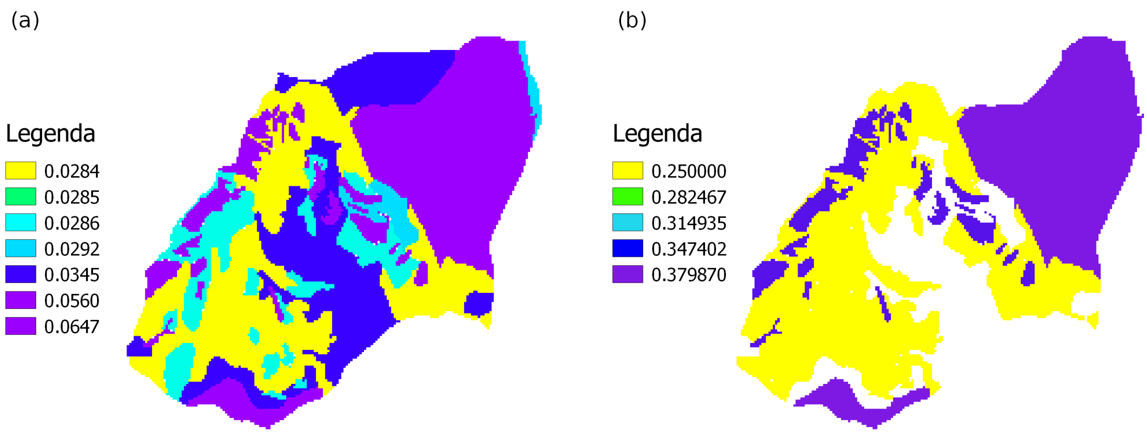


Figure 4: (a) K factor derived from equation and (b) GRASS GIS K factor comparison.

3.2 Other factors

As shown in Figure 5, the LS factor is characterized by a range of values from 0 to 100, since, as explained previously, we assigned NULL values to the streams (25-100).

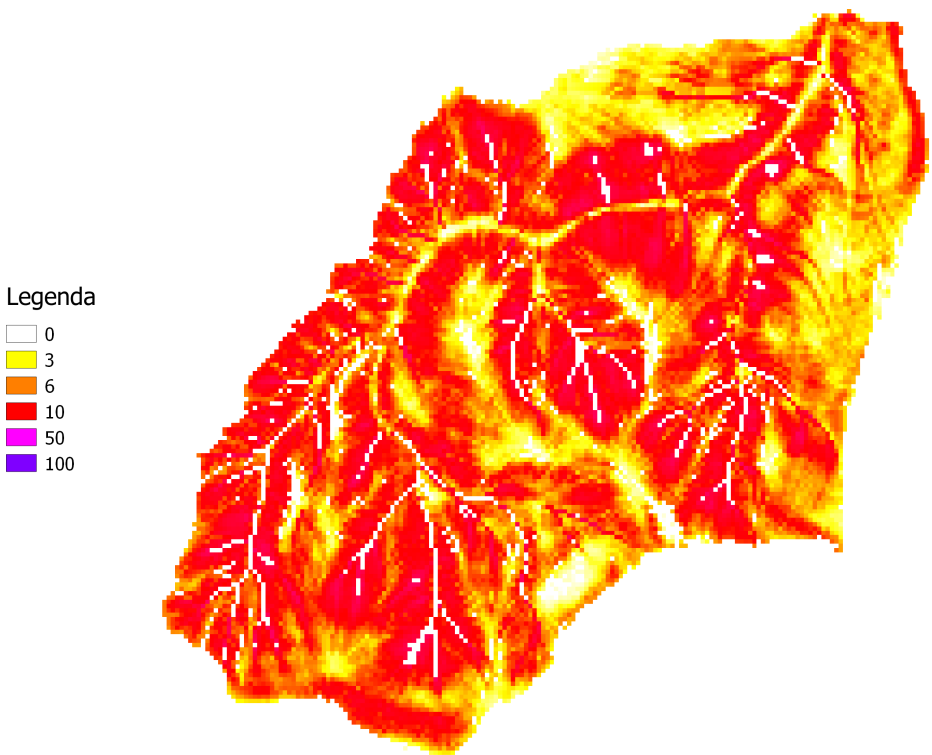


Figure 5: LS factor. Range 0-100 with 25-100 as NULL values.

The C and P factors, instead, are reducing factors because they are both smaller than 1. C factor ranges from 0.003 to 0.45, while P ranges from 0.85 to 1.

3.3 RUSLE results

In Figure 6 the erosion rate is shown for each year. The temporal variations are due exclusively to changes in the R factor because it is the only one changing during the study period; whereas the spatial variation is due to the other factors.
Field observations gave us a reference value of soil loss of 35 t ha^{-1} y^{-1} with variations from 10 to 50 t ha^{-1} y^{-1}. Thanks to r.univar GRASS GIS statistical command, we could obtain summarizing and comparable results.
Although the total annual precipitation was maximum in 2009 (Table 1), the maximum erosion rate occurred in 2007 (Table 4). This is mainly due to the intensity of precipitations. The year 2009 had 107 events of which only 8 were characterized by an intensity higher than 10 mm/h; on the contrary 2007 had only 58 events, 7 of them with intensity higher than 10 mm/h. These results demonstrate that erosion events occur when heavy rainfalls occur.

Year	Mean erosion rate (t/ha y)	Maximum erosion rate (t/ha y)	Erosion (t/y)
2006	36.99	726.72	6879.92
2007	37.17	730.23	6913.21
2008	30.33	595.92	5641.65
2009	36.72	721.39	6829.51
mean	35.31	693.56	6566.07

Table 4: Erosion rates and mean values from RUSLE model.

Several different statistical routines are available in GRASS GIS, of which r.univar is the simplest; we tested our results also with r.report, however, which divided the data in categories previously defined by the user through r.recode.
In this study we divided results in six classes: minimum erosion (0-5 t ha^{-1} y^{-1}), low erosion (5-10 t ha^{-1} y^{-1}), moderate erosion (10-30 t ha^{-1} y^{-1}), high erosion (30-80 t ha^{-1} y^{-1}), very high erosion (80-300 t ha^{-1} y^{-1}), extreme erosion (higher than 300 t ha^{-1} y^{-1}). The r.report routine calculated the area and the percentage for each of these erosion classes.
Once more, table 5 shows year 2007 as the one with highest erosion. In particular it has a lower presence of minimum erosion class pixel, but the highest presence of pixel characterized by high-level classes. Nevertheless in the study period 2006-2009 there has been no great varations of erosion rates.
On a spatial basis, areas with higher erosion rates are those characterized by gullies "calanchi" land cover and pedology, already known in the literature as vulnerable areas.

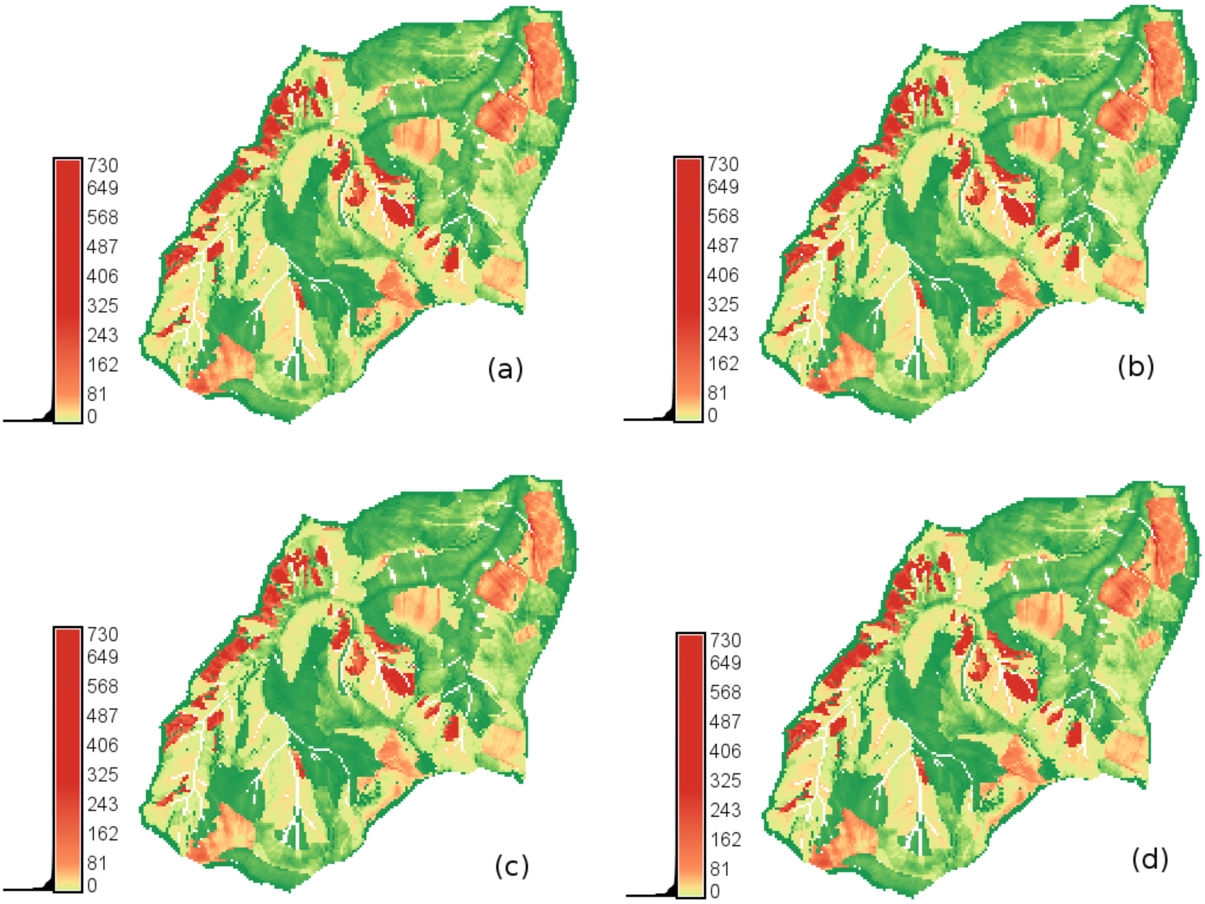


Figure 6: RUSLE Erosion rate (t ha^{-1} y^{-1}) for (a) 2006, (b) 2007, (c) 2008, (d) 2009.

2006			
Erosion class	Range ($t\ ha^{-1}\ y^{1}$)	Area (ha)	Percentage
1 minimum	0-5	78.49	42.20
2 low	5-10	22.28	11.98
3 moderate	10-30	46.82	25.17
4 high	30-80	18.69	10.05
5 very high	80-300	13.33	7.17
6 extreme	>300	6.37	3.43
2007			
Erosion class	Range ($t\ ha^{-1}\ y^{1}$)	Area (ha)	Percentage
1 minimum	0-5	74.59	40.11
2 low	5-10	24.57	13.21
3 moderate	10-30	47.78	25.69
4 high	30-80	19.13	10.29
5 very high	80-300	13.48	7.25
6 extreme	>300	6.43	3.46
2008			
Erosion class	Range ($t\ ha^{-1}\ y^{1}$)	Area (ha)	Percentage
1 minimum	0-5	81.76	43.96
2 low	5-10	24.51	13.18
3 moderate	10-30	45.32	24.37
4 high	30-80	17.39	9.35
5 very high	80-300	13.24	7.12
6 extreme	>300	3.76	2.02
2009			
Erosion class	Range ($t\ ha^{-1}\ y^{1}$)	Area (ha)	Percentage
1 minimum	0-5	75.11	40.39
2 low	5-10	24.43	13.14
3 moderate	10-30	47.73	25.66
4 high	30-80	19.01	10.22
5 very high	80-300	13.39	7.20
6 extreme	>300	6.31	3.39

Table 5: Annual distribution of erosion classes.

4 Conclusions

This study not only has been a good practical exercise for students, but also a validation of the GRASS GIS 7.0 RUSLE routine. Results are comparable with the field measurements, giving a mean erosion rate of about 35 $t\ ha^{-1}\ y^{-1}$. In conclusion GIS technologies can be assumed as practical and reliable tools to estimate a catchment erosion rate through the RUSLE model; moreover, they are easy to apply and don't need too large an amount of input data. They are really flexible, making it possible to analyze both small and extended areas, with one or more meteorological stations. On the other hand, it could be difficult to obtain all the required data (high resolution DEM, soil characteristics, etc.) and to decide what kind of algorithm to apply in selected study areas.

Concerning the K factor, it would be useful to interact with r.uslek developer once more, to better fix it and to improve the available documentation, rather

poor as of now. Nevertheless, GRASS GIS software is certainly a good instrument for the computation of a reliable estimate of erosion rates also if there wouldn't be a large amount of data.
Finally, additional statistical analysis tools would be useful to validate the results and the whole methodology.

References

- ✔ Confluence Inc. (2009). *Lower Clark Fork Tributaries Sediment TMDLS And Framework For Water Quality Restoration – Attachment 2 - Sediment contribution from hillslope erosion* (p. 65).
- ✔ Desmet, P. J. J., & Govers, G. (1996). A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *Journal of Soil and Water Conservation*, *51*, 427–433.
- ✔ Diodato, N. (2004). Estimating RUSLE’s rainfall factor in the part of Italy with a Mediterranean rainfall regime. *Hydrology and Earth System Sciences*, *8*(1), 103–107. doi:10.5194/hess-8-103-2004
- ✔ Farabegoli, E., Rossi Pisa, P., Costantini, B., & Gardi, C. (1994). Thematic cartography for the study of the erosion at the catchment scale. *Italian Journal of Agronomy*, *28*, 356–363.
- ✔ Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Koppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*, 259–263. doi:10.1127/0941-2948/2006/0130
- ✔ Moore, I. D., & Wilson, J. P. (1992). Length-slope factors for the Revised Universal Soil Loss Equation: Simplified method of estimation. *Journal of Soil and Water Conservation*, *47*, 423–428.
- ✔ Neteler, M., & Mitasova, H. (2008). *Open Source GIS A GRASS GIS Approach* (Third Edit, p. 406).
- ✔ Pieri, L., Bittelli, M., Wu, J. Q., Dun, S., Flanagan, D. C., Pisa, P. R., … Salvatorelli, F. (2007). Using the Water Erosion Prediction Project (WEPP) model to simulate field-observed runoff and erosion in the Apennines mountain range, Italy. *Journal of Hydrology*, *336*, 84–97. doi:10.1016/j.jhydrol.2006.12.014
- ✔ Pieri, L., Ventura, F., Gaspari, N., Salvatorelli, F., Vitali, G., & Rossi Pisa, P. (2013). Runoff in cultivated hilly areas as influenced by crops and land management. *Italian Journal of Agrometeorology*, *XVII*(1), 23–32.
- ✔ Renard, K. G., Foster, G. R., Weesies, G. A., & Porter, J. P. (1991). RUSLE: Revised universal soil loss equation. *Journal of Soil and Water Conservation*, *46*, 30–33.
- ✔ Servizio statistica e informazione geografica - Regione Emilia-Romagna. (2013). DBTR2013 - Curve di livello da Carta Tecnica Regionale 1:5000. Bologna: Archivio Cartografico - Regione Emilia Romagna.
- ✔ Servizio Sviluppo dell’Amministrazione digitale e Sistemi Informativi

Geografici Regione Emilia-Romagna. (2011). Ortofoto AGEA 2008 RGB. Retrieved from http://servizigis.regione.emilia-romagna.it/wms/agea2008_rgb

- ✔ Soil Survey Staff. (1975). *Soil Taxonomy, A Basic System of Soil Classification for Making and Interpreting Soil Surveys*. *USDA Agriculture Handbook* (Vol. 436, p. 754). Washington, DC: Soil Conservation Service, USDA.
- ✔ USDA NRCS. (2000). RUSLE Erosion Prediction. In *Field Office Technical Guide* (pp. 1–199). Ohio.
- ✔ Vianello, G., Lorito, S., Bigi, A., & Pavanelli, D. (2004). Stima dell’erosione e bilancio dei solidi del Sillaro: applicazioni del modello RUSLE e attraverso con il monitoraggio del trasporto in sospensione. *Arparivista*, *6*, 21–23.
- ✔ Wischmeier, W. H., & Smith, D. D. (1978). *Predicting rainfall erosion losses - a guide to conservation planning*. *USDA Agriculture Handbook* (p. 58). Washington, DC: Science and Education Administration, U.S. Dept. of Agriculture.
- ✔ Zampoli, M. (2007). *Analisi dei processi di erosione del suolo nel bacino del fiume Tusciano*. Università degli studi di Napoli “Federico II.”

MapboxGL + Protobuf + d3.js = liquid fast massively scalable interactive web map & data visualization!

Benjamin Wyss

GEM Foundation

Abstract

At the 2014 FOSS4G conference in Portland, Oregon I presented a web mapping and data visualization technique using raster tiles, this stack included Leaflet.js, UtfGrids, and D3.js. It provides a scalable and versatile approach to accessing, visualizing and interacting with data through a web mapping interface.

I have now updated the stack to use vector tiles, this stack includes Mapbox-GL, Protobufs, and D3.js which provides even more flexibility, speed and scalability for geospatial data visualization.

In this presentation I briefly summarize the raster stack, I will discuss the pros and cons of raster and vector tiles. I will also demonstrate the technology, and explain why vector tiles are even faster and more scalable then the raster approach. This presentation is not on the use of D3.js but rather how to get massive amounts of geospatial data from a web map to D3.js.

I am currently working on an open source project called OpenQuake. As a part of this project we are developing a platform which serves as a hub for integrated risk assessment. It allows users to combine seismic hazard, risk and social vulnerability in many different ways in order to obtain output for science, risk assessment, risk awareness and risk management.

The Platform thus needs to provide users with access to large geospatial data sets and models. This is why I am always on the search for faster and more scalable approaches to geospatial data visualization techniques.

Microservices and Geo - yeah it's that good

Steven Pousty

Red Hat

Abstract

You have started to hear about microservices and you want to learn more. We will start with an intro. to the basic ideas behind microservices. From there I will go over some pros and cons, covering the process I went through in constructing http://wwww.flatfluffy.com, a multi-device application for gamified recording of roadkill. I will focus on the geo pieces of the app. Come in curious, leave with some solid ideas to get started on your first microservice architected application.

BruGIS, a webGIS for Brussels Urban Planning: Past, present and future

Vandenschrick Gregoire[1], Micha Laurence[2]

[1] Centre Informatique pour le Région Bruxelloise (CIRB)/ Brussels Region Informatics Center (BRIC)

[2] Bruxelles Développement Urbain, Service Public de la Région Bruxelloise (BDU, SPRB)/ Brussels Urban Planning, Brussels Regional Public Service (BUP, BRPS)

Abstract

Since 2002, a small team of Brussels Urban Planning (BUP), part of the Brussels Regional Public Service (BRPS), worked on integrating and managing its multiple geomatic data sources (only a few of them being proper spatial data at the time) to ease their access, consulting and usage. The first aim was to be able to consult as many information sources available as possible through one integrated mapping solution in order to answer to public and private requests on the matters related to urban planning.

This integration lead in 2006 to the development and use of a cartographic website named BruGIS. This platform was first dedicated to the Civil Servants of BUP only and was accessible only on the Ministry Intranet. For that first version, the spatial database, data publication and website used proprietary solutions (Hexagon Intergraph Geomedia Webmap Server and client). Such choice was taken after a requirements and market analysis that reflected the state of the art at the time. Infrastructure was a mere intranet Windows Server.

With growing interest and use of the website, decision was taken in 2009 to open the platform to other entities, such as local authorities and citizens. In the meantime, a more robust geospatial database solution was adopted (Oracle Spatial) and some webservices were put online as well (WMS mainly). Shared databases maintained by Brussels Region Informatics Center (BRIC) were used. The Geomedia solution was moved on an Internet accessible Windows Server. BruGIS was online.

The load on the website increasing, a better technical solution was needed to serve it and its webservices. Tests were done, first with Open Source software (Geoviewer API) to develop a better client. OpenBruGIS was successfully put online in December 2011. It still gave access to the proprietary Geomedia Webmap Server webservices (WMS only).

Since 2011, The team focused all its efforts to adapt and deploy a complete OpenSource solution to serve its data and offer fully OGC compliant webservices. Geoserver was chosen and replaced the Intergraph solution in 2013 (2.1.3 version at the time of launch). The aging OpenSource Geoviewer client was replaced by a new development based on Geoexplorer, renamed BruGIS.

Currently, BruGIS is maintained through a complete development/staging/

production environment of 13 virtual machines running Windows 2008 SP2 Servers hosted by the BRIC. BruGIS Geoservers and website client are run on Apache Tomcat 8 application servers. Daily, BruGIS has a mean of 500 to 600 users.
The next challenges for the future will be to migrate all the data from Oracle Spatial databases to Postgres/PostGIS ones. Another path will be the development of a responsive design BruGIS for mobile viewing.
The presentation will focus first on the evolution of BruGIS since its inception to its current version. The different technical solutions chosen throughout its development and life as well as the infrastructure used will be explained. Planned development for the future will be given. A live demo will end the talk.

QGIS Server Python Plugins

Alessandro Pasotti

ItOpen

Abstract

A recent addition to QGIS Server is Python Plugins support.
Since many years QGIS provides full access to its C++ API through Python bindings and Python plugins played an important role in extending the functionalities of QGIS Desktop.
Following the same principle, Python plugins for QGIS Server allow to easily add new services and/or modify the behaviour of core C++ services (WMS, WFS etc.) without the need to write C++ code and follow the standard core committing path normally required for core contributions.
This means quick development, easy deployment and fast availability of new features for QGIS Server, effectively bringing the power of Python to QGIS Server development.
This paper provides an introduction of the core concepts behind Python plugins implementation for QGIS Server, a description of the new API and a few representative examples of what can be easily implemented using the new Python bindings.

GIS.lab – news from development of technology for rapid deployment of complete geospatial infrastructure with supercow features

Ivan Minčík

GIS.lab project

Abstract

Presentation will summarize last year of active development of GIS.lab [1] technology with overview of new features and plans for near future. Also, it will provide a brief description of DevOps tools and principals used for development, testing and fully automatic deployment of scalable infrastructure equipped with free geospatial software.

GIS.lab is capable to deploy a complete GIS infrastructure in local area network (LAN), data center or cloud in a few moments. It provides comprehensive set of free geospatial software, seamlessly integrated in to one, easy-to-use system with desktop, web and mobile client interfaces. GIS.lab lowers deployment and ownership cost of complex GIS infrastructure to absolute minimum, while still keeping whole technology in house and under full control.

GIS.lab can be used in places and conditions where deployment of any other technology wouldn't be affordable or technically possible. GIS.lab is capable to turn bunch of heterogeneous or broken computers in to crisis management command center, flawlessly working in very hard conditions of natural disaster with power and Internet outages. It is also ideal system for education or just for Open Source technologies popularization.

References

✔ [1] - GIS.lab home page (http://imincik.github.io/gis-lab)

Let's get parallel

Graeme Bell

Norwegian Forest and Landscape Institute

Abstract

During the last ten years, computer components such as CPUs, disks, and GPUs have seemed to hit a limit in terms of the speed of their individual components. Increasingly, modern computer servers instead provide parallel computing resources (multi-core CPUs, wide disk arrays) in order to increase total system performance.

In domains such as bioinformatics and computer games, software is often capable of making good use of such parallel resources. Unfortunately, most GIS applications, databases and operations make only limited use of parallelisation. Consequently, there have been only moderate performance improvements seen by GIS users as the world has moved from single processor servers to 64 processor servers. So, what can a GIS user do when they have a powerful computer and an urgent deadline to meet? How do they carry out complex national-scale analysis several times per day rather than per month?

At the Norwegian Forest and Landscape Institute, we have GIS/database servers with 4 to 32 CPU cores and up to 16 disks. We've repeatedly faced the challenge of extracting the maximum utility from these servers in order to finish projects that traditionally would take days, weeks or months to run to completion. In one recent project, we simulated up to 100 years of change in forested areas at high resolution, on a national scale, and on a year-by-year basis. We transformed and intersected 4 national forest datasets, each with millions of records, along with other diverse datasets (water, roads, and conservation areas). Using our fastest server, with a high degree of parallelisation and 'divide and conquer' techniques, a simulation run can be completed in around an hour.

The first part of the presentation demonstrates a variety of techniques that can be used to obtain maximum performance from modern computing equipment. This can mean substantially increasing the detail of your work, or substantially decreasing the time taken to perform ordinary GIS operations. Some techniques are as simple as making system-wide configuration changes at the hardware, operating system or database level. Others depend on appropriate use of long-established open source software such as Gnu Parallel (Tange, 2011) and Make, or newer systems such as Drake or Bpipe (Sadedin et al., 2012).

The second part of the presentation introduces new approaches that are enabled by the GIS parallelisation tools that we have developed locally at our institute. These new tools allow for an order of magnitude improvement in some geometry and raster GIS operations. Here, we are sharing them with the open source community under an open source license.

References

- O. Tange (2011). GNU Parallel - The Command-Line Power Tool. ;*login: The USENIX Magazine*, February 2011, 42-47.
- Sadedin et al. (2012). Bpipe. *Bioinformatics, 28(*11), June 2012, 1525-1526.

What happens when you put 1 billion points into Postgis Topology?

Lars Aksel Opsahl

Norwegian Forest and Landscape Institute (SogL, NIBIO)

Abstract

In this presentation we will take a closer look at how Postgis Topology can be used for large, national scale datasets. The building blocks of Postgis Topology that are available on internet are generally ready to cope with “big data”, but we need to use these building blocks very carefully in order to be able to handle extremely large datasets in an efficient way.

Our institute is currently running a full scale test of moving from a simple feature map representation in Postgis to a Postgis Topology representation for the land resource map (Ar5) for Norway. Ar5 contains 947967792 points and 7989051 surfaces, and covers all of Norway with out gaps. The Ar5 layer is administered as a single entity – it is not split up into separate administrative units. The distance between the points may vary from a few meters to many thousand meters and the points are recorded with centimeter scale accuracy. Norway's country shape is very long and thin, including a long Norwegian coast line and also very long roads traveling from south to north which must be represented. There are also data attributes connected to the lines as well as the surfaces. The layer is updated by a variety of people in many organizations, and the same map area might be updated at the same time by different people.

For us, there are several advantages to moving to a topology model. Our national standards require topological representations wherever possible, and historical data is recorded in a topological format. We need to be certain that we do not have multiple representations of land status or missing coverage; we can avoid well known GIS problems like 'slivers' and 'holes' in our coverage. However Postgis Topology has seen less development than the rest of Postgis, which leads in part to the discussion of the presentation.

This paper presentation explores how we made the migration from Postgis simple feature to Postgis Topology and how we plan to use it in a production environment for map production and updates.

Key topics that will be discussed:

- Generation of a content based grid model, where cell area varies allowing the number of geometries per cell to be kept balanced; enabling efficient parallelized divide and conquer algorithms.
- Migrating from linestring and representation points, to a Postgis Topology model.
- How to model and process line and surface attributes in Postgis Topology.

- Handling updates of the topology layer from clients used by non-GIS users.
- Improving performance by caching simple feature interpretations of the topology.
- Java, Hibernate Spatial.

The talk will start from models and delve into relevant sql code to explain how the underlying postgis topology features enable us to build the tools that are needed.

Ice: lightweight, efficient rendering for remote sensing images

Julien Michel[1], Gabriele Facciolo[2]

[1] CNES, DCT/SI/AP

[2] CMLA, ENS-Cachan

Abstract

Rendering large remote sensing images is always challenging. Recent web based mapping technologies have pushed new ergonomy standards since any viewer provides almost instant zooming and panning as a baseline feature. Yet web mapping relies on heavily preprocessed data, so that lightweight, multi-resolution, and optimised 8-bits images could be transmitted efficiently to the end-user. But when it comes to navigation in a stack of raw satellite images, with wider dynamic, many spectral bands, and several gigabytes of data each, then these standards may be difficult to achieve.

This talk introduces Ice (Michel, 2014), a side project of the open-source software Orfeo ToolBox (http://www.orfeo-toolbox.org). Ice is an OpenGL only rendering engine written in C++, which also comes with a demo viewer based on GLFW. Two key concepts that make Ice attractive are inspired from the pvflip lightweight image viewer (Facciolo, 2013): loading the images on the GPU memory as 32-bits textures, and render them using OpenGL Shaders. Having the images available in the GPU memory enable instant navigation operations. While rendering with shaders permit instant contrast manipulation, resulting in fun rendering tools like real-time contrast adaptation based on the intensity of the pixel under mouse pointer.

From these two core concepts, inherited from pvflip, Ice implements specific features for remote sensing images. First, Ice does not load entire images into the GPU memory, instead it only loads the required tiles at the required resolution, thanks to the overview feature from GDAL (http://www.gdal.org). Then, Ice distinguishes between a heavy and a lightweight rendering cycle: the first one may involve data transfers between hard drive and GPU, the second does not. This permits to design user interactions that are very fast, because they only rely on the lightweight cycle and therefore mainly involve GPU operations.

But there is more. OTB has a nice feature, which is made possible thanks to OSSIM (http://www.ossim.org), that allows to convert between coordinates in different reference systems, including non georeferenced raw sensor images. On the other hand OpenGL can map textures on any quadrilateral. Therefore, in Ice the viewport geometry is not bounded to the actual image geometry, since coordinates of the corners of a tile can be converted to arbitrary quadrilaterals on which textures can be mapped. You can have a stack of raw satellite images, and use a cartographic viewport geometry. Ice will perform rough

orthorectification of those images on the fly, even considering a digital elevation model if one is provided! This means that you can open a set of images from different sensors and have them instantly mosaicked together.
Ice is a rendering engine: you can use its API to bring remote sensing image rendering into your application. This is already the case for our end-user software Monteverdi2 (Grizonnet & Michel, 2013), which renders images through Ice. There are lots of nice features that were not mentioned here: vector data rendering, image rotation, fun shaders, and more. The talk will walk through all these features and will also include a live demonstration of Ice!

References

- ✔ Facciolo, G. (2013). pvflip (Python vflip): OpenGL accelerated image viewer. https://github.com/gfacciol/pvflip
- ✔ Grizonnet, M., & Michel, J. (2013). Monteverdi 2.0 - Remote sensing software for Pleiades images analysis. Paper presented at the Proceedings of the 33th EARSeL Symposium, Matera (Italy), 3 - 6 June (pp. 467-476), EARSeL and CNR.
- ✔ Michel, J. (2014). ICE: fast OpenGl rendering API for remote sensing images. https://hg.orfeo-toolbox.org/Ice/

Web services at the State Office for Spatial Information and Land Development, Baden-Württemberg, Germany

Alexander Schmidt, Andreas Schmidt

Landesamt für Geoinformation und Landentwicklung, Baden-Württemberg, Germany

Abstract

In recent years a lot of the challenges in the publication of online spatial data has arisen and the public authorities have to adapt to the new AFIS, ALKIS and ATKIS (AAA) data model (AdV 1997) and the Infrastructure for Spatial Information in the European Community (INSPIRE 2007). Every new directive is a major change for the whole infrastructure and a chance for open source technology for the state agency for Geo-Information and Rural Affairs of Baden-Württemberg which has the mandate to provide a variety of spatial data.

The new AAA data model is very complex and therefore difficult to grasp, understand and time-consuming for the end user. To overcome this issue, a user oriented spatial data model (NOrA_BW) was developed. The data model was developed in cooperation with the Environmental Agency of Baden-Württemberg that contains all important information for customers but uses a flat and simple data structure. As INSPIRE is still a priority it was also developed to represent the INSPIRE data model. All data is stored in PostGIS database and is planned to be updated daily in the final version therefore customers who access the data via Web Map Service (WMS) or Web Feature Service (WFS) will have the most recent and most reliable data. In addition to vector data, there are vast amounts of raster data (e.g. aerial imagery) which also has to be made available. The vector data is completely distributed by GeoServer whereas the aerial imagery is currently made available by ArcGIS Server or ERDAS Apollo Server.

To provide fast raster services with high availability for (web) applications, all the highly frequented WMS services are cached in a GeoWebCache and distributed as Web Map Tile Service (WMTS). The GeoWebCache has to update the affected tiles since the vector data is going to be updated on a nightly basis and the aerial imagery will also be frequently updated – there can be a few hundred changes within the vector data which spread all over the state. The implemented GeoRSS update mechanism was not ideal for the data since it spans an envelope between two objects, which in a worst case scenario then reseeds the whole area of the state, therefore the changes are instead clustered and then reseeded with an adapted algorithm.

The only access point to the data is an OpenGIS Web Service Proxy which restricts the access to WMS, WMTS and WFS services. This is especially important since all the data is not Open Data and has to be purchased. The OWS-Proxy is based on the deegree OWS and was extended to support different constraints and to log all user access to charge their accounts. Additionally, the OWS distributes static capabilities to fulfill the INSPIRE standards.

To enable the broad public to use our services a new OpenLayer client is in development which supports simple routing based on GraphHopper and is able to geocode spatial entities by utilizing the full text search by PostgreSQL. On top of that, several extensions will also be developed.

References

- ✔ Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland [AdV]. (1997). AFIS-ALKIS-ATKIS-Modell. Retrieved from http://www.adv-online.de/AAA-Modell/
- ✔ European Parliament. (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community. Retrieved from http://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32007L0002&from=EN

S2P: a new open-source stereo pipeline for satellite images

Julien Michel[1], Carlo de Franchis[2], Enric Meinhardt-Llopis[2], Gabriele Facciolo[2]

[1] CNES, DCT/SI/AP

[2] CMLA, ENS de Cachan

Abstract

One of the great features of modern Very High Resolution satellites (VHR) is their ability to acquire several views of the same landscape in a short time frame. For instance, the French Pléiades satellites can perform three acquisitions of the same scene from the same orbit with slightly different viewing angles, all in about ten seconds. A direct application of such acquisitions is the reconstruction of a high resolution Digital Surface Model (DSM), which is made easier by the small viewing angles difference and the short delay between the views.

Yet, while 3D reconstruction is a common task when dealing with photographs or aerial images, its application to satellite imagery remains complex, and hardly achievable with open-source solutions, with the notable exceptions of the NASA Ames Stereo Pipeline (Moratto et al., 2010) and the French national geographic institute MicMac software (Deseilligny, 2007).

The most important issue when dealing with VHR satellite images is that they are not acquired with a regular camera, but rather by a single line of detectors called a pushbroom sensor and the images are formed line by line as the satellite moves. Fundamental matrix modelling of the line of sight cannot be applied to pushbroom images and the so called epipolar geometry, where height or depth variation results in image disparities in horizontal direction only, does not formally exist. Thus, the software must take into account a more complex, less common sensor model that prevents from using many state-of-the-art 3D reconstruction algorithms from the computer vision field, because these methods are usually bound to work on epipolar rectified images.

Secondary issues are the size of the images, which requires out-of-core processing, the specific content of the images (earth surface viewed from 700 km above), and the variety of formats and exogenous data that must be considered (such as Spatial Reference Systems, or pre-existing low-resolution Digital Elevation Model for solution initialization).

In this talk, we will introduce and announce the first open-source release of S2P (De Franchis et al., 2014), a software jointly developed by CMLA and CNES in the frame of the Pléiades mission. S2P is a processing chain written in Python and C, that tries to address the issues listed above and provides a new open-source alternative to 3D reconstruction from VHR satellite images. From a user perspective, S2P aims at providing a simple, almost automatic solution. From a scientific point of view, the core concept of S2P is to create controlled and optimized piecewise epipolar approximations to enable the use of state-of-the-

art stereo matching algorithms from the computer vision field. Indeed, the matching part is modular and anyone can plug his own algorithm, while S2P provides the end-to-end plumbing for tile-wise processing, geometric alignment, point cloud colorization, and raster DSM synthesis. As such, it greatly facilitates the assessment and comparison of matching algorithms performance on complex remote sensing datasets.
The presentation will provide a detailed description of the S2P workflow and algorithms, and will also include a live demonstration.

References

- ✔ De Franchis, C., Meinhardt-Llopis, E., Michel, J., Morel, J. M., & Facciolo, G. (2014). An automatic and modular stereo pipeline for pushbroom images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1*, 49-56.
- ✔ Deseilligny, M. P. (2007). B. 1 MicMac, un logiciel pour la mise en correspondance automatique d'images dans le contexte géographique. *Bulletin d'Information Scientifique et Technique de l'IGN n*, *77*, 1.
- ✔ Moratto, Z. M., Broxton, M. J., Beyer, R. A., Lundy, M., & Husmann, K. (2010, March). Ames Stereo Pipeline, NASA's open source automated stereogrammetry software. In *Lunar and Planetary Science Conference* (Vol. 41, p. 2364).

Spatial tools for LiDAR based watershed management and forestry analysis

Andrea Antonello[1], Silvia Franceschi[1], Francesco Comiti[2], Giustino Tonon[2]

[1] HydroloGIS Bolzano

[2] Free University of Bolzano

Abstract

The increased availability of Airborne Laser Scanning (ALS) surveys and ALS-derived datasets in the last decade pushed scientists to find their way to properly exploit the detailed information contained to extract different features and forest and terrain characteristics. The geomorphological tools based on LiDAR derived raster data increased their importance in watershed management in particular for the evaluation of natural hazard: floods, landslides and avalanches.

JGrassTools[1] is an open source geospatial processing library containing tools for geomorphometry analysis, models for the evaluation of maximum discharges, hillslopes stability and hydrological indexes, utilities to handle raster and vector data and to prepare the base inputs for external complex models (HECRas, NewAge). Beside this, in 2014, we started the development of LESTO (LiDAR Empowered Sciences Toolbox Opensource): a set of modules for the analysis of LiDAR point cloud with an Open Source approach, with the aim of improving the extraction of the position of the trees, the volume of biomass and other vegetation parameters on large areas for mixed forest structures. The extraction follows the single tree based approach starting with the implementation of some of the most used algorithms in literature. These have been tweaked and applied on LiDAR derived raster datasets (DTM, DSM) as well as point clouds of raw data. The methods range between the simple extraction of tops and crowns from local maxima, the region growing method, the watershed method and individual tree segmentation on point clouds. LESTO contains a set of modules for data handling and analysis implemented within the JGrassTools spatial processing library. The main subsections are dedicated to 1) preprocessing of LiDAR raw data mainly in LAS format (utilities and filtering); 2) creation of raster derived products; 3) flight-lines identification and normalization of the intensity values; 4) tools for extraction of vegetation and buildings.

On a watershed level, the resulting extracted trees with position and main characteristics can be used for forestry management (evaluation of the volume of biomass) or for the evaluation of natural hazard related to the hillslopes stability and to the possibility of large wood transportation during flooding events. Within the last Google Summer of Code (2014) we developed a set of tools to evaluate large wood recruitment and propagation along the channel

1www.jgrasstools.org, https://github.com/moovida/STAGE/releases

network based on a simplified methodology for monitoring and modeling large wood recruitment and transport in mountain basins implemented by Lucía et al. (2014). These tools are also integrated in the JGrassTools project as a dedicated section in the Hydro-Geomorphology library. The section LWRecruitment contains 10 simple modules that allow the user to start from very simple information related to geomorphology, flooding areas and vegetation cover and obtain a map of the most probable critical sections on the streams.
During the presentation a detailed list of the available tools and some examples of application will be presented on test sites.

References

✔ A. Lucía, A. Antonello, D. Campana, M. Cavalli, S. Crema, S. Franceschi, E. Marchese, M. Niedrist, S. Schneiderbauer and F. Comiti. Monitoring and Modeling Large Wood Recruitment and Transport in a Mountain Basin of North-Eastern Italy. Engineering Geology for Society and Territory - Volume 3, edited by Giorgio Lollino, Massimo Arattano, Massimo Rinaldi, Orazio Giustolisi, Jean-Christophe Marechal, and Gordon E. Grant, 155–58. Cham: Springer International Publishing, 2015. http://link.springer.com/10.1007/978-3-319-09054-2_31.

Use of indexes on geospatial database with the PostgreSQL DBMS

Giuseppe Broccolo

2ndQuadrant Italia

Abstract

The technological progress in GIS has resulted in the management of increasingly large and complex data, so it is more and more necessary to lean on databases that simplify and especially speed up data access.
The (open source) PostgreSQL DBMS provides many useful features, making it particularly suitable for the management of geospatial databases. In this work I will present the available PostgreSQL tools to manage the simplest but also one of the most adopted geospatial datatype in GIS, from geolocation to LiDAR surveys: the point.
Point indexing is crucial for the performance of GIS applications. The choice of the best index type strongly depends on the kind of operation that we have to do on the points, making the indexing of this particular geospatial datatype more complicated than others. PostgreSQL provides many kinds of indexes implementing the most common indexing algorithms for spatial data, with each one more suited than the others for specific operations.
Moreover, multiple kinds of 'points' are defined depending on the information that must be stored together with the geospatial one. PostgreSQL has defined "in core" a first type of two-dimensional geometric points. The PostgreSQL GIS extension, PostGIS, increases the collection of points data types, introducing also three and four-dimensional objects (i.e. points with an additional information), and geographical points (two-dimensional points referred to a given datum frame identified by its SRID[1]). For points that need higher number of related information, as in the case of LiDAR survey points, a second PostgreSQL extension (PointCloud) has been created, adding more complex kind of points (whose size on disk is not fixed[2] but varies with the actual accuracy of the stored data) and features to manage a large amount of information.
All these datatypes support the various indexing algorithms in PostgreSQL, and the extensions mentioned above expand the set of available algorithms as well. Some simple (but also frequently used in GIS) examples will be presented here to better understand the potential of PostgreSQL, such as the search of the k closest points to a given one (using the kNN algorithm) and the search of a

1 Spatial Reference system IDentifier: a unique value used to unambiguously identify projected, unprojected, and local spatial coordinate system definitions following the Open Geospatial Consortium (OGC) standard. See the OGC website (2015): http://www.opengeospatial.org/.
2 Both PostgreSQL standard 2D point and PostGIS 2D geometric point are implemented in double precision. See the OSGeo rants about that (2015): http://trac.osgeo.org/postgis/wiki/ToleranceDiscussion.

subset of points included inside a given bounding box. It will be shown which indexes are the best ones for each specific operation and for each kind of considered points datatype, understanding how the query strategy adopted by PostgreSQL changes in the various cases.

References

- ✔ The PostgreSQL geometric datatype (2015): http://www.postgresql.org/docs/9.4/static/datatype-geometric.html
- ✔ Regina O. Obe, Leo S. Hsu, & Paul Ramsey, Manning (Eds.). (2013). *PostGIS in Action* (2nd ed.). Stamford, CT.
- ✔ Using PointCloud extension for PostgreSQL, the OpenGeo Suite (2014): http://suite.opengeo.org/4.1/dataadmin/pointcloud/index.html

GDAL 2.0 overview

Even Rouault

SPATIALYS

Abstract

Initiated in 1998, the GDAL/OGR (Geospatial Data Abstraction Library) library has dramatically grown, supporting from 20 raster and vector formats in its early 1.1 version to more than 200 in the latest v1.11. In the meantime, it has been successfully adopted by hundreds of software projects, being at the foundation of many free and open source or proprietary GIS stacks.

In this session, after an overview of the current capabilities, as well as the state of the current user and development community, we will focus on recent developments and achievements that will be available in GDAL 2.0:

- unification of the raster and vector driver models,
- performance and capability improvements in raster operations : resampling, block cache thread scalabality improvements, GeoTIFF special reading modes
- support of ISO SQL/MM curve geometry types,
- new vector attribute data types : 64 bit integer values and identifiers, boolean values, ...
- vector fields not-null constraints and default values,
- multiple geometry fields per feature,
- support of new formats (GeoPackage vectors and rasters, KEA, etc) and enhancements in existing ones (PostGIS, GML, MapInfo .TAB etc).

The impact on existing downstream software and the porting efforts will be addressed.

Finally, we will explore potential future directions for the project : planar topology abstraction and drivers, geography networks, Cmake unified multi-platform build system, etc.

References

✔ GDAL/OGR website: http://gdal.org/

✔ GDAL Request For Comments list: https://trac.osgeo.org/gdal/wiki/RfcList

Classifier to map the burned areas in the Canary Islands with MODIS data

Jose R. Garcia-Lazaro[1], Jose A. Moreno-Ruiz[1], Manuel Arbelo[2], Pedro A. Hernandez-Leal[2], Alfonso alonso-Benito[2]

[1]Grupo de Tratamiento de Imágenes, Departamento de Informática, Universidad de Almería, 04120 La Cañada de San Urbano, Almería, Spain

[2]Grupo de Observación de la Tierra y la Atmósfera (GOTA), Departamento de Física, Universidad de La Laguna, 38200 La Laguna, Spain

Abstract

A Bayesian classifier from WEKA (Waikato Environment for Knowledge Analysis) Data Mining Software (Hall et al., 2009), originally developed and validated for the detection of burned areas in boreal forest using images from NOAA-AVHRR sensor at 0.05º (~ 5 km) of spatial resolution (Moreno-Ruiz et al., 2012), was applied to a time series of Terra-MODIS images of forest areas of the Canary Islands (Spain) for the period 2002-2012. From daily images of reflectance and surface temperature of MOD09GQ (250 m) and MOD11A1 (1km) MODIS products respectively, 10 days product compounds were constructed using the criterion of maximum temperature. The statistics variables used in the classifier were GEMI (Global Environmental Monitoring Index) and BBFI (Burn Boreal Forest Index) vegetation indices and the MODIS NIR spectral band, for the previous year and the year of occurrence of the fire. Polygons reference of the fourteen fires greater than 100 hectares identified in the analyzed period were created using Landsat-TM post-fire images and official information from the national database of wildfires of the Ministry of Agriculture, Food and Environment (MAGRAMA). The detection algorithm was trained using a fire in the south of Tenerife in July 2012 which affected more than 6000 ha. The results show the detection of 13 out of 14 fires. The total burn area detected represents a 66.5% of MAGRAMA reference data and a 78.4% by Landsat data. These differences mainly come from omission error (near 50%) in the estimation of burned area of a large fire in the summer of 2007. The proposed methodology, still in the experimental stage, could improve these results using other statistic variables and/or vegetation indices more suited to the spectral response of forest cover and burned areas of the Canary Islands.

Acknowledgements

This work was supported by the Ministerio de Economia y Competitividad (Spain) under Grant CGL2013-48202-C2.

References

- ✔ Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.
- ✔ Moreno-Ruiz, J. A., Riaño, D., Arbelo, M., French, N., Ustin, S. L., Whitinget, M. L. (2012). Burned area mapping time series in Canada (1984–1999) from NOAA-AVHRR LTDR: A comparison with other remote sensing products and fire perimeters. *Remote Sensing of Environment*, 117, 407-414. doi: 10.1016/j.rse.2011.10.017

Earning Your Support Instead of Buying it: A How-to Guide to Open Source Assistance

Ian Turton

Envitia Ltd, North Heath Ln, Horsham, RH12 5UX, UK; ian.turton@envitia.com

Abstract

More organisations are moving to use FOSS4G software to cover shrinking budgets. It is very appealing to an organization's leaders to ditch their current proprietary software solution with the attendant saving on per user licences and ongoing maintenance costs. Obviously, if you switched to FOSS4G to get better features and scalability you should consider buying a support contract from one of the many vendors that offer them, these companies support many of the core developers directly. This way you get all the advantages of open source, prompt support and often the chance to ask for new features. However, if you (or your boss) are looking to save money then you are moving from a cash economy to a gift economy. In a gift culture you need to build up your "capital" before attempting to take too much out.

For example, you've downloaded the software and installed it, and all looks good. Then disaster hits, you have a demo for the CIO and something isn't working; Time to hit the user list, the developer list, stack exchange. Why can't you get an answer? Remember just because your issue is urgent to you the developers might be in the middle of a new release or adding a new feature and have much more important (or fun) things to do with their time. They will notice they have never seen your name before on the list, or on Stack Exchange that you have a reputation in the low single digits – thus you are a newbie. There's no harm in that but wouldn't it be better to have got that out of the way before your emergency. You could have built up your reputation by asking some questions earlier – especially questions like "what can I do to help?" or "I found an unclear paragraph in the install instructions, how do I fix it for you?" on a mailing list. On Stack Exchange you can build reputation by asking good questions and by answering other people's questions.

Once you've banked some capital there are still good and bad ways of asking a question. Developers are busy people (the GeoTools users list has 20-30 messages a day for example) no one has time to read all of them closely. If you use a poor subject (e.g. "Help!!!!") or don't provide a clear description of the problem (e.g. "it crashes, your software sucks") then the odds of being ignored are huge. It can be tempting once you have found a helpful (or friendly) developer to keep emailing them directly, but this is likely to lead a polite(ish) reminder to keep to the list so that everyone can benefit or silence.

This talk will attempt to show how to be a better open source citizen and get a better answer than RTFM or STFW when your project is stuck and the demo is the next day. The author will share his experience with helping users and

developers on the GeoTools and GeoServer mailing lists and, more recently, as a moderator on gis.stackexchange.com.

Case Study: Mapping Mangroves and Coastal Wetlands

Christopher Dubia, Kent Lewis, Andy Long

MapWorks Learning

Abstract

Mapping the Mangroves (MTM), a project of MapWorks Learning, provides formal and informal education, and gives citizen scientists and the larger scientific community the ability to engage with and explore mangroves and their ecosystems. Mangroves are a cornerstone species and play important roles in habitat formation, stabilization of coastal environments, and carbon sequestration. The MTM open curricula and GIS tool provide opportunities for anyone to learn about authentic applications of GIS in the field, explore mangroves and their ecosystems, and share their findings. Learners develop an understanding and appreciation for the role mangroves play in a healthy environment and how GIS can aid in conservation.

References

✔ About MapWorks Learning. (2014, April 8). Retrieved March 4, 2015, from http://www.mapworkslearning.org

✔ Baldwin, D., Dubia, C., & Long, A. (2013, July 7). MapWorks Learning. Retrieved March 4, 2015, from http://www.oercommons.org/groups/mapworks-learning/204

✔ Balga, P. (2013, November 11). Mangrove Restoration in Guanabara Bay - Rio de Janeiro - Brazil. Retrieved February 24, 2015, from http://mappingthemangroves.org/reports/view/351

✔ Dubia, C. (2012, April 1). Mapping the Mangroves. Retrieved March 4, 2015, from http://www.mappingthemangroves.org

✔ Meinke, B. (2014, August 28). MapWorks Learning combines OER and open data to protect threatened biodiversity. Retrieved February 24, 2015, from http://creativecommons.org/weblog/entry/43541

✔ Ushahidi. (n.d.). Retrieved March 4, 2015, from http://en.wikipedia.org/wiki/Ushahidi

An Overview of OpenSensorHub for SensorWebs and IoT

Mike Botts[1], Alex Robin[2]

[1] Botts Innovative Research Inc.

[2] Sensia Software

Abstract

OpenSensorHub (http://www.opensensorhub.org) is an open source software community being built around the development and deployment of a new software stack supporting the full vision of the OGC® Sensor Web Enablement (SWE) suite of standards, as well as some IoT protocols (https://github.com/sensiasoft/sensorhub). The goal is to allow easy deployment all types of sensors/actuators (in-situ and remote sensors, mobile or static, etc.) within a cross-domain context and interconnect them through the web in a fully interoperable manner.

OpenSensorHub supports the full range of sensors and actuators (both simple and complex) while providing capabilities for discovery, quality assessment, data streaming, tasking, visualization, and processing (onboard or online). Yet a major driver for OpenSensorHub is to ultimately provide this functionality to sensor deployers and exploiters in a package that is easily deployed and configured with minimal development required. The OpenSensorHub community is focusing on all aspects of this challenge, from web service support to javascript clients to online middleware components.

The main development team for OpenSensorHub has been heavily involved for 15+ years in designing sensor web standards and implementing software solutions for interoperability of disparate sensors. This includes editorship of two of the OGC SWE specifications, chairmanship of the SWE Domain Working Group, and development of various server and client tools such as the Space Time Toolkit. But the team is very eager to work with other developers in order to grow the OpenSensorHub community and bring to reality the vision of a fully connected and interoperable web of sensors, actuators, and processes.

This presentation will provide an overview of the goals and architecture of OpenSensorHub and will define the potential roles of deployers and software contributors within the open source software community. In addition, both live and recoded demonstrations of the current capabilities of OpenSensorHub will be provided. Finally, the presentation will discuss how the FOSS4G community can help advance the project by being adopters and code contributors.

References

- ✔ Botts, M., Percival, G., Reed, C., & Davidson, J. (2006). OGC Sensor Web Enablement: overview and high level architecture. DOI:10.1007/978-3-540-79996-2_10 Conference: GeoSensor Networks, Second International Conference, GSN 2006, Boston, MA, USA, October 1-3, 2006, Revised Selected and Invited Papers.
- ✔ Botts, M. (2011). OpenGeo Sensor Web Enablement (SWE) Suite. White Paper: http://boundlessgeo.com/wp-content/uploads/2013/09/OpenGeo-SWE.pdf.
- ✔ Botts, M. (2014). OGC SensorML: Model and XML Encoding Standard, Version 2.0, OGC 12-000. https://portal.opengeospatial.org/files/?artifact_id=55939.
- ✔ Robin, A. (2011). OGC SWE Common Data Model Encoding Standard, OGC 08-094r1. http://portal.opengeospatial.org/files/?artifact_id=41157.

Using Free and Open Source Solutions in Geospatial Science Education

Vaclav Petras[1], Anna Petrasova[1], Keren Cepero Perez[1], Markus Neteler[2], Luca Delucchi[2], Martin Landa[3], Helena Mitasova[1]

[1] Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, USA

[2] Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy

[3] Faculty of Civil Engineering, Czech Technical University in Prague, Czech Republic

Abstract

Free and open source software (FOSS) is a crucial part of open and reproducible science and is becoming increasingly important for the geospatial industry. We address these needs in our geospatial education program by integrating FOSS into our courses. In our talk, we share our experience with teaching geospatial modeling courses at North Carolina State University (NCSU), focusing on the recent improvements, usage of new GRASS GIS 7 and incorporations of innovative technologies in the course materials. Moreover, we will show how the course material can be used in different countries with location specific data.

Since 2008 our graduate students have the opportunity to take a course on geospatial modeling and analysis taught with both open source and proprietary software. In the hands-on assignments students use the free and open source package GRASS GIS and the proprietary package ArcGIS to perform similar geospatial tasks. Each assignment is accompanied by a software-independent lecture on the theory behind the particular geospatial concepts. By using two software packages, our goal is to ensure that students can distinguish more easily between GIS concepts and specifics of GIS software, become true experts in GIS and gain flexibility to choose optimal solutions for their future work.

Recently we have redesigned our course materials to be easier to follow and navigate for our on-campus and distance education students, but also to make them easier to share and modify for the FOSS community of educators. Our materials are static HTML websites which are highly portable and simple to edit. We use the Git revision control system in combination with GitHub to manage and share the source files. We developed Python scripts to build the website and to extract all GRASS GIS assignment instructions into a script which can be executed in GRASS GIS in order to test and ensure all instructions are correct and give expected results. Our teaching material is available online, being shared under Creative Commons Attribution-ShareAlike (CC BY-SA) license.

The assignments in our courses are based on the "North Carolina sample dataset", which is also used in GRASS GIS manual pages and tutorials. However, instructors in other regions often prefer to teach with local datasets which makes sharing of hands-on teaching material challenging. Therefore, we are currently transitioning to an updated version of the North Carolina sample dataset which includes frequently used and commonly available data with standardized names such as elevation, streets, or lakes. This will allow any instructor to create new datasets in different

parts of the world with names compatible with our teaching materials, and thus helping the students to learn geospatial concepts in the local context. We will present sample results from these shared assignments for North Carolina, Puerto Rico, Italy and the Czech Republic.
Link to the geospatial modeling and analysis course:
http://courses.ncsu.edu/gis582/common/
Link to list of datasets with examples:
http://grasswiki.osgeo.org/wiki/GRASS_GIS_Standardized_Sample_Datasets

Processing, serving and rendering huge point clouds on Mobile devices and Web pages

Manuel de la Calle[1], Diego Gómez-Deck[1]

[1] Glob3 Mobile

Abstract

Rendering 3D point clouds on mobile devices is a hard task because we are working online and we work on limited performance devices. We have developed a server-client library that allow developers to serve any size point clouds on any environment

Keywords

Point clouds, HTML5, Android, iOS,3D, LiDAR

1 Mobile Map Technology (Glob3 Mobile)

Mobile Map Tools is a library to build Native Mobile Map Apps. Visualize huge points cloud in a 3D environment is a hard task that could be face using a server-client approachment.

1.1 G3M capabilities

Glob3 Mobile is an API for the development of native map applications on mobile devices.
The main capability of this library is the Multiplatform approach, it have the very same API in all environments thanks to coding translation.
Developing with Glob3 Mobile you can save time and resources when you face a mobile development having all advantages of native development (Performance, UI, Access to disk, sensors, etc) and the simplicity of an API thought for GIS developers.
Glob3 Mobile has been developed thinking in the usability and the UI of mobile devices.
Currently Glob3 Mobile is working in the next platforms:

- iOS
- Android
- Google Glass
- html5-webgl

The capabilities list is huge but the main are:

- Raster data
- Vectorial data
- Point Clouds

- 3D models (Buildings, cities, vehicles, ...)
- 4D Data
- Real Time
- Simbology
- Offline – Online -> Cache Handling
- 3D- 2D – 2,5D views
- Scenarios
- Animations
- Cameras
- Tasks

3 The problem

Render huge point clouds on mobile devices (and on the web) is not an easy task, we have some problems:

- Connectivity and data transmisions
- 3D handling
- Device capabilities

Moreover we have all the problems because we are working on a multiplatform environment.

4 The implemented solution

We face this problem using techniches like stream the data and the LoD

- Very big point cloud could be served using a streamed format
- The points serverd save the shape of the cloud
- Using LoD techniques and MMT you can see the point cloud in 3D on any mobile device or web page.
- The point cloud must be pre-processed

The library developed has the following products

- A tool to import the point clouds from diferent formats
- A library to save this data on a Berkeley DB (To improve the speed of reading of data)
- A library to pre-process fastly this points
- A server to give the correct points on every tile to the clients apps. This server streams the data.

5 About g3m

Mobile Map Technology is a library developed by Glob3 Mobile Inc, and is under the Location Tech's umbrella.

5.1 A live demo

http://point-cloud.glob3mobile.com/

References

- ✔ https://github.com/glob3mobile/g3m
- ✔ http://point-cloud.glob3mobile.com/
- ✔ https://2015.foss4g-na.org/session/processing-serving-and-visualizing-huge-point-clouds-mobile-devices-and-web-pages
- ✔ https://vimeo.com/107476722

Shortest Path search in your Database and more with pgRouting

Daniel Kastl[1]

[1] Georepublic

Abstract

pgRouting extends the PostGIS / PostgreSQL geospatial database to provide shortest path search and other network analysis functionality.

This presentation will show the inside and current state of the pgRouting development, from its wide range of shortest path search algorithms to driving distance calculation or “Traveling Sales Person” (TSP) optimization. Additionally we will give a brief outlook and introduction of upcoming new features like the “Vehicle Routing Problem” (VRP) solver, and what we have in mind for future releases.

We will explain the shortest path search in real road networks and how the data structure is important to get better routing results. Furthermore we will show how you can improve the quality of the search with dynamic costs and make the result look closer to the reality. You will also learn about difficulties and limitations of the library, and when pgRouting might not be not the right tool to solve your routing problem.

References

- ✔ pgRouting Project - Open Source Routing Library (http://pgrouting.org/)
- ✔ pgRouting - Routing on PostgreSQL (https://github.com/pgrouting/pgrouting)

How Open Source and INSPIRE can be used as a tactical weapon for economical growth in Europe

Dirk Frigne[1], Torsten Friebe[2], Giacome Martirano[3]

[1] geomajas

[2] deegree

[3] SMESpire

Abstract

During this presentation, Dirk Frigne (founder and spiritual father of the geomajas project - www.geomajas.org) will combine observations and lessons learned from different projects to point out some ideas and possibilities how to launch an ambitious initiative where Europe and here SME's can form a strong player on the global market to offer solutions for business needs world wide and to support tackling the ever increasing environmental challenges.

References

- ✔ http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/index_en.htm
- ✔ http://www.smespire.eu
- ✔ https://joinup.ec.europa.eu/community/are3na/home
- ✔ http://en.wikipedia.org/wiki/Comparison_of_free_and_open-source_software_licenses#cite_note-gnucopyleft-9
- ✔ https://www.gnu.org/licenses/agpl-3.0.html
- ✔ http://www.bsa.org/country/Public%20Policy/~/media/Files/Policy/Security/General/sw_factsfigures.ashx

Sunlumo – open source spatial data management platform

Dražen Odobašić[1]

[1] Candela-IT

Abstract

Sunlumo (Candela-IT, 2015) is an open source spatial data management platform built on top of the QGIS application framework. The platform is defined by a QGIS Project file and provides services to various user groups. For example, an expert user could use the QGIS Desktop for any purpose, a basic user will use the same project with a customized QGIS Desktop with limited basic user functionality, and an anonymous (web user) would use a Sunlumo WebGIS interface to browse and query spatial data.

The Sunlumo platform also provides services, like SimilaritySearch, which are reusable for any purpose. SimilaritySearch provides a searchable index for any QGIS vector layer and enables 'equals', 'startswith' and 'contains' operators. An index is constructed per layer for a specific combination of that layer's attributes.

Sunlumo platform is based on QGIS Python API, Django framework and PostgreSQL/PostGIS database. Current development is mainly focused on reimplementing the QGIS Server and QGIS WebClient on Django framework by utilizing QGIS Python API.

The backend is built on a standard Django framework with unit test. The frontend development uses modern Javascript/CSS frameworks like OpenLayers3 and Mithril.JS backed by a Node.js toolchain.

References

Candela-IT. (2015). Sunlumo - an open source spatial data management platform. Retrieved March 1, 2015, from https://github.com/candela-it/sunlumo

Implementing INSPIRE with HALE

Simon Templer[1], Thorsten Reitz[2]

[1] wetransform GmbH, Germany

[2] data harmonisation panel, Germany

Abstract

The INSPIRE directive aims at bringing an interoperable spatial data infrastructure to Europe. This is an opportunity for better cooperation across regions and countries all over Europe, but also a challenge for data providers to implement. Data providers have the obligation to provide data compliant to the INSPIRE data specifications, served via download and view services.

With the HUMBOLDT Alignment Editor (HALE), users create transformation mappings between different data models. A common application of HALE is to transform data from a custom data model to an INSPIRE application schema. Once defined, the mapping can be reused for all data that adheres to the source schema, independent of the format. The declarative representation of the mapping allows for easy maintenance, e.g. in case of changes to the data models.

HALE is designed specifically to deal with the challenges of complex data models like the INSPIRE application schemas, where there are Feature Types with complex sub-structures and recurring attributes. For a quick start working on a mapping to INSPIRE, HALE isbundled with template projects and integration with the INSPIRE registry for code lists. HALE is not only a tool that can be used for INSPIRE though, it provides generic support for XML schemas and GML application schemas, and thus can be used with formats like GeoSciML, XPlanGML or CityGML.

For creating the actual schema mapping, users work with the source and target models in conjuntion with a set of sample source data. Every time the mapping is updated, users can directly see how the transformation is applied and what the results will look like – in map and table views.

The transformation mapping is done step-by-step, connecting source and target model type and attribute elements via relations. For each step, transformation and validation are performed, and users can inspect and analyze source and transformed data.

To provide INSPIRE download services, the transformation can be set up in multiple ways, the following are two commonly used possibilities:

- create an *Atom-based pre-defined data set download service*
- Supply data via a *WFS-T interface* to compliant software like deegree.

Integrating HALE with existing infrastructure is made easy by a command line interface and the capability to deploy HALE as a service with web frontend and REST API.

As a technology, HALE is based on Java, the user interface is based on Eclipse RCP. The modular architecture provides a large number of extension points for adding functionality, such as additional data or model formats, custom transformation functions and UI components. Advanced users often work with the powerful Groovy scripting functions as part of the transformations.

HALE is Open Source software licensed under the GNU Lesser General Public License (LGPL) v3.0 and welcomes new users and developers to its community.

Push it through the wire!

Jachym Cepicky[1]

[1] Geosense s.r.o.

Abstract

Today's web browsers, their renderers and JavaScript interpreters are able to display relatively big amounts of vector data. Moving from DOM rendering to Canvas (and further to WebGL) enables us to display thousands of vector features. With this possibility, new problem has to be solved: how to send such amount of data through limited internet connection? Presentation will demonstrate some possible solutions for this problem, using tiled vectors, generalization, TopoJSON format and other.

School on the Cloud – implementing Cloud-based FOSS4G technologies in education to support digital citizenship development

Elżbieta Wołoszyńska-Wiśniewska[1], Panos Papoutsis[2], Karl Donert[3], Luc Zwartjes[4], Kostis Koutsopoulos[5], Sofie De Cupere[6]

[1] UNEP/GRID-Warsaw Centre (Poland)

[2] Doukas School S.A. (Greece)

[3] Innovative Learning Network Ltd. (United Kingdom)

[4] Geography Department, Ghent University (Belgium)

[5] European Association of Geographers (EUROGEO) (Belgium)

[6] GO! Education of the Flemish Community (Belgium)

Abstract

Since the beginning of the 21st century cloud computing technologies have been entering into different sectors of our economy, including education (Johnson, 2012). Finally, in 2012 the first education Cloud has arrived in Europe. 1,200 schools and more than 350,000 teachers and students in Northern Ireland has become involved in a new dynamic way to educate that aligns with the way we nowadays think, share, learn and collaborate (Meyer, 2012).

From that time on, Cloud-based solutions are being more and more adopted in education. Easy access to high-speed networks along with open standards, content and technologies are encouraging individuals to use ICT and empower education process. As learning becomes increasingly digital, online access becomes the necessary vehicle for the emerging Cloud-based developments (Donert & Bonanou, 2014). Mobile technologies enable the integration of inquiry project support into everyday life situations of learners, providing the ability to learn anytime and anywhere. Mircea and Andreescu (2011) confirm the Cloud allows teachers and learners to access real time information from anywhere in the world in a matter of seconds through applications (apps) and other useful tools for free. Flexibility, as well as low cost, is compelling advantage of the Cloud. Therefore, many schools and educational organisations are considering migrating their activities to the Cloud.

UNEP/GRID-Warsaw Centre, the institutional affiliation of the main author, is an NGO, exploring the potential of the Cloud in education, with particular focus on FOSS4G applications. As Wołoszyńska and Rusztecka (2012) indicates modern society is depending increasingly on spatial information. FOSS4G solutions has become a way to foster development of responsible digital spatial citizenship – an ability to participate in societal spatial decision making through the reflexive use of geoinformation – by simplifying the use of geospatial technology and access to open geodata. However, Koutsopoulos and Kotsanis (2014) claim that

there is a need to develop a new approach to education based on cloud computing, taking simultaneously into account, not only technological aspect but also new role of teachers and students. Therefore, the Centre cooperates in both: the network of the ICA-OsGeo Labs gathered under umbrella of Geo4All initiative (http://www.geoforall.org/) as well as “School on the Cloud” education network, SoC with 57 partners in 18 countries (http://schoolonthecloud.eu/).

Main goal of the ICA-OsGeo Lab in UNEP/GRID-Warsaw Centre is to widely promote implementation of FOSS4G solutions in- and outside the school. On the other hand, SoC network focuses on understanding the changing roles that all actors involved in educational process play, gathering partners in 4 work groups:

- i-Manager (sharing experiences of issues related to leadership and management);
- i-Teacher (learning and teaching issues connected with Cloud-based learning, examination of barriers and key competences required);
- i-Learner (exploration of the opportunities of Cloud-based personalized learning 'at any time, any place, by any one');
- i-Future (dealing with topics like the impact of open (education) resources, the availability of free and available information (open data), new generation tools for the Cloud, etc.).

References

✔ Donert, K., & Bonanou, H. (2014). Education on the Cloud 2014: State of the Art. Retrieved 3 March, 2015, from http://www.schoolonthecloud.eu

✔ Johnson, K. (2012, February 29). Education in the Cloud: Exploring Opportunities. Message posted to http://www.wired.com/2012/02/education-in-the-cloud/

✔ Koutsopoulos, C. K., & Kotsanis, Y. C. (2014). School on Cloud: Towards a paradigm shift. *Themes in Science and Technology Education*, 7(1), 47-62. Retrieved from: http://earthlab.uoi.gr/theste/index.php/theste/article/view/147/96

✔ Meyer, L. (2012, December 4). Europe's First Education Cloud to be Implemented for all Northern Ireland Primary Schools. Message posted to http://thejournal.com/articles/2012/04/12/europes-first-education-cloud-to-be-implemented-for-all-northern-ireland-primary-schools.aspx

✔ Wołoszyńska, E., & Rusztecka, M. (2012). GI Technologies in Secondary School Education as the Key to Responsible Spatial Citizenship. In T. Jekel, A. Car, J. Strobl & G. Griesebner (Eds.), *GI_Forum 2012: Geovisualization, Society and Learning* (pp. 302-308. Berlin, Germany: Wichmann.

✔ Mircea, M. & Andreescu, A. I. (2011). Using Cloud Computing in Higher Education: A Strategy to Improve Agility in the Current Financial Crisis. *Communications of the IBIMA (CIBIMA)*, 2011(2011). doi: 10.5171/2011.875547

Flow analysis using sUAS and lidar data

Helena Mitasova[1], Justyna Jeziorska[1,2] , Anna Petrasova[1], Vaclav Petras[1], Thomas Zajkowski[1]

[1] North Carolina State University (NCSU), Raleigh, NC, USA,

[2] Wroclaw University, Wroclaw, Poland

Abstract

Topographic mapping using small Unmanned Aerial Systems (sUAS) is becoming a mature technology providing extremely high resolution 3D data with rapidly growing range of applications. High efficiency and low cost of data acquision make it ideal for repeated surveys of eroding landscapes (d'Oleire-Oltmanns eta al., 2012) and for precision land management including control of surface runoff (Barreiro et al., 2014). We present an application of GRASS GIS (Neteler et al., 2012) for sub-meter resolution analysis and modeling of overland flow patterns on agricultural fields using Digital Surface Model (DSM) derived from sUAS-imagery and lidar-based DEM. The sUAS data were collected by NextGen Air Transportation (NGAT) Program at NCSU using Trimble UX5 aircraft. Orthophoto and DSM were derived at 0.04 m and 0.15 m resolutions respectively. The lidar data were collected by Leica ALS-70 aerial LiDAR sensor system for the entire county and the DEM was interpolated at 0.30 m resolution. The study area includes tilled fields, ditches and unpaved roads where understanding of surface flow patterns is imporant for maintaining access to the fields and reducing pollutant transport into the neighboring streams.

We compare the properties of surfaces captured by lidar and sUAS and investigate techniques to combine the lidar and sUAS data to incorporate the parts of watershed outside the sUAS mapping area while preserving landscape features critical for flow modeling. The surface flow analysis was performed in GRASS GIS using two flow routing methods: a least cost path, multiple flow direction technique and a path sampling method for solution of bivariate overland flow equation (Mitas & Mitasova, 1998). We discuss the handling of microtopography by these methods and their suitability for different aspects of surface flow analysis. As expected, the least cost path method results were most useful for capturing channels and deriving watershed boundaries. The path sampling method provided dynamic representation of water depth distribution, including pooling of water in tilled rows and in depressions on the service roads.

Weekly mapping by sUAS is planned starting in spring 2015, providing multitemporal imagery and elevation data to support monitoring of crop growth and surface processes. The GRASS GIS temporal framework will be used to manage the resulting time series. Special focus will be on development of workflows to use the collected sUAS data to improve our predictive capabilities to support precision runoff and erosion control in the fields and along service roads.

References

- ✔ Barreiro, A., Domínguez, J.M., Crespo, A.J., González-Jorge, H., Roca, D., & Gómez-Gesteir M. (2014). Integration of UAV Photogrammetry and SPH Modelling of Fluids to Study Runoff on Real Terrains. PLoS ONE 9(11): e111031.
- ✔ d'Oleire-Oltmanns, S., Marzolff, I., Peter, K. D., & Ries, J. B. (2012). UAV for monitoring soil erosion in Morocco. Remote Sensing, 4(11), 3390-3416.
- ✔ Mitas, L., & Mitasova, H., (1998). Distributed erosion modeling for effective erosion prevention. Water Resources Research, 34(3), 505-516.
- ✔ Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. Environmental Modelling & Software, 31, 124-130.

3D virtual representation of drones' flights through Cesium.js and Oculus Rift

Matteo Di Paolantonio, Cesar González Fernández, Manuel José Latorre, Felix Pedrera García, Abraham Roldán París

Elecnor Deimos, Ronda de Poniente, 19, 28760 Tres Cantos, Madrid, felix.pedrera@deimos-space.com

Abstract

This work describes the representation of drone's flights in web browsers using 3D globes technology Cesium.js and virtual reality headset Oculus Rift. An interactive visualization is performed using telemetry data, digital elevation models, images and mosaics. These data are processed by open source software and served through OGC web services. This allow users to virtually experience the flight and to visualize in an interactive way through the immersive interface offered by Oculus Rift.

GeoGigTools: OpenLayers plugin for versioned geodata

Francesco Bartoli[1]

[1] Geobeyond Srl

Abstract

GeoGig is a new fashionable tool to handle the history of spatial data in the same way as we maintain the versioning of source code. Inspired by the workflow of a GIT source control management, GeoGig, formerly known as GeoGit, is an open source project developed and mantained by Boundless Spatial to build a distributed system for the versioning of spatial data. It is a J2EE server that keeps trace of edits to datasets and their related geospatial features. This time machine stores all the commits for a certain dataset into a repository either when the geometry and/or the alfanumeric piece of a feature has changed. GeoGig repositories can be local or remote, synchronizable through web API by using push and pull operations, supporting the load of data from many spatial database systems and formats with a powerful CLI. GeoGig has been also well-integrated as GeoServer datastore to publishing OGC WxS interfaces and thus any WFS-T client can automatically perform such edits from a map. Once an Openlayers webgis client is able to interact with layers from a GeoGig repository then it is possible to retrieving all the versions for a layer, all the features for a version and the Map object for such a version. This need is the challenge from which GeoGigTools comes, a lightweight OpenLayers extension that supports map visualisation along the timeline of any versioned layer. The library has been equipped with few but useful functions: the recordset list of all versioned layers with name, WMS url and BBOX; all available versions for a determined layer including information associated to each commit; all features of a determined layer with their information at a specified commit in the past; a layer at a specified commit ready to be passed to an OpenLayers map object for mapping a past version in a WebGIS; the classic getFeatureInfo adapted for the current visualized version of a selectable geometry in a map.

This library has been used proficiently in a WebGIS for implementing a spatial versioned grid system that divides airports into operational areas where tickets are assigned to workers in the field.

References

- ✔ GeoGig formerly GeoGit – http://geogig.org
- ✔ GeoServer – http://geoserver.org
- ✔ OpenLayers – http://openlayers.org
- ✔ GeoGigTools code repository - https://github.com/geobeyond/GeoGigTools

mapmap.js: A High-Level Client-Side API for Interactive Thematic Maps

Florian Ledermann[1]

[1] Technische Universität Wien, Vienna, Austria

Abstract

Creating interactive thematic maps using web technology is still a complex task. Web mapping APIs can help us by hiding some of the details of the technical platform and offering high-level building blocks - existing open source solutions like the Leaflet API (Agafonkin, 2010) made available a greatly simplified way to display tile-based maps in browsers. From a cartographer's perspective, such tile-based mapping APIs offer little excitement by themselves, as the cartographic quality lies mostly in the tiles and the tools used to produce them (on the server), and not in the methods provided by the API. Furthermore, tile-based maps are not always well suited for fully interactive & animated mapping applications.

In recent years, interest in the possibility to render maps in the web client entirely, instead of downloading pre-rendered tiles, has increased, and solutions that render interactive maps using web technologies like Canvas, SVG or WebGL are currently actively investigated. Low-Level APIs like D3.js (Bostock, Ogievetsky, & Heer, 2011), which is not specifically a web mapping API but rather an abstraction of how to interact with document content and connect it to data, have been successfully used to create interactive maps and other spatial information visualizations. However, as cartographers, we primarily deal with map elements, symbolization, metadata and user interaction, but usually when creating web maps we have to think a lot more about technical details like asynchronous requests, event callbacks or the document object model. The necessity to cater for all those technical aspects to be able to produce a state-of-the art online map is a profound obstacle in creating and teaching interactive web maps, and obscures cartographic decisions and processes in the source code.

With mapmap.js[1], we present our approach to implementing a high-level, open source API for creating interactive thematic maps in the browser. mapmap.js supports geometry, data and metadata loading and transformation, selection and symbolization of map elements as well as specifying interaction through built-in or user-provided methods. Internally, mapmap.js uses state-of-the-art JavaScript features like promise chains or functional inheritance to allow us to expose a clear, simple and concise API to the user. Despite its simplicity, the mapmap.js API aims for transparency, meaning simple things should be simple,

1 https://github.com/floledermann/mapmap.js

but every aspect of the internal implementation should be open to modification or the provision of an alternative implementation by the user.

Figure 1 shows an example of a complete implementation of an interactive choropleth map, showing population data loaded from a csv file.

```
var map = mapmap('#mapEl')
  .geometry('austria_admin.topojson', 'iso')
  .data('unemployment_AT_2014.csv', 'code')
  .meta({
    'unemploym':{
      domain: [0,0.15],
      colors: colorbrewer.YlOrRd[5],
      numberFormat:'%',
      label: "Unemployment Rate"
  }})
  .identify(['iso','name'])
  .choropleth('unemploym')
  .hoverInfo(['name','unemploym'])
  .highlight('Vienna')
  .on('click', mapmap.zoom());
```

Figure 1 - An interactive choropleth map, implemented using the mapmap.js API

The listing in Figure 1 shows only a trivial example - mapmap.js comes with sensible defaults, but every line can be further modified by adding parameters or injecting user-provided functions, and there are many more methods in the API. Mapmap.js is implemented using SVG (Ferraiolo & ed., 2001) and D3 internally, which means generated maps can be styled using CSS and extensions can be implemented using D3. The API works with imperfect data, comes with powerful data transformation tools, and has been successfully used with messy real-world datasets without the need for external preprocessing.

References

- ✔ Agafonkin, V. (2010). Leaflet — an open-source JavaScript library for interactive maps. Retrieved March 9, 2015, from http://leafletjs.com/
- ✔ Bostock, M., Ogievetsky, V., & Heer, J. (2011). D3: Data-Driven Documents. IEEE Transactions on Visualization and Computer Graphics, 17(12), 2301–2309. http://doi.org/10.1109/TVCG.2011.185
- ✔ Ferraiolo, J., & ed. (2001). Scalable Vector Graphics (SVG) 1.0 Specification.

SATELLITE SNOW COVER PRODUCTS EVALUATION AND VALIDATION PLATFORM DEVELOPED ENTIRELY WITH FLOSS SOFTWARE

Vasile Crăciunescu, Anișoara Irimescu, Gheorghe Stăncălie, Simona Catană

Romanian National Meteorological Administration

Abstract

The monitoring of snow cover extent is important for the management of natural resource, extreme events prediction such as snowmelt floods, avalanches etc. The current status is that the network of weather stations is too sparse in regions with seasonal snow cover to provide reliable snow monitoring and impact applications. Remote sensing can regularly provide maps of snow cover extent, under limitations imposed by satellite cycles or cloud cover. A number of daily or synthesis snow cover extent products, covering Romania, with different resolutions and specifications, are available for free (e.g. GLOBSNOW, CryoLand, H-SAF, IMS). These products were homogenized and included, along with reference and in-situ data, into an application that make possible for user to inspect, process, analyze and validate the information, using a web based interface. The platform, created by National Meteorological Administration of Romania offers services based on Open Geospatial Consortium standards for data retrieval (WMS, WCS, WFS) and server-side processing (WPS, WCPS). The services were built upon open source solutions such as GeoServer, OpenLayers, GeoExt, PostgreSQL, GDAL, rasdaman. The application is composed of several software modules/services. The modules are split into two categories: server-side modules/services and client side modules - responsible for interaction with the user. A typical usage scenario assumes the following steps:

1. The user is operating the client functionality to select a temporal and spatial slice from a product cube (e.g. 5 months archive of daily CryoLand FSC data);
2. The users select a statistic method to be applied;
3. The request is sent to the server side processing applications wrapped as WPS or WCPS calls;
4. The process will trim/slice the coverage cube, perform the statistic operation for the pixels within the ROI for each day in the selected time interval;
5. The results are sent back encoded in a standard file format;
6. The web client display the results in a relevant form.

LiDAR data as a tool for forest fuel classification

Alfonso Alonso-Benito[1], Alejandro Lorenzo-Gil[1], Manuel Arbelo[1], Pedro A. Hernandez-Leal[1], Jose R. Garcia-Lazaro[2], Jose A. Moreno-Ruiz[2]

[1] Departamento de Fisica, Universidad de La Laguna, La Laguna, Spain

[2] Departamento de Informatica, Universidad de Almeria, Almeria, Spain

Abstract

The updating of fuel maps represents a key element in forest fires management. On this sense, new technologies based on Remote sensing data provide a valuable improvement in fuels characterization. Both multispectral satellite images and active sensors like the Light Detection and Ranging (LiDAR), have been used previously with differrent successs (Alonso-Benito, A., 2013). In this work only airborne LiDAR data were used to map fuel types in a forested area of Tenerife (the Canary Islands, Spain) to evaluate its potential to be used as an updating tool. The LiDAR data, (provided by GRAFCAN) are discrete and unclassified, with an average density near to 2 points/m^2. A field campaign was made through 82 circular plots during 2011 summer season. Fuel types were identified according to the PROMETHEUS project based on the vertical distribution of the vegetation cover. Using PAM (Partitioning Around Medoids) method included in R package 'cluster' (R Core Team, 2013), a cluster analysis of the cloud LiDAR points within the field plots was carried out. Once the forest vertical structure was characterized, a classification of fuels using a decission tree was performed. The results show a global accuracy of a 82%. This reliability was assessed by analyzing the error due to both quantity and allocation disagreement, which resulted in a 12% error in quantity and a 6% in allocation. Results clearly remark the usefulness of these data for the fuels mapping, despite the low density of LiDAR points on the flights provided.

Acknowledgents

This work was supported by the Ministerio de Economia y Competitividad (Spain) under Grant CGL2013-48202-C2.

References

Alonso-Benito, A., Arroyo L.A., Arbelo, M., Hernández-Leal, P., González-Calvo, A. (2013). Pixel and object-based classification approaches for mapping forest fuel types in Tenerife Island from ASTER data. *International Journal of Wildland Fire*, 22, 3, 306-317. doi: 10.1071/WF11068.

R Core Team, 2013, R; A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. Available online: http://www.R-project.org/ (accessed on september 2013).

Integrating FOSS4G into an enterprise system for labour management

Daniel Urdă[1], Florin Iosub[1]

[1] TeamNet

Abstract

In order to help the Romanian Labour Inspection Authority efficiently assume its responsibilities, our company has developed an integrated system that provides the Authority with the various tools required for the inspection and control of companies active in Romania, ensuring their compliance with the local labour legislation. One of these tools was dedicated to helping field inspectors search for and identify the companies' offices, where their presence is required. Naturally, this application required a GIS component, providing off-site location services and an intuitive user interface customized for tablets. We chose to implement this component by using OSS, more specifically building upon existing projects such as OpenGeoSuite and the OpenSource Routing Machine, and leveraging open data provided by OpenStreetMap. Thus, GeoServer and Geowebcache were used to serve an optimized local copy of OSM data, as well as other data relevant to the business scenario. The custom-built user interface was based on GeoExplorer, extended to provide routing (using an OSRM backend) and geocoding (using OSM's Nominatim service). Furthermore, OSM data was used, along data collected in the field by our company, as input for a custom geocoding engine that allowed us to locate over 300,000 companies and their offices throughout Bucharest.

Sensor Widgets: Configurable graphical components for your SOS data

Oscar Fonts[1], Martí Pericay[2]

[1] Freelance (oscar.fonts@geomati.co)

[2] Freelance (marti@pericay.com)

Abstract

Despite of the efforts in the last two years, one of the limiting factors for the adoption of the Sensor Observation Service (SOS) OGC standard the lack of attractive visualization tools, or clients.

The classical SOS client provides a complex user interface, in the form of a full-screen control panel, which exposes all the standard's internal structure and concepts to the final user, such as "Offering", "procedure", "feature of interest" or "observable property", difficult to understand for those not familiar with it. If the user isn't aware of the data actually available in a given service, it is possible that discovering and visualizing it in a proper way can be challenging for a newcomer. The kinds of visualization and interaction with data are also limited to a few predefined ones (typically a time chart, a raw data table and a feature map), which, despite being generic, are not best suited for some use cases. These visualizations are not easily adaptable or reusable in other contexts.

Moreover, due to the excessively verbose SOS data messages (GetObservation XML responses being much redundant), data fetching and parsing can be slow. Hopefully, the SOS version 2.0 standard, and its 52n implementation (the SOS v.4.0 server), provide mechanisms to deliver data in a more web-friendly (JSON format) and compact way, which client implementations could benefit from.

Trying to overcome the aforementioned limitations, and taking profit of the new standard and implementation versions, we developed the "Sensor Widgets", consisting of an extremely thin Javascript SOS client implementation, and an easily configurable and extensible collection of graphical components providing the interactive visualization for sensor data. A Wizard can be used to interactively configure the widgets and generate the code to embed them in other web pages.

Each widget can provide a multi-purpose generic view, or specialize in a particular data magnitude for a specific use case. So, apart from the typical "Timechart", "Table" and "Map" widgets, there are some others such as a "compass" widget to visualize instantaneous wind direction, or a more complex "wind rose" widget can which will display aggregated wind regimes (direction and speed) over a period of time. Other provided Widgets are an instant measure "Panel", a "Thermometer", a "Progressbar" and a "Gauge".

The widgets are very lightweight (about 50-100 lines of code) and designed to be extended and reused, so understanding their mechanics to create a new widget should be very affordable. All the code is released under the MIT license and comes with examples and documentation.

References

http://sensors.fonts.cat
https://github.com/oscarfonts/sensor-widgets

How to get your entire university excited about (open) spatial data and tools

Govert Schoof, Leon van der Meulen

Geodienst, Research and Innovation Support, Centre for information technology, University of Groningen, The Netherlands

Abstract

FOSS4G tools and applications are improving faster and faster, but at universities students, staff and researchers often stick to the traditional proprietary tools.

We are a geo technology startup within the University’s centre for information technology, focusing on stimulating the use of spatial data and analysis. Supporting staff and students with a growing degree of FOSS4G tools and open data and.. a mindset to match. (Leeuwarder Courant, 2013) (Welcome to Geo Services!, 2014).

We as geo-geeks are convinced that using spatial data really enhances research and projects. There is nearly always a spatial component and you should use it! By giving pitch talks, workshops, lectures, organizing knowledge cafés and visiting scholars we spread the idea and help people enhance their work. We are very active at half of our faculties, other universities and companies and institutions. We’ve grown from around 20-40 calls a month to well over a 100 in the past year. We helped researchers publish over 10 scientific publications (GIS analysis, data acquisition and mapmaking) and are working on some seriously innovative projects such a disaster risk management and high performance calculations for spatio-economic correlations. Since this is only the first year we exist, you could argue we are promoting the use of GIS pretty well.

We love open source and are using open source tools more and more. Our department of High Performance computing has great 10k core linux clusters and fast mass storage which we use for big data projects. Since last summer we offer support for Qgis and made the application available on every university computer for students and staff.

Open Data has become a hot item now, but we have been working with it for quite a while. Projects with the Dutch elevation model, Buildings, statistical data are all based on open data.

Currently we are developing a web platform that aims to bring people companies together based on their location and knowledge or expertise. How nice would it be for example to see if one of your neighbor companies has experience with export to a certain country? This is 100% open source based and for a great part also on open data!

References

- ✔ Leeuwarder Courant. (2013, september). High tech & innovatie Special. Leeuwarder Courant.
- ✔ Welcome to Geo Services! (2014, August 27). [Motion Picture]. University of Groningen. Retrieved from Unifocus: http://www.unifocus.nl/site/pagina.php?id_item=445

Mapbender3

Emde Astrid[1]

[1] WhereGroup, Mapbender Team

Abstract

Mapbender3 is a client framework for spatial data infrastructures. It provides web based interfaces for displaying, navigating and interacting with OGC compliant services.

Mapbender3 has a modern and user-friendly administration web interface to do all the work without writing a single line of code. Mapbender3 helps you to set up a repository for your OWS Services and to create indivdual application for different user needs.

The software is is based on the PHP framework Symfony2 and integrates OpenLayers and MapQuery. The Mapbender3 framework provides authentication and authorization services, OWS Proxy functionality, management interfaces for user, group and service administration.

In the presentation we will have a look at some Mapbender3 solutions and find out how powerful Mapbender3 is! You will see how easy it is to publish your own application.

References

- ✔ http://mapbender3.org/en/gallery
- ✔ http://demo.mapbender3.org

QGIS Graphical Modeler on topographic beach surveys

Lucas Terres de Lima[1], Cristina Bernardes[1], Paulo Baptista[1]

[1] Centro de Estudos do Ambiente e do Mar (CESAM), Departamento de Geociências, Universidade de Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal. lucasterres@ua.pt, cbernardes@ua.pt, renato.baganha@ua.pt

Abstract

The *Graphical Modeler,* a tool of *QGIS 2.8.1 Wien* allows to create complex models using a simple and easy-to-use interface. When working with a Geographical Information System (GIS), most analysis operations are not isolated, but rather part of a chain of operations (QGIS, 2014). Using the graphical modeler, that chain of processes can be wrapped into a single process, which makes it easy and convenient to execute – on a different set of inputs – just as a single process. No matter how many steps and different algorithms are involved, a model is executed as a single algorithm, thus saving time and effort (QGIS, 2014). In the present work, it is intended to create a model with *Processing modeler tool* in order to obtain surface images from topographic data, acquired in the aim of a seasonal monitoring program of sandy beaches on the northwest coast of Portugal. The beach topography is was monitored by the INSHORE system, a land based survey system in which Global Positioning System in Differential mode (DGPS) is combined with laser measurements (Baptista et al., 2011). The model created performs the standard procedure to the type of data collected. First, the model asks to add the table with the 3D survey coordinates (X, Y, Z). The table goes through the operation *'Points layer from table'* that turns the table to points in *shapefile* format. After this procedure, the points with the value of elevation simultaneously pass through the algorithm '*Ordinary Krigin (global)*' and '*Fixed distance buffer*'. The '*Ordinary Krigin (global)'* tool creates the interpolated surface elevation of each point and the tool *'Fixed buffer distance'* creates an area spanning ten meters around the points, generating a study limit to the beach. The results of the two operations follow through the '*Clip grid with polygon'* tool, generating the final output with the clipping boundary of the beach. The volume of the surface using the algorithm '*Grid volume*' is also calculated. The model created by *QGIS 2.8.1 Wien* generated good results and prove to be efficient. With this created model, any user can obtain reliable results without the need of specific knowledge about the processes involved. The trend for future projects is to optimize operations with the creation of models similar to this.

References

- ✔QGIS, P. (2014). QGIS User Guide. *Creative Commons Licence*. MA: QGIS Project
- ✔Baptista, P., Bernardes, C. & Cunha, T.R. (2011). The validation analysis of the INSHORE system - a precise and efficient coastal survey system. MA: *Environmental Monitoring And Assessment*. 179, 1-4, 589-604

Sharing software for INSPIRE implementation, use and reuse

Robin S. SMITH[1], Michael LUTZ[1], Leda Bargiotti[2]

[1] EC Joint Research Centre, Ispra (VA), Italy

[2] PwC, Brussels, Belgium

Abstract

1. Introduction

Several activities are supporting public sector organisations to implement the INSPIRE Directive but how can they easily collaborate with software providers to have interoperable and reusable solutions? Similarly, how can we best involve FOSS actors to make INSPIRE fully operational so that, in turn, we can protect the European environment, reuse INSPIRE in other sectors and make best use of its data in mobile apps and other applications? This paper outlines some answers to these questions but also poses others to understand the FOSS community's role at the conference.

2. Creating open source tools in ARE3NA

Action 1.17 of the EU's Interoperability Solutions for European Public Administrations (ISA) Programme: A Reusable INSPIRE Reference Platform (ARE3NA) was presented at the FOSS4G CEE conference in 2013, shortly after its launch. Led by the technical coordinators of INSPIRE, the Joint Research Centre of the European Commission (JRC), the project has been supporting INSPIRE implementation and reuse through open source interoperability solutions, including

- a testbed for access control to spatial data and services (JRC, 2015a)
- the Re3gistry, software to manage data codes and labels, deployed in the INSPIRE Registry (JRC, 2015b)
- A candidate Download Service using OGC Sensor Observation Services (52North, 2014)
- Reusing INSPIRE metadata in open data portals (Goedertier, 2014)

This experience, and the support to the INSPIRE Maintenance and Implementation Group (MIG), has shown benefits and identified opportunities. As more organisations implement INSPIRE, a sustainable approach is needed beyond ARE3NA to support the supply and demand of solutions. Already, FOSS developers are playing an important role in providing INSPIRE-related solutions, as seen through a list of tools drafted by OSGEO members (OSGEO, 2014).

3. Developing platforms for implementation

Our notion of a 'reference platform' includes an arena where implementers find information about software, while developers can showcase their tools or apps/applications using the infrastructure. It should also be a virtual setting for stakeholders to discuss how tools supporting INSPIRE implementation and reuse can be improved or new tools made to fill gaps, and to explore how such developments could be funded. ARE3NA is building this platform incrementally and needs feedback from the FOSS community. Some key questions include:

- Do FOSS actors share INSPIRE core values and consider themselves to be partners?
- Do they know their solutions fit INSPIRE's needs?
- Do they see enough/potential business in INSPIRE?
- What do they see as barriers to collaboration with implementers, and who else could be involved?
- What basic information do (re)users of INSPIRE data need to readily build applications/ apps?

To help this discussion, an initial survey in ARE3NA gathered a list of around 40 topics that stakeholders felt would help INSPIRE tools to function better and 'bridge the gap' between INSPIRE and other sectors. Work is planned for the 2015 INSPIRE conference to explore this topic with implementers (and some software developers). This paper and discussion will allow FOSS actors to provide their views before ARE3NA aims to move collaboration online.

References

JRC (2015a). *ARE3NA study on Authentication, Authorization and Accounting for Data and Services in EU Public Administrations*. Retrieved March, 9, 2015, from https://joinup.ec.europa.eu/asset/are3na-aaa/description

(JRC, 2015b) *INSPIRE Registry.* Retrieved March, 9, 2015, from http://inspire.ec.europa.eu/registry/

52North (2014) *Sensor Observation Service.* Retrieved March, 9, 2015, from http://52north.org/communities/sensorweb/sos/

Goedertier, S (2014) *DCAT application profile for data portals in Europe.* Retrieved March, 9, 2015, from https://joinup.ec.europa.eu/asset/dcat_application_profile/description

OSGEO (2014) *INSPIRE tools inventory*. Retrieved March, 9, 2015, from http://wiki.osgeo.org/wiki/INSPIRE_tools_inventory

ERSAF's audiovisual heritage of historical land use data for EXPO 2015

Marco Cicala[1], Franco Guzzetti[2], Marco Stevan[3], Paolo Viskanic[4]

[1] R3 GIS S.r.l. – Via Kravogl, 2 – 39012 Merano (BZ) – marco.cicala@r3-gis.com

[2] Dip. ABC Politecnico di Milano – Via Ponzio, 31 – 20133 Milano (MI) – franco.guzzetti@polimi.it

[1] R3 GIS S.r.l. – Via Kravogl, 2 – 39012 Merano (BZ) – marco.stevan @r3-gis.com

[1] R3 GIS S.r.l. – Via Kravogl, 2 – 39012 Merano (BZ) – paolo.viskanic@r3-gis.com

Abstract

ERSAF (Ente Regionale per i Servizi all'Agricoltura e alle Foreste) owns a huge database of historical data about land use, agriculture, breeding, forests and protected areas in Lombardy. For EXPO 2015 ERSAF decided to make the data of the last 60 years of changes available to the public in an attractive and simple but at the same time complete and scientifically correct way.
R3 GIS decided to develop a web application for exploring data and data changes in many different ways and using linear and non-linear navigation, through product cards and maps, itineraries and 3D map navigation. The application is based on Free and Open Source components and is published with a FOSS Licence.

1 The data

The first challenge was to get an overview of all of ERSAF's data. Datasets reach from statistical historical data about agricultural and forest land use (DUSAF, collected in 1954, 1980, 1999, 2005, 2007, 2009 and 2012) to typical products (IGP) and DOC, DOCG vine catalogues, to a video library of interviews about agriculture and past mountain pasture lifestyles.
Founding those heterogeneous types of data into an easy-to-use web application was the second challenge.
The basic idea was to split the data into 4 main categories: The Terrain, The Products, The Stories and The Itineraries, each of them having a specific type of data visualization and representation. So while terrain and itineraries are focused in map views (3D and time slider), product data and video interviews are show in a more traditional way, through data cards.

Figure 1: The splash screen of the WebGIS application

2 Tag based navigation

To link this data into a didactic application, a traditional menu driven navigation would have been too restrictive. So a tag based navigation has been developed and added to the menu navigation.
All data inserted into the application's database will be checked and tagged using both, geographical and non geographical tags.

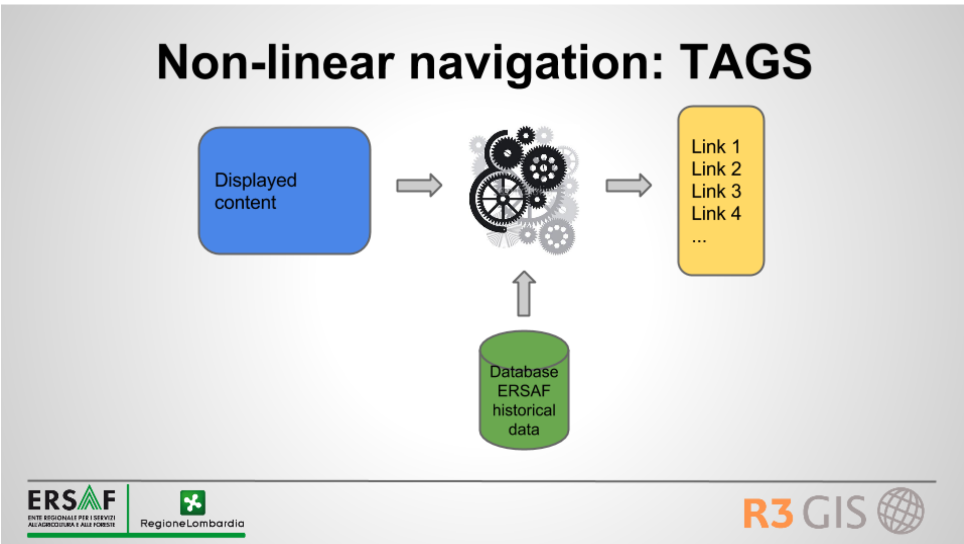


Figure 2: Tag-based navigation schema

Once a content is displayed in the application, a special referencing engine, developed for this needs, looks for similar content checking geographical and mnemonic tags and shows a list of links to related articles.
As a welcome side-result the application can be updated with new information at any time, without the need to modify the code.

3 Time navigation and 3D exploring

To make things even more attractive, some geographical data, the one where changes over the last 60 years were available and interesting, are shown in a basic map window with 4 time-layers and a slider. Moving the time-slider through the years from 1954 to 2012, shows how the territory has changed in the last 60 years.

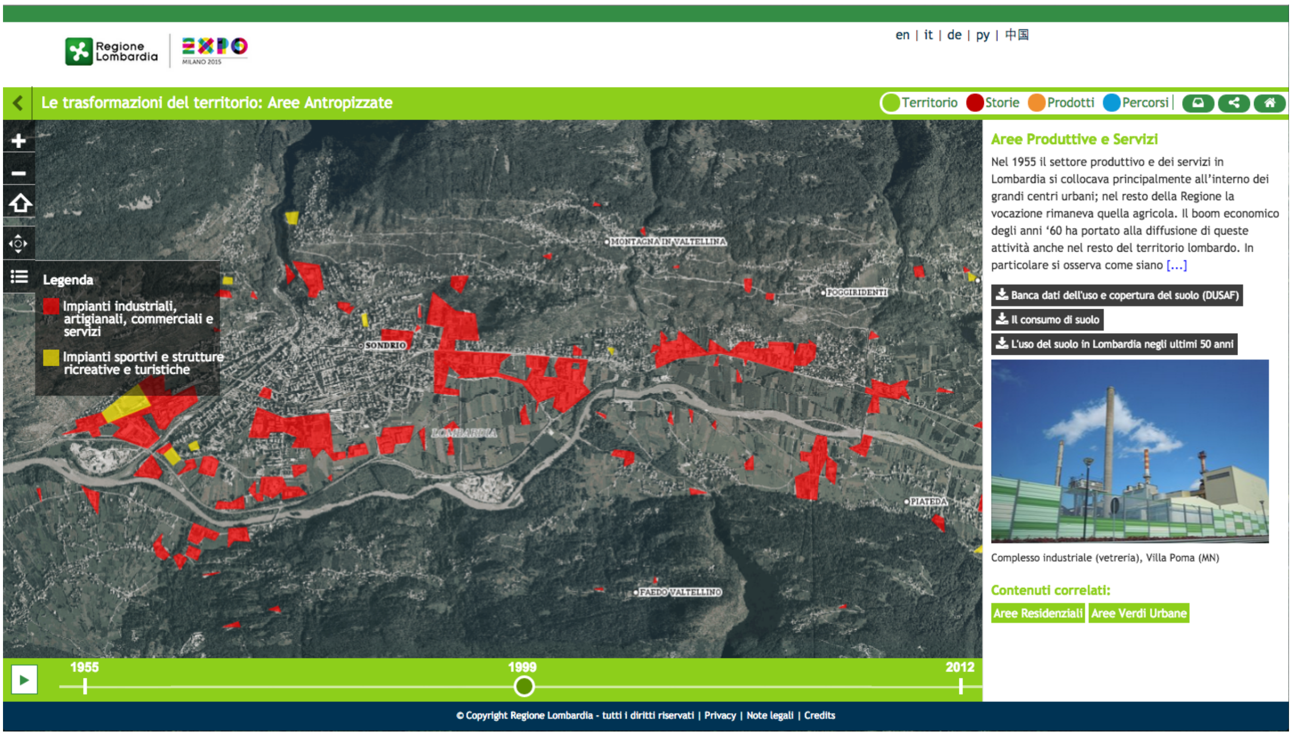


Figure 3: representation of typical aspects of the territory in different periods

Since the territory of Lombardy is composed of plains, hills and mountains, making some geographical data available through 3D terrain navigation was another required feature. Extent and location of distinct vineyards are easier to understand, if seen with their height-component. For this purpose the 3D engine Cesium JS was used.

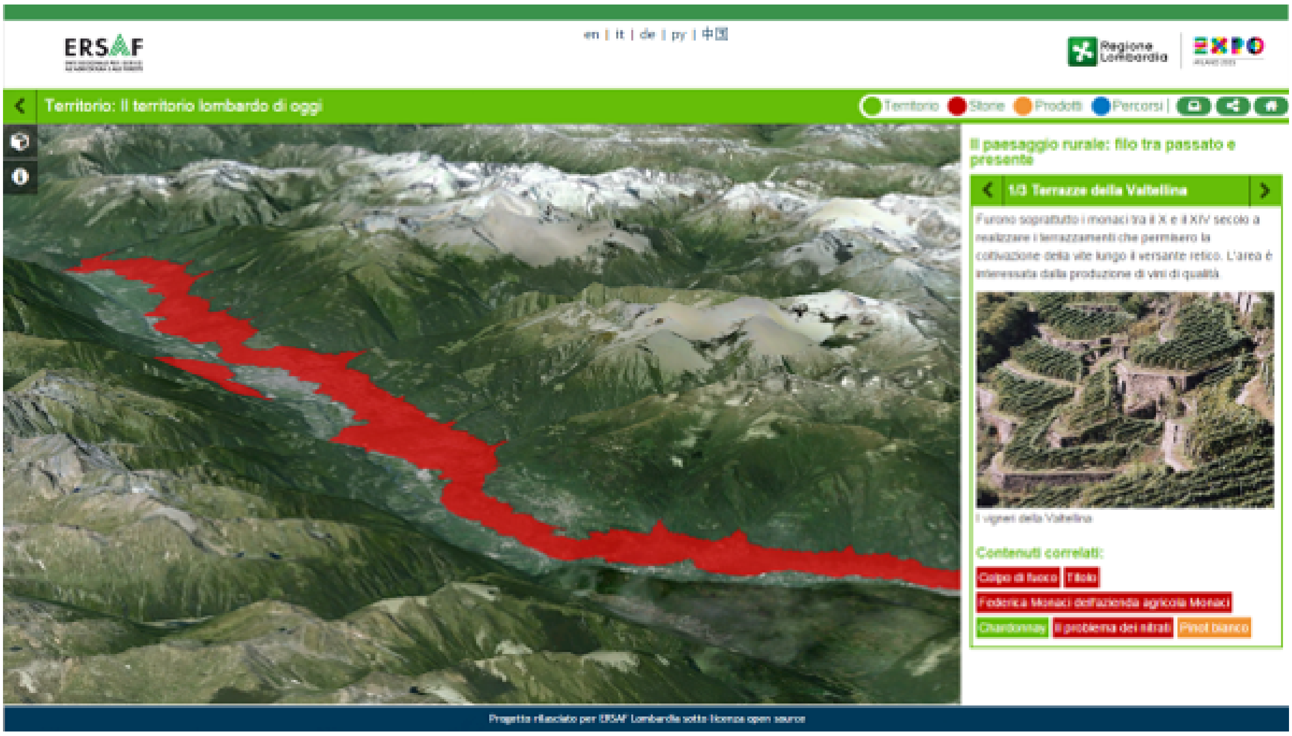


Figure 4: 3D representation of particular aspects using Cesium 3D

4 The technical challenge

The average user of this application is a non technician visitor of Milan's EXPO 2015 who has only limited experience with GIS applications. So the application had to be designed with a user-friendly interface, and with no navigation instructions needed. Moreover the application needs to be usable also with tablets and smartphones and will be used in EXPO 2015 through a totem with touch-screen.

Following it's vocation as an open source company, R3 GIS decided to assemble the most advanced and modern technology available FOSS.

The stack of software on which the application is based are a CentOS operating system, a PostgreSQL/PostGIS Database, Apache Web Server and UMN Mapserver for spatial data rendering.

For 3D navigation the choice fell on CESIUM JS (cesiumjs.org), an open source WebGL Virtual Globe and Map Engine, for both, its advanced technical features and its active open-source community.

Since an open-source OL3-Cesium library is available to switch smoothly between 2D and 3D, OpenLayers 3 is the best candidate as a client for the GIS data.

Since this is very recent technology, the compatibility with all browsers and mobile operating systems is not jet guaranteed, but the community working on Cesium is very active and outlooks are good to have most devices/browsers compatible until summer 2015.

5 Beyond the EXPO2015

In order to keep the application alive also after the EXPO 2015, it has been build to auto-configure itself just out of its underlying database.
To add additional contents and data in the future, a simple data entry system has been developed. Through a user friendly interface any new content can be easily placed into the PostgreSQL/PostGIS database and linked to the correct categories. Tags connect these new contents to the existing data and bind it into the application, making it immediately available.

Advanced Cartographic Rendering in GeoServer

Andrea Aime[1], Simone Giannecchini[2]

[1] andrea.aime@geo-solutions.it, GeoSolutions SAS

[2] simone.giannecchini@geo-solutions.it, GeoSolutions SAS

Abstract

Various software can style maps and generate a proper SLD document for OGC compliant WMS. However, in most occasions, the styling allowed by the graphical tools is pretty limited and not good enough to achieve good looking, readable and efficient cartographic output. For those that like to write their own styles CSS also represents a nice alternatives thanks to its compact-ness and expressiveness.This presentation will provide hints, examples and enough information to master SLD in order to create beatiful maps with GeoServer.

Several topics will be covered, providing examples in both SLD and CSS for each, including: mastering multi-scale styling, using GeoServer extensions to build common hatch patterns, line styling beyond the basics, such as cased lines, controlling symbols along a line and the way they repeat, leveraging TTF symbol fonts and SVGs to generate good looking point thematic maps, using the full power of GeoServer label lay-outing tools to build pleasant, informative maps on both point, polygon and line layers, including adding road plates around labels, leverage the labelling subsystem conflict resolution engine to avoid overlaps in stand alone point symbology, blending charts into a map, dynamically transform data during rendering to get more explicative maps without the need to pre-process a large amount of views.

The presentation aims to provide the attendees with enough information to master SLD/CSS documents and most of GeoServer extensions to generate appealing, informative, readable maps that can be quickly rendered on screen.

Managing MetOc and Remote Sensing data with GeoServer

Simone Giannecchini[1], Andrea Aime[2]

[1] simone.giannecchini@geo-solutions.it, GeoSolutions SAS

[2] andrea.aime@geo-solutions.it, GeoSolutions SAS

Abstract

This presentation will provide detailed information on how to ingest and configure spatio-temporal data in GeoServer, to be served using OGC services, with examples from WMS and WCS services. Topics covered are as follows:

- Discussion over existing data formats and how to preprocess them for best serving with GeoServer
- Configuring SpatioTemporal raster and vector data in GeoServer
- Serving SpatioTemporal raster and vector data with OGC Services
- Tips and techniques to optimize performance and allow maximum exploitation of the available data

The attendees will be provided with the basic knowledge needed to preprocess and ingest the most common spatiotemporal data from the MetOc and Remote Sensing field for serving via GeoServer.

Managing Web GIS applications using the Goemajas Deskmanager

Oliver May[1]

[1] Geomajas developer and PSC member

Abstract

Geomajas is a Web GIS framework that makes it possible to present and manage GIS data from different sources into a web browser. The Geomajas server adds extra functionality such as security, advanced filtering, caching and so on to existing data sources like WFS, postgis etc.

To ease up the configuration the Geomajas team has developed the Deskmanager plug-in. This plug-in consists of a management (web) user interface where GIS maps, so called 'geodesks' can be managed. It is possible to add external data sources, configure default visibility of data layers, apply authorization, represent data layers as a (configurable) layer tree, apply SLD's and so on.

The target users are organizations that want to provide Web GIS maps to their audience, such as local governments or SaaS providers.

During the presentation the current state of the Deskmanager plug-in will be discussed, there will also be a quick peek under the hood. Next to this there will be a demonstration of the Deskmanager management application in action.

Finally the future of the plug-in is discussed, as well as some thoughts on integration with other web GIS clients (such as Openlayers, Leaflet, etc.) and how to manage this using open standards.

Management and publication of geospatial public data using GFOSS: the case study of Vicenza municipality

Roberto Marzocchi[1], Roberta Fagandini[1], Lorenzo Beggiato[2],Eugenio Berti[2], Rosario Ardini[2], Tiziano Cosso[1], Marcello Missagia[2]

[1] Gter srl Innovazione in Geomatica Gnss e Gis, (University of Genoa Spin-Off company) – Via Greto di Cornigliano 6r, 16152 Genova (Italy)

[2] Comune di Vicenza, Settore: Risorse umane, organizzazione e segreteria generale, Corso Palladio,98 - 36100 Vicenza

Abstract

In recent years , the use of Gegraphic Free and Open Source Software (GFOSS) by the Italian public administration is more diffused (CAD, 2013) even if the suspicion that proprietary software is more able to completely solve the problem of the management of a complex geoportal of a big municipality is still present. The SitVi 2.0 (Geographical Information System of Vicenza Municipality version 2.0) is an interesting case study that shown how it is possible to resolve the issue properly using GFOSS.

Vicenza is a town of the North-East of Italy with a resident population more than 110'000 people (from the 2013 Census). The Information System department of Vicenza municipality use since a few years open source software (e.g. MapServer and Openlayers) for the visualization of geospatial data. Since 2014 they have decided to totally migrate to open source software for the visualization of data, the management of associated database system and the integration of geospatial data with cadastrial, census and planning information.

The new SitVi (version 2.0) hence, from the first layer, use PostgreSQL/PostGIS as GeoDatabase; this GeoDBMS is actually a copy of a proprietary database. Nevertheless it ensures the same performance in terms of reliability and a better management of topological information.

As application server Geoserver was preferred to Mapserver due to the presence of a user friendly interface that allows also to technicians of Spatial Information System department, not necessarily computer science experts, to manage and add new layers by itself.

Finally a graphical user interface for visualize geospatial and related data is realized using Openlayers, ExtJS and GeoExt javascript frameworks. Several modules were implemented using also php web programming language for a user-friendly and secure interface with Postgresql DBMS and are integrated with the SitVI 2.0 webGIS platform. Two different version of SitVi were realized; a first one is available for internal users on intranet network allowing access to several information using profiled access with username and password of LDAP system internal municipality system, while a second version of SitVi is available

for public not-profiled users.

References

- ✔ CAD (2013) Codice Amministrazione digitale - Decreto Legislativo n. 82 7 marzo 2005. Last update 11 november 2013 - available online (last access 7march 2015) http://www.agid.gov.it/agenda-digitale/codice-amministrazione-digitale

Publication of wind and wave monitoring and forecasting geospatial data using GFOSS: the "Wind, Ports, and Sea” project

Roberta Fagandini[1], Roberto Marzocchi[1], Tiziano Cosso[1], Massimiliano Burlando[2], Patrizia De Gaetano[2], Marina Pizzo[2], Maria Pia Repetto[2], Marco Tizzi[2], Giovanni Solari[2]

[1] Gter srl Innovazione in Geomatica Gnss e Gis, (University of Genoa Spin-Off company) – Via Greto di Cornigliano 6r, 16152 Genova (Italy)

[2] Dipartimento di Ingegneria Civile Chimica ed ambientale (DICCA), Università degli studi di Genova, Via Montallegro 1, 16145 Genova

Abstract

"Wind, Ports, and Sea” project follows a previous European project, named “Wind and Ports”, that focused on the safety management of port areas with reference to the problem of the wind effects on the land side, through the realisation of an integrated system including in situ wind monitoring, wind climatology evaluations, and the implementation of algorithms for medium- and short-term wind forecast. “Wind, Ports, and Sea” integrates “Wind and Ports”, focusing on the sea side of the port areas, through the acquisition of new in situ wind and wave monitoring devices and the implementation of a wave forecast system coupled to the already existing wind forecast system.

One of the aim of the project is to realize new tools that will provide the Port Authorities and their stakeholders with further information on the wind and sea state conditions. To this reason the webGIS realized for the previously project “Wind and Ports” has been completely updated in order to obtain a new webGIS with better performances and an interface more user friendly.

The old webGIS and also the new one have been totally developed using Gegraphic Free and Open Source Software (GFOSS). Aim of the webGIS is to visualize geospatial data coming from computational algorithms implemented using fortran programming language which produce as output not georeferenced grid of data. The computational results are imported in a PostgreSQL/PostGIS geodatabase (Holl & Plum, 2009) using python scripts that use the Psycopg2 adapter. Using the geodatabase the computational results are related to a geospatial grid in order to create vector grid of data points with wind and wave forecasting. Mapserver has been used as application server, because it supports the used Open Geospatial Consortium (OGC) standards and our purpose is to maintain the same software used in the previous version of the webGIS; while the new client interface has been realized using Openlayers, ExtJS and GeoExt javascript framework and the Heron Mapping Client library. The use of these applications allows to obtain a client interface more user friendly respect of the use of p.mapper, which is the chosen software in the previous “Wind and Ports” webGIS.

All the system has been optimized in order to visualize real-time updated information coming from in situ monitoring and forecasting. A mobile version of the client interface has been also realized in order to be compatible with a mobile application implemented to guarantee a simple access for stakeholders to monitoring and forecasting data.

References

- ✔ Holl S. & Plum H. (2009). PostGIS, powerful Gis function in the database system PostgreSQL. GeoInformatics, 03/2009:34-36
- ✔ OGC, Open Geospatial Consortium. http://www.opengeospatial.org/

DATE OSSIM PlugIn: a new open source tool for digital automatic terrain extraction

Martina Di Rita[1], Andrea Nascetti[1], Mattia Crespi[1]

[1] Geodesy and Geomatics Division – DICEA, "Sapienza" University of Rome, Italy

Abstract

Nowadays high-resolution satellite imagery (HRSI) provides global coverage and allows for accurate and reliable terrain characterization and data extraction. In particular the Digital Surface Model (DSM) is one of the major outputs of the photogrammetric processing for satellite imagery and represents an important component of geospatial database. Consequently, over the years, a wide range of software (commercial and scientific) able to generate DSMs starting from HRSI has been developed.

In the last two decades an increase of free and open source software for geospatial data processing has been witnessed. This kind of software features a source code available to anyone and for any purpose, thus allowing a constant improvement and updating process coming from the developer community.

In the geospatial field OSGeo (Open Source Geospatial) is a leading player: it is a non-profit and non-governmental organization whose mission is to support and promote the collaborative development of open geospatial technologies and data (Mitchell, 2012).

It is in this context and following an open source vision that the present work has been conceived.

In this paper we present *DATE – Digital Automatic Terrain Extractor* – a new Free and Open Source Software for Geospatial (FOSS4G) developed in the framework of the 2014 Google Summer of Code program. *DATE* is conceived as an extension to OSSIM (Open Source Software Image Map), a suite of geospatial libraries and applications for imagery, maps, terrain and vector data processing, belonging to the wide OSGeo family. The tool is a free and open source PlugIn able to perform a totally automatic DSMs generation chain starting from a high resolution satellite stereo-pair. In particular the tool can accommodate images acquired by the most common sensors (e.g GeoEye1, WordlView1-2, Quickbird, Pleiades).

Its peculiarity consists in performing a “hybrid” processing chain based on both rigorous photogrammetric approach and Computer Vision techniques: there is a combination of OSSIM photogrammetric capabilities with the OpenCV library algorithms (Bradski & Kaehler, 2008).

DATE processing chain is essentially based on automatic Tie Points (TPs) identification and Semi-Global Matching disparity map generation (Hirschmüller, 2008) both exploiting OpenCV algorithms implementation.

It consists of some principal steps:

- Stereo-pair projection: the images are projected in a ground geometry using a coarse DSM or a constant-height plane
- Automatic TPs identification, with a subsequent filter to remove outliers
- Quasi-epipolar image generation
- Disparity Map generation, using the Semi-Global-Block Matching (SGBM) algorithm implemented in the OpenCV library
- Disparity to height conversion
- DSM geocoding

Several tests have been performed in two morphologically different areas and the DSM accuracy has been assessed using a high-resolution LiDAR DSM as a reference surface.
The first stereo-pair used for the assessment is a GeoEye1 over the city of Trento (Italy), a morphologically complex terrain: a Root Mean Square Error (RMSE) of approximately 5 meters with respect to the reference LiDAR has been achieved. The second stereo-pair is a GeoEye1 over the centre of Rome (Italy), characterized by narrow streets and high buildings: a RMSE of approximately 8 meters has been reached.

References

✔ Bradski, G., & Kaehler, A. (2008). *Learning OpenCV. Computer Vision with the OpenCV Library.* Sebastopol, CA: O'REILLY.

✔ Hirschmüller, H. (2008). Stereo processing by Semiglobal matching and mutual information. *IEEE transaction on pattern analysis and machine intelligence*, 30(2).

✔ Mitchell, T. (2012). Annual report 2010. *OSGeo Journal*, (9), from http://www.osgeo.org/journal

✔ Zhang, C., & Fraser, C. (2009). An improved approach for DSM generation from high-resolution satellite imagery. *Journal of Spatial Science*, 54(2), 1-13. doi:10.1037/1065- 9293.59.4.286

The OpenQuake Integrated Risk Modelling Toolkit

Paolo Tormene

GEM Foundation

Abstract

The OpenQuake Integrated Risk Modelling Toolkit is an open-source project developed by the GEM Foundation. Its main component is a PyQt4-based plugin for QGIS (https://plugins.qgis.org/plugins/svir/). It provides a GUI for creating and editing socioeconomic indicators and composite indices that represent characteristics within communities that create the potential harm from natural hazard events. The plugin enables users to combine these with estimates of physical earthquake risk (i.e. estimates of human or economic losses) for earthquake risk assessment and communication.

A number of open-source tools cooperate seamlessly through an intuitive workflow. One such tool is the OpenQuake Engine (https://github.com/gem/oq-engine/), which can export the results of physical risk calculations as csv files that can be loaded by the plugin as vector layers. The plugin can connect via https with another tool, the OpenQuake Platform (https://platform.openquake.org/), in order to browse socioeconomic indicators and to obtain a vector layer from the selected indicators, for chosen geographical zones. A spatial aggregation workflow enables the user to aggregate the physical risk variables by the zones for which the socioeconomic indices are defined, incorporating all the information in a single layer. The plugin adds data transformation algorithms to the standard QGIS "field calculator" methods, allowing, e.g., to normalize the available indicators in a mutually consistent way.

A javascript D3 Collapsible Tree chart (http://bl.ocks.org/mbostock/4339083) can be used inside a web browser embedded in a QGIS dialog, in order to calculate composite indices for each geographical zone: a "Risk Index" combining risk variables, a "Social Vulnerability Index" taking into account socioeconomic parameters, and an "Integrated Risk Index" combining the two into the final outcome. The tree represents a "project definition" mapping all the steps necessary to produce the same results afterwards, starting from the data contained in the layer. It can be edited dynamically, assigning a weight to each node and choosing how to combine nodes together (e.g., calculating a weighted sum). At each modification of the tree, the composite indices are automatically re-calculated and graphically displayed, styling the QGIS layer accordingly, so the user can have an immediate feedback of the outcomes produced by changing the project definition tree.

The user can upload their work to the OpenQuake Platform, which produces a Geonode layer along with a description of the tree. Then, the owner of the

project can choose to share it with other platform users, allowing them to exploit all the standard functionalities offered by Geonode.
Through the plugin it's also possible to list the projects available on the OpenQuake Platform, download one, edit it locally, and afterwards possibly upload an updated version to the Platform, engaging the scientific community in an iterative and collaborative process.

IMPACT: Portable GIS Tool for Image Processing and Land Cover Mapping

Dario Simonetti

European Commission, Joint Research Centre, Institute for Environment and Sustainability (IES), Forest Resources and Climate Unit

Abstract

Did you ever try to produce a reliable land cover map from Earth Observation data? How many steps are involved and how many different tools do you need? Did you succeed in a reasonable amount of time?

EO imagery has been identified as a key source of information for mapping and monitoring land cover changes or degradation over time but data extraction, layerstack, radiometric calibration, normalization, mosaicking, automatic classification and segmentation are only some of the pre-processing steps that a ICT user has to undertake to obtain a raw land cover map; lastly, post-processing as class editing, land cover validation and statistics extraction, is assigned to RS experts though out the use of available GIS tools.

The IMPACT tool offers a combination of elements of remote sensing, photo interpretation and processing technologies in a portable and stand-alone GIS environment, allowing non ITC users to easily accomplish all necessary pre-processing steps while giving a fast and user-friendly environment for visual editing and map validation.

Based on open sources technologies, IMPACT makes use of GeoExt, OpenLayers and Mapserver for the graphical user interface and raster/vector rendering; EO data processing is carried out by GDAL and Python. No installation or virtual machines are required and therefore it could be copied on a portable device for easy execution and data sharing. Internet connection is not essential although might offer access to additional information like very high resolution WMS layers from Google, NASA, OpenStreetMap or other sources.

A variety of EO data like Landsat, Spot, RapidEye, DMC could be ingested and converted to top of atmosphere reflectance to be then pre-classified using the built-in pixel based classifier; ultimately, single-date or multi-date segmentation is performed throughout Baazt segmenter libraries (Inpe's TerraAIDA Operators); generated objects are automatically labeled according to class majority derived from the pixel based classification. This last step is essential to offer a preliminary land cover maps (for each input date/epoch), reducing the photo interpreter's verification workload.

IMPACT contains modules for visual identification of forest degradation, ground through collection and spatial statistics all characterize by easy to learn and user friendly interfaces.

This tool has been well received by the biodiversity and conservation communities (e.g. protected areas mapping and monitoring) as well as research institutes in developing countries collaborating with the Joint Research Centre throughout projects like MESA, PacsBio and ReCaREDD.

An open source framework for processing daily satellite images (AVHRR) over last 28 years

Sajid Pareeth[1,2,3], Luca Delucchi[1], Markus Metz[1], Nico Salmaso[2], Markus Neteler[1]

[1] GIS and Remote Sensing unit, Department of Biodiversity and Molecular Ecology, Research and Innovation Centre (CRI), Fondazione Edmund Mach (FEM), Trento, Italy

[2] Limnology and River Ecology unit, Department of Sustainable Agro-Ecosystems and Bioresources, Research and Innovation Centre (CRI), Fondazione Edmund Mach (FEM), Trento, Italy

[3] Department of Biology, Chemistry and Pharmacy, Freie Universität, Berlin, Germany

Abstract

The space-borne sensor Advanced Very High Resolution Radiometer (AVHRR) nowadays offers high resolution data for last 30 years which are publicly available through the National Oceanic and Atmospheric Administration (NOAA) archives. The data are available at level 1B in its original spatial resolution (1.1 km) and known as Local Area Coverage (LAC) products (Kidwell, 1996). The data archive includes images acquired by multiple NOAA polar orbiting satellites – from TIROS-N, NOAA (6 – 19). These data are valuable to regional level studies for assessing long term environmental and climatic changes. Achieving sub-pixel accuracy in image to image geo-correction is crucial in avoiding spurious trends from time series analysis. However, due to on-board malfunctions and instrument decaying especially with the older NOAA satellites, earlier images from AVHRR are often difficult to calibrate and georectify accurately. The main contributors to the errors are the well documented clock drift and attitude shifts due to decaying of the instruments (Kaufmann et al., 2000).
In this study, we are leveraging the power of existing open source spatial tools to establish an automated workflow to correct images for the entire 30 years time period covering the study area of Northern Italy. Concisely, we developed time series of daily calibrated and geo-corrected thermal bands (Brightness Temperatures – BT) from 1986 onwards by integrating multiple open source geospatial tools. We used the Pytroll libraries (www.pytroll.org) to read the original data format, to correct for bow-tie effects, clock drift and attitudinal errors followed by calibration of the AVHRR LAC images (Kaufmann et al., 2000; Khlopenkov & Trishchenko, 2008). We extended the data readers in pytroll to support the LAC data. Sub-pixel level accuracy between temporal images in geo-rectification is accomplished by using SIFT image matching technique implemented with Orfeo Toolbox (Lowe, 2004)(www.orfeo-toolbox.org). Final steps of polynomial rectification, cloud removal, filtering for outliers, and temporal database development we performed in GRASS GIS 7.0 (Neteler, Bowman, Landa & Metz, 2012). The resulting temporal database can be used to

calculate Lake Surface Water Temperature (LSWT) for the sub-alpine lakes in the Northern Italy facilitating ongoing project in monitoring the long term trend in surface temperature variation.

References

- ✔ Kaufmann, R. K., Zhou, L., Knyazikhin, Y., Shabanov, V., Myneni, R. B., & Tucker, C. J. (2000). Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(6), 2584–2597. doi:10.1109/36.885205.
- ✔ Khlopenkov, K. V., & Trishchenko, A. P. (2008). Implementation and Evaluation of Concurrent Gradient Search Method for Reprojection of MODIS Level 1B Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(7), 2016–2027. doi:10.1109/TGRS.2008.916633.
- ✔ Kidwell, K. B. (1996). AVHRR Data Acquisition, Processing and Distribution at NOAA. In G. D'Souza, A. S. Belward, & J.-P. Malingreau (Eds.), *Advances in the Use of NOAA AVHRR Data for Land Applications* (pp. 433–453). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-009-0203-9_18.
- ✔ Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, *60*(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94.
- ✔ Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, *31*, 124–130. doi:10.1016/j.envsoft.2011.11.014.

ODK for public transport

Laura Carcano, Barbara Li Gregni, Annarita Polacchini
ASF Autolinee Srl, Via Asiago 16/18, 22100 COMO

Abstract

ASF Autolinee is the public transport company of Como and its province and annually transports 20 Millions of passengers for a total of 13 millions of kilometers. In view of the change of tariff plan provided by Lombardy regional laws now in force, ASF Autolinee needs to know how its customers use its buses to move inside Como province.
In fact, in public transport, knowing people needs in terms of journeys and behaviors is fundamental in order to plan the transport service and tariffs. ASF Autolinee is searching for tools that facilitate the collection of data related to people movements. Open Data Kit (ODK) seems to be a very useful instrument when the transport company has to carry out market investigations substituting the old paper questionnaires. ODK is in fact a free and open-source set of tools developed by researchers at the University of Washington and supported by a growing community of developers, implementers and users. ODK provides three systems: ODK Build or XLS FORM, to create a questionnaire; ODK Collect, that properly installed on a device (tablet, smartphone) allows the collection of data and their sending to the server; ODK Aggregate, that allows the configuration of a server to contain the collected data. The company is testing the use of ODK within public transport and want to introduce ODK as a new technology that will substitute the previous way of carrying out questionnaires or surveys. ODK provides indeed a useful informatics technology about surveys, (questionnaires and cordon), which are very important element for a transport company planning and management. In particular, the adoption of ODK allows to collect data from a device (tablet or mobile phone) which can be or not connected to the internet network and to send them to a server. Once the data are on the server, they can be extracted as a database and directly analyzed. This method allows to speed up the times of data collection and data transmission and save the times spending to transcribe the collected data on a PC and reduce the errors. In particular the company is testing ODK tools to perform:

- interviews to customers about origin, destination, frequency and modality of their travel;
- cordon surveys, done at some major bus stops, registering for each bus the number of passengers on the bus and those getting on and off;
- studies of the load factor, done registering along a bus journey, the number of people getting on and off the bus at each stop.

In addition, ODK also provides the collection of the GPS position of the device: this is very useful when the collected data have to be combined to geographic data. For example in case of load factor studies, registering the number of customers getting on and of the bus together with the position of the device allows to immediately insert the collected data inside a GIS software and to view them in a referenced way, in order to perform geographic analysis.

Distribution of OSGeo software in the INSPIRE ecosystem

Friebe Torsten, Goerke Sebastian

lat/lon GmbH

Abstract

The Infrastructure for Spatial Information within the European community (INSPIRE) provides the world's biggest spatial data infrastructure. As there is a considerable number of OSGeo projects which are capable to provide services which fulfill the INSPIRE regulations, it is interesting to ask for their distribution within the INSPIRE ecosystem. This talk presents an evaluation based on the registered member state's network services from the INSPIRE registry utilized by the harvester component for the new EU INSPIRE Geoportal.
For this evaluation significant characteristics of the main OSGeo projects which provide OGC webservices were identified and the complete INSPIRE ecosystem was scanned for implementations of these OSGeo packages.
The most important result is a map showing how OSGeo software in general is distributed within the INSPIRE ecosystem and, more specifically, how widespread the different packages are used in the INSPIRE member states.

References

- ✔ INSPIRE registry: http://inspire.ec.europa.eu/registry/
- ✔ OSGeo Foundation: http://osgeo.org

deegree webservices – essential open source for INSPIRE

Goerke Sebastian[1], Friebe Torsten[2]

[1] lat/lon GmbH

[2] lat/lon GmbH, deegree initiative

Abstract

deegree is the most comprehensive open source implementation of OGC standards within the stack of OSGeo projects. The mission statement of the deegree initiative points out, that deegree is able to tackle simple as well as complex challenges. One of the big themes where tackling complex challenges in the world of spatial data infrastructures is INSPIRE, the European SDI. This presentation outlines the technical challenges of fulfilling the requirements of the INSPIRE technical guidance documents and how these are coped by deegree webservices.

Overall, there are two big challenges, the technical implementation of INSPIRE brings along. The first one is to harmonize data to follow the INSPIRE data specification. This means, datasets have to be compliant to the complex INSPIRE GML schemas and the underlying OGC GML 3.2.1 standard. The second challenge consists of the deployment of discovery, download and view services which are fulfilling the requirements for INSPIRE network services. Here deegree webservices come into play. deegree is OGC reference implementation for nearly all underlying OGC standards which are required by INSPIRE (GML 3.2.1, WFS 2.0, WMS 1.3.0) and others (WFS 1.0.0, WFS 1.1.0, WMS 1.1.1). A must requirement specified in most of the INSPIRE Technical Guidance documents is that an implementation has to be compliant to the specific ISO/OGC standards.

Putting the two mentioned challenges together results in a hypothesis like: "INSPIRE network services must be able to provide INSPIRE harmonized data through compliant service interfaces". The presentation shows how this "super-challenge" can be handled using deegree webservices.

References

✔ deegree - open source software for spatial data infrastructures and the geospatial web: http://deegree.org/

✔ Mission statement of the deegree initiative: http://wiki.deegree.org/deegreeWiki/MissionStatement

✔ OGC compliant products details for deegree: http://www.opengeospatial.org/resource/products/details/?pid=1139

✔ INSPIRE – Infrastructure for spatial information in the European community: http://inspire.ec.europa.eu/

Raster Data in Django

Daniel Wiesmann[1] and Mike Flaxman[2]

[1] Urbmet Limited, Lisbon, Portugal

[2] Geodesign Technologies Inc, San Francisco, California, USA

Abstract

Integrating raster data in modern web applications is challenging even for experienced web developers. Existing software to create raster based GIS web services are not directly integrated into modern web frameworks such as Django, Ruby on Rails, or Drupal. Developers striving to create online GIS services including raster data are often bound to learn, install and maintain software external to their main web framework.

In this talk we present first results of our ongoing effort to integrate raster data into GeoDjango, the GIS module of the Django web framework. Django is a popular modern web framework written in the Python programming language (Django Software Foundation [DSF], 2015a). It contains a GIS module called GeoDjango which has the goal to "make it as easy as possible to build GIS Web applications and harness the power of spatially enabled data" (DSF, 2015b). GeoDjango has custom bindings to GDAL (2015) and GEOS (2015) and connects to multiple spatial backends (PostGIS, MySQL, Oracle Spatial and SpatiaLITE). It has extensive GIS capabilities for vector data, but did not have any support for raster data up to the current release (Django 1.7).

In collaboration with Django core developers, we have contributed to adding basic raster bindings to the GDAL interface of Django, which will be released in version 1.8 of Django (DSF, 2015c). Creating the bindings to the GDAL raster capabilities was the first step in adding raster as a data type into the Django model structure (DSF, 2015d). The goal for the Django 1.9 release is to add a Raster Field to Django, which will allow writing and retrieving raster data with the same ease as writing ordinary text or numerical data.

Our contribution to this raster integration is inspired by our applied work with raster data in large scale scenario planning. For our applied work, we have been developing a plugin to Django as an external python package called django-raster (Wiesmann & Flaxman, 2015), which we are actively using in our applications. The package still relies on GDAL's own Python bindings, but otherwise functions similarly to what is being integrated directly into Django. Through our real world use cases, we identify the main difficulties in serving rasters through Django early on. Furthermore, we develop and test essential features that will be necessary in any applied case such as color-schemes and legend generation. With the django-raster package, and the upcoming deeper integration into GeoDjango and the Django web framework, the creation of raster based online services such as Web Map Services or Tile Map Services becomes relatively straightforward, even for web developers with limited GIS experience.

At the end of our talk, we showcase how we serve raster tiles through Django in our applied work.

References

- ✔ Django Software Foundation. (2015a). *Django – The web framework for perfectionists with deadlines.* Retrieved March 9, 2015, from https://djangoproject.com/
- ✔ Django Software Foundation. (2015b). *GeoDjango*. Retrieved March 9, 2015, from https://docs.djangoproject.com/en/dev/ref/contrib/gis/
- ✔ Django Software Foundation. (2015c). *Django 1.8 Release Notes.* Retrieved March 9, 2015, from https://docs.djangoproject.com/en/dev/releases/1.8/
- ✔ Django Software Foundation. (2015d). *Add a RasterField to GeoDjango*. Retrieved March 9, 2015, from https://code.djangoproject.com/ticket/23804/
- ✔ GDAL. (2015). *Geospatial Data Abstraction Library.* Retrieved March 9, 2015, from http://gdal.org/
- ✔ GEOS. (2015). (Geometry Engine - Open Source). Retrieved March 9, 2015, from http://geos.osgeo.org
- ✔ Wiesmann, D. & Flaxman, M. (2015). *Django-raster python package.* Retrieved March 9, 2015 from https://github.com/geodesign/django-raster

istSOS: latest developments and first steps into the OSGeo incubation process

Massimiliano Cannata[1], Milan P. Antonovic[1], Mirko Cardoso[1], Priska Pietra[2]

[1]Institute of Earth Science – SUPSI

[2] (GSoC Student)

Abstract

istSOS (https://geoservice.ist.supsi.ch/projects/istsos) is an OGC SOS server implementation entirely written in Python. istSOS allows for managing and dispatching observations from monitoring sensors according to the Sensor Observation Service standard. istSOS is released under the GPL License, and should run on all major platforms (Windows, Linux, Mac OS X). The presentation will go through the details of all the new features that will be packed in the next release. In particular the presenters will introduce enhancements that include the Advanced Procedures Status Page and the istSOS Alerts & Web Notification Service. The istSOS Advanced Procedures Status Page is a new section of the Web graphical user Interface, offering at a glance a graphically representation of the Sensor Network health. Administrators can easily figure out common issues related with sensor data acquisition and transmission errors. The istSOS Alert & Web Notification Service are the result of the Google Summer of Code 2014 outputs. This service is a REST implementation that take inspiration from the OGC Web Notification Service (OGC, 2003; OGC, 2006a) and the Sensor Alert Service (OGC, 2006b) which currently are OpenGIS Best Practices. Alerts are triggered by customized conditions on sensor observations and can be dispatched through emails or social networks.

This year istSOS is entering into the OSGeo incubation process, this new challenge will permit to enhance the software quality and consolidate the project management procedures. The presenters will present the incubation status and discuss about the next steps.

References

- ✔ Open GIS Consortium Inc. (2003). Web Notification Service, Version: 0.1.0. OpenGIS Discussion Paper. Editor: Ingo Simonis, Andreas Wytzisk. Retrieved from http://portal.opengeospatial.org/files/?artifact_id=1367
- ✔ Open GIS Consortium Inc. (2006a). Web Notification Service, Version: 0.0.9. OpenGIS Best Practices Paper. Editor: Ingo Simonis, Johannes Echterhoff. Retrieved from http://portal.opengeospatial.org/files/?artifact_id=18776
- ✔ Open GIS Consortium Inc. (2006b). Sensor Alert Service Candidate Implementation Specification, Version: 0.9. OGC Best Practices. Editor: Ingo Simonis. Retrieved from http://portal.opengeospatial.org/files/?artifact_id=15588

Debian and Ubuntu GIS

Van de Wauw Johan

GIS Consultant, RealDolmen, A.Vaucampslaan 42, 1654 Huizingen, Belgium

Abstract

Debian, is one of the oldest and most widely deployed open-source operating systems, not only among end users but also as a basis for other systems. Ubuntu is a linux-distribution which is derived from Debian.
Anyone using Debian or Ubuntu is used to find much of his prefered software as packages in the repositories. These include a growing number of GIS related libraries and applications, packaged by Debian GIS and Ubuntu GIS. These packages form the basis for OSGeo live, and are increasingly being used in automated tests for many projects.
This talk has two goals: 1) introducing the community behind these projects. 2) highlighting what steps are needed to have your application included in Debian and/or Ubuntu. Or put differently "How to make sure you favourite application does (not) reach Debian/Ubuntu GIS".
The target audience are not only developers, but also users of Debian/Ubuntu: In fact many tasks that are handled by the team (license review, documentation, testing) and to some extent the packaging don't require in-depth programming knowledge. And ... all of us started as users!

References

- ✔ Debian Project, Debian: http://www.debian.org
- ✔ Debian GIS Project: https://wiki.debian.org/DebianGis
- ✔ UbuntuGIS Project: https://wiki.ubuntu.com/UbuntuGIS
- ✔ OSGeo Live: http://live.osgeo.org

OpenQuake: an Open Platform for Earthquake Risk Awareness and Assessment

Paul Henshaw

GEM (Global Earthquake Model) Foundation

Abstract

The GEM Foundation (http://www.globalquakemodel.org) is a private-public partnership that drives a global collaborative effort for transparent assessment of earthquake risk and risk management around the globe. The OpenQuake platform (http://platform.openquake.org) is GEM's open source, interactive web-based suite for earthquake risk awareness and risk assessment. The platform provides access to software, datasets, models and guidelines and also encourages users to share their contributions with the rest of the community via specialised data collection tools.

The OpenQuake software suite also comprises a calculation engine for seismic hazard and physical risk analysis, and desktop tools to facilitate the construction of integrated risk models. The components and their interactions are illustrated in Figure 1.

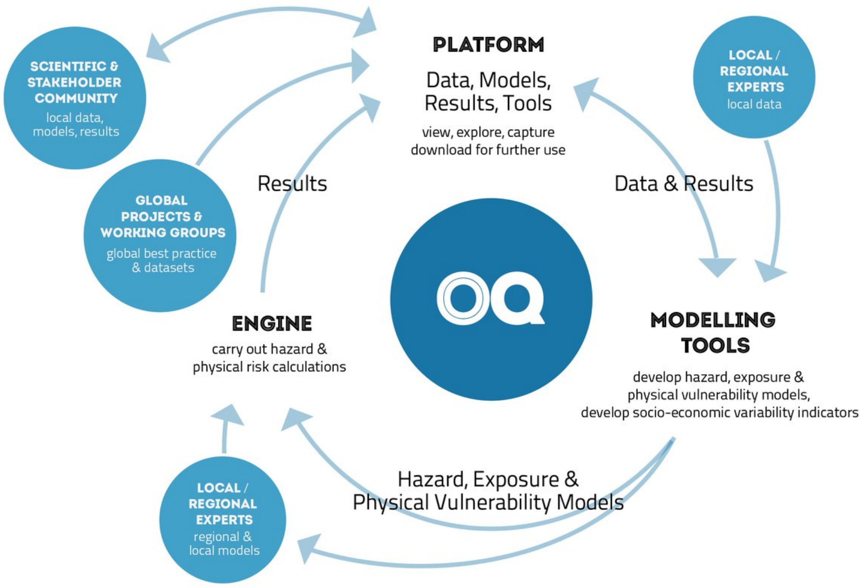


Figure 1: OpenQuake components and community interaction.

Users can explore data and models present in the platform and download elements for use in their own calculations. The modelling toolkits can then be used to produce an input model from the downloaded data, perhaps in combination with local datasets. The OpenQuake engine can then be used to perform seismic hazard and risk analyses. The results of hazard analyses may be used as input for physical risk models. The results of physical risk analyses may be combined with socio-economic indicators to produce a model of integrated risk. Users can upload models and results to the platform and optionally share them with the rest of the community who in turn can use them to create and share new products. Much of the data currently available on the platform was produced through collaborative projects with scientific consortia and regional experts.

The platform was built using GeoNode (http://geonode.org) leveraging its built-in support to manage maps, documents and users. Some GEM data products however are not suitable for display using the standard features and require custom handling. Two distinct techniques were adopted for adding domain-specific features: in some cases existing components were extended, for example a graphical workflow for crowdsourcing Active Fault information using GeoExplorer; in other cases, separate Django applications were used to construct REST API endpoints and interactive applications for exploring structured non-map information using technologies such as Leaflet, D3 and UTFgrid.

The Integrated Risk Modelling Toolkit combines web-based applications and APIs provided by the platform with a QGIS (http://qgis.org) plugin to form an integrated workflow allowing the user to select and download socio-economic indices, construct and explore a model of integrated risk and upload their work to the platform. This combination of FOSS4G tools both on the desktop and on the server results in a powerful integrated workflow for scientific collaboration.

OpenQuake software source code is available under the terms of the AGPL from GitHub (https://github.com/gem). GEM data products are made available from the platform under the terms of Creative Commons licenses. Using open data and open source software permits transparent and repeatable calculations important not only for scientific credibility, but also facilitating regulation compliance and community collaboration.

QGIS Support for Map Projection Distortions Visualization

Dražen Tutić

University of Zagreb, Faculty of Geodesy, Croatia

Abstract

Most important properties of a map projection, i.e. reference coordinate system for spatial data are size and distribution of angle, distance and area distortions. If performed in an inappropriate way, without consideration of map projection distortions, calculations in plane reference coordinate system can lead to unreliable results and analysis. GIS software usually have support for these problems, for example, it is possible to get distances, azimuths or areas on sphere or ellipsoid. It can be assumed that awareness of these facts is not widely accepted by regular GIS users. Expert users acquainted with theory of Earth's shape and its projection into the plane are exception to this assumption.

Usual way for visualisation of map projection distortions are Tissot's indicatrices which are not usually supported by GIS and are not easy to understand by regular GIS user. In this work, different way to visualize these distortions is proposed to be implemented into QGIS. Map projections distortions can be interpreted as field which mean that raster representation of its values is appropriate. When raster is calculated it can be visualised by known methods for thematic maps, or contours can be extracted as vector representation. Having this, GIS user has convenient and easy way to estimate impact of map projection distortions and subsequently to decide whether is it important to take then into account.

PROJ4 has function pj_factors which calculates scales (meridian, parallel, areal), convergence, meridian-parallel angle, angle distortion, max and min scale (Tissot's indicatrix axes) for a given point. Verbose output for a projected point is controlled by –V option of PROJ4 executable. Function pj_factors is not included in PROJ4 API.

The proposed way to include calculations and visualizations in QGIS is to include pj_factors in PROJ4 API in a such way that data structure can be changed in future. Next step is to bind this function into QGIS core and its class QgsCoordinateTransform. This way output from pj_factors is available for core functions or plugin development. Later is chosen as solution for adding this functionality to QGIS. Alternative approach is also implemented which calls PROJ4 executable and parse distortion values from its output. This has advantage that plugin can be used without modifying existing PROJ4 API or QGIS core at the expense of speed. Since raster can have large number of points, first approach is far better solution, and alternative can only be useful for limited usage.

Plugin is designed in a way that it takes extent of existing layer, or extent is manually entered, user choses scales, distortions, convergence etc., as raster cell value, raster size or resolution and filename. When raster is created, using existing QGIS functionalities, user can change how raster is visualised, extract contours or perform any other meaningful analysis of newly created raster.

Intelligent SDIs with MapMint 2.0

Bozon Nicolas[1], Fenoy Gérald[1], Raghavan Venkatesh[2]

[1] GeoLabs SARL, FR

[2] Osaka City University, JP

Abstract

This conference aims at presenting the status of the MapMint open source project and its most recent 2.0 version. The upgrade to newer versions of its core open source components will first be explained. The extensive use of OGC standards through ZOO-Project 1.5, GDAL 1.11 and MapServer 7 is indeed making MapMint an even more stable and efficent foundation to build an open source and standar-compliant spatial data infrastructure. The new MapMint responsive user interfaces based on OpenLayers 3 and Bootstrap will then be presented. Both code and documentation improvements will also be detailed. The newly added functionnalities in MapMint 2.0 will finally be explained from the developer and user point of views, based on case studies and live examples.

From mobile weather sensors to open weather data communities

Le Glaunec Julien, Freyre Jérome, Rappo Daniel, Ertz Olivier

HEIG-VD, School of Business and Engineering Vaud, University of Applied Sciences Western Switzerland

Abstract

Weather data are usually collected by fixed weather stations. Skywatch Windoo is based on a different approach. JDC Electronic and HEIG-VD have partnered to think about a mobile weather station enhancing a smartphone with new sensors: wind speed, temperature, humidity and pressure. Called the Windoo device, it works without any battery, just powered through the jack plug. Then, data are transmitted to the smartphone still through the jack plug using signal analysis and a dedicated asynchronous communication protocol. In order to abstract these low level communications, a hardware API was developed to access the Windoo device from any Android and iOS application. Everybody is now able to build its own application and users community (e.g. Windoo for ski touring, Windoo for paragliding, Windoo for nature conservation, etc.). For now, API users develop applications in the fields of windsports, environment and ballistics.

Beside the hardware API for developers, the second aspect of the project concerns the Skywatch Windoo branded application which is used by the community of weather enthusiasts. This iOS/Android native application allows instant measurements, graphical display and measurements history.

Moreover the project does open interesting perspectives for Volunteered Geographic Information (VGI). Currently, the two versions of the native application let the user share measurements on the Windoo.ch Community Map and afterwards on Facebook and Twitter. When sharing, she/he is invited to associate a picture taken at the time of measurement. Thus, the platform intents to gather measurements and pictures sent by the contributors. These ones may complete their user profile and sharing parameters so that contributions are recognized and everyone is able to visualize each user's own measurements on a map. Some share a breathtaking view from a high summit while other just shot a mindless but “meteo-tagged” picture from their balcony.

The Community Map is based on opensource components: LeafletJS on client-side using OpenStreetMap as a basemap, the Laravel PHP web framework on server-side, and MySQL for data storage. The data can be accessed through a JSON service. Windoo application may be improved in several ways and there are on-going discussions on how to federate common development interests.

After several monthes, we can observe that the Community Map has its own life with many people ready to produce weather VGI, although it was not a major initial intent of the project. Nonetheless, this observation leads us to some interesting insights for the future, especially to stimulate the community

of weather enthusiasts. Firstly, even though developers have at disposal a simple API to access the device, they still have to build entirely their custom application to present, store and share data. Therefore it would be valuable to provide an enriched API more oriented on “community data sharing”. This would imply to purpose some ready-to-use widgets and to offer better data structures, describing sensors data, the user, the application, the device. Consequently, we could imagine to offer access to the community data themselves, with or without a Windoo device. And finally, these insights would require to build a solid cloud infrastructure, allowing any application provider to solve the hosting question.

References

- ✔ Skywatch Windoo :: The weather station: http://www.windoo.ch/en
- ✔ JDC Electronic SA: http://www.jdc.ch/en
- ✔ HEIG-VD: http://www.heig-vd.ch/international

Managing the Flemish Functional Cycle Network: a FOSS4G solution

Karel Dieussaert[1], Jan Pelckmans[2]

[1] Geo Solutions N.V., karel.dieussaert@geosolutions.be

[2] Afdeling Beleid, Mobiliteit en Verkeersveiligheid, jan.pelckmans@mow.vlaanderen.be

Abstract

Everyday people make a lot of movements between their home and attraction poles in cities and municipalities such as work, school, stores,.... A lot of these functional movements could be done by bike if the appropriate infrastructure would be there. More than 10 years ago, The Flemish Ministry of Mobility and Public Works defined a Functional Cycling Network (FCN) to realise a logical and safe infrastructure for functional cycle movements between these environments.

The GIS architecture that was set up in that period was not appropriate anymore to serve the current goals and possibilities. Figure 1 shows this former architecture; the base is a central database (Oracle). Every work cycle, distributed datasets (Personal geodatabases) were created for editing in a Desktop GIS (ESRI). After editing the information was validated, corrected and re-entered in the central database. Reporting (statistics & static maps) was based on the central database.

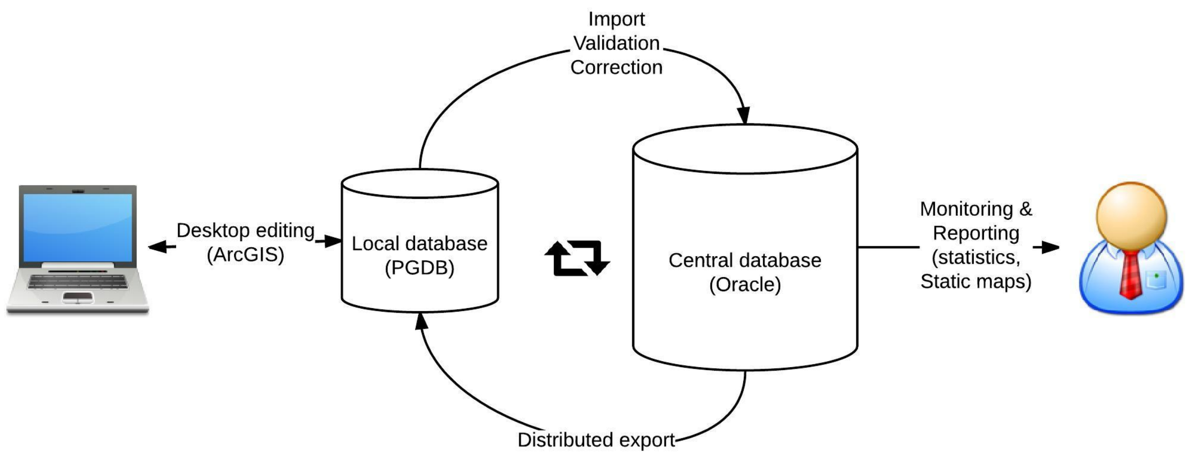


Figure 1: Former architecture with distributed datasets for desktop editing.

The scope of this project was to provide a GIS-based web tool to manage the data and provide tools to monitor, analyse and report about the FCN. This to improve the FCN according to defined standards for cycling infrastructure. As this tool is going to be used by the responsibles in every municipality of Flanders, they expect a reliable, performant and stable system that can

manage this kind of data.
Figure 2 shows the state-of-the-art architecture, based upon several FOSS4G components (PostGIS, Geotools, Geoserver, Openlayers 2/3) and OGC standards (WMS, WFS, WPS).

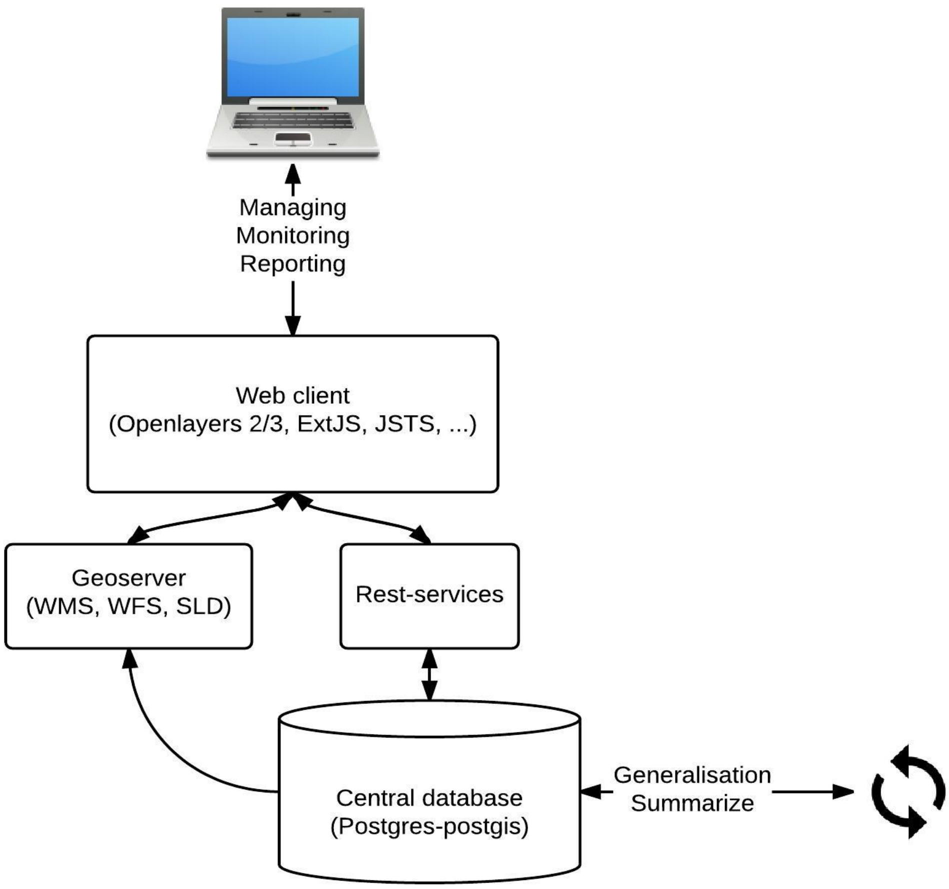


Figure 2: New architecture for GIS based web tool based on FOSS4G.

Postgres/Postgis serves as the central database where all the (geo)data is stored. The baseline road network is lineair referenced and all the attribute information about the FCN is stored as event tables.
For performance issues the attribute information is generalized and materialized for visualisation purposes. Every night, data is summarized to serve predetermined reports.
The map in the web client is mainly based on OpenLayers 2. Only for the editing window OpenLayers 3 is introduced for its capabilties to turn the map to show bicycle roads in a horizontal manner (Figure 3). Furthermore the web client integrates all functionalities for editing, monitoring and reporting the FCN (Figure 4).
Together with the renewed architecture, a new baseline road network was introduced for the FCN. Because the new baseline differs geometrically from the former baseline, OpenLR was used to convert the FCN. OpenLR is a technique to encode routes in a map agnostic format so it can be decoded to a

geometrically different road network. Afterwards, validation and manual correction of the transformed FCN was still necessary.
All attribute information, referenced to the former baseline was transformed to new events on the new baseline using a combination of Lineair Referencing Postgis Queries.
This project shows that FOSS4G components can be used to create a reliable and stable GIS application and have the ability to overcome typical migration and linear referencing problems.

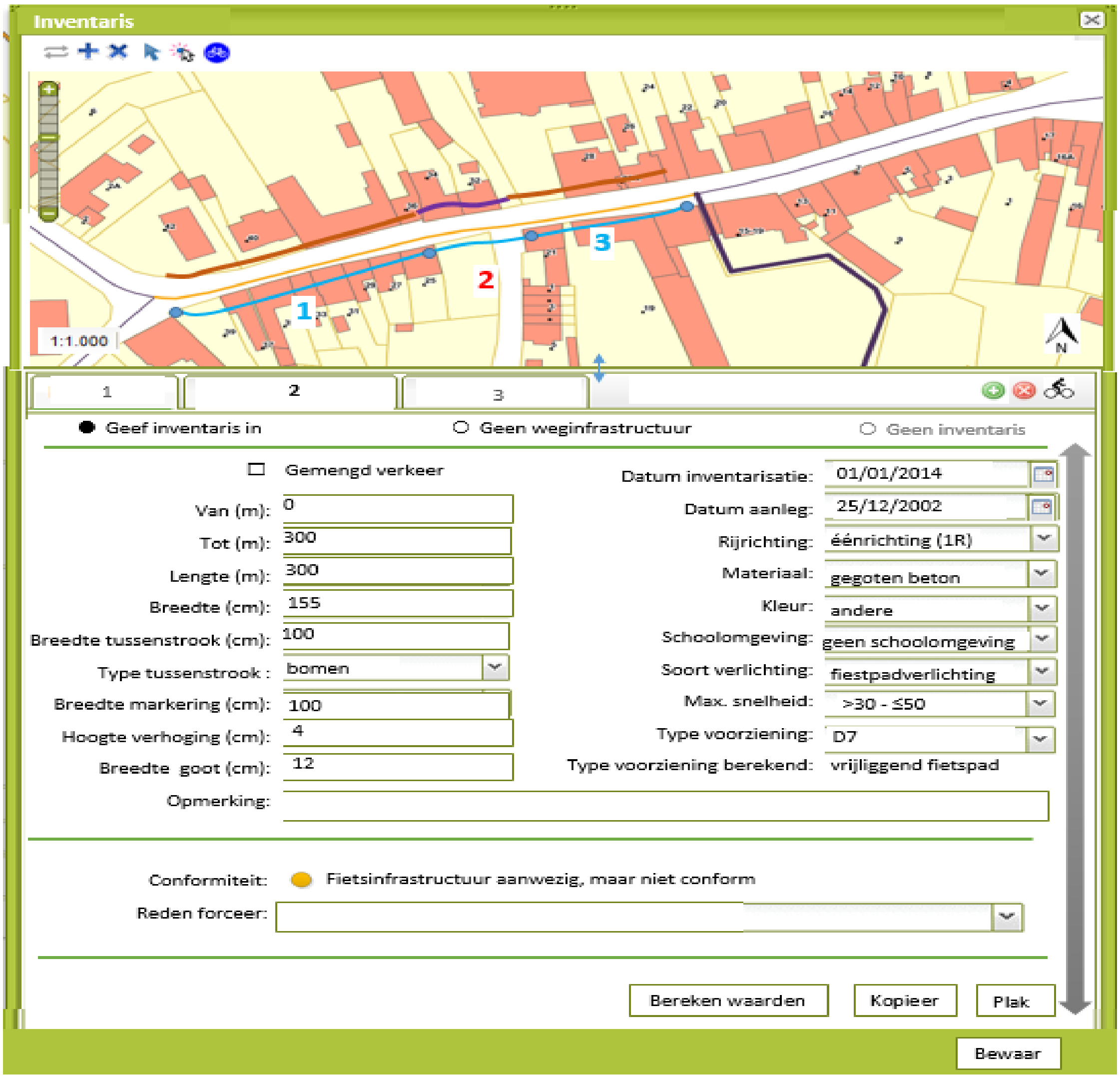


Figure 3: GIS-based web client editing capabilities.

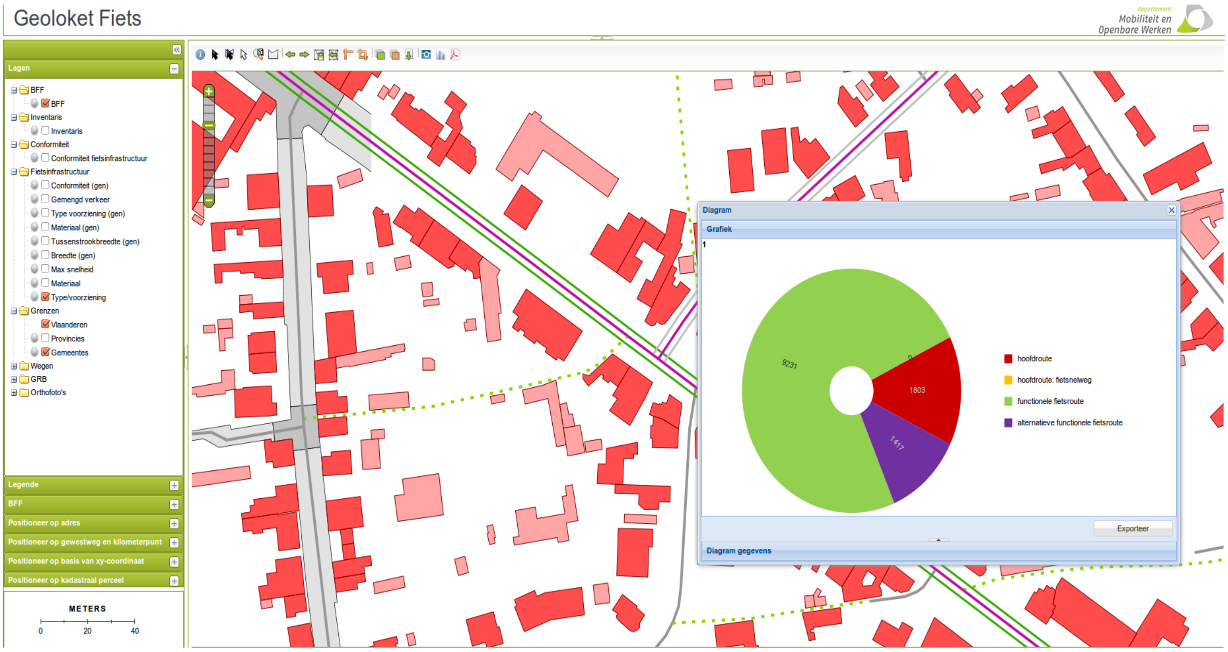


Figure 4: GIS-based web client reporting general statistics.

References

- ✔ Geo Solutions N.V. (2014). *Business analysis Functional Cycle Network*. Kontich, Belgium: Geo Solutions N.V.
- ✔ Geo Solutions N.V. (2014). Functional *analysis Functional Cycle Network*. Kontich, Belgium: Geo Solutions N.V.
- ✔ TomTom International B.V., & DLR, German Aerospace Center. (2012). *OpenLR: An open standard for encoding, transmitting and decoding location references in digital maps* (White Paper version 1.5). Amsterdam, The Netherlands: TomTom International B.V. Retrieved from http://www.openlr.org/data/docs/OpenLR-Whitepaper_v1.5.pdf

OL3-Cesium: 3D for OpenLayers maps

Beraudo Guillaume

Camptocamp

Abstract

OL3-Cesium is a new WEB library used to create and synchronize a Cesium 3D globe from an OpenLayers3 map. The library takes care of automatically reading the raster and vector layers from the map and synchronize them to 3D. Additionally, the view parameters — centre, resolution, rotation — get synchronized bidirectionally between the map and the globe, allowing interactions to span naturally between 2D and 3D.
The library was started by three companies from the OpenLayers community. Having the generic and complex synchronization logics implemented in one place allows easier maintenance and porting to new versions of OpenLayers3 and Cesium. It also helps create new applications more quickly, with less code and focus on application specific needs; setting up a stacked Cesium globe on top of OpenLayers3 is as straightforward as invoking "*var ol3d = new olcs.OL3Cesium({map: ol3Map});*".
In this talk we will demonstrate practical cases of map enhancement with 3D, show code and explain what happens under the hood. Our main interests will be:
- easy kick start of a side-by-side or stacked application;
- handling of different raster and vector projections;
- positioning vectors on terrain or at given absolute 3D coordinates;
- editing and picking Point Of Interests in a unified 2D/3D manner;
- streaming and displaying vector data;
- streaming and displaying buildings.

References

✔ https://github.com/openlayers/ol3-cesium
✔ http://openlayers.org/ol3-cesium/examples/
✔ http://openlayers.org/ol3-cesium/apidoc/olcs.OLCesium.html

What is this new Rust language and why should a GIS developer care about it?

Alex Morega

geo-spatial.org

Abstract

Rust is a new systems language, developed at Mozilla, which provides static analysis of memory safety, and borrows ideas from functional and object-oriented programming. With powerful and elegant abstractions, it's a great language to develop GIS applications, processing data with minimal overhead, without sacrificing code readability. We explore the specifics of the Rust programming language, the packaging system, how they fit the needs of GIS developers, and take a look at the library ecosystem, including libraries for processing spatial data.

References

✔ The Rust programming language - http://www.rust-lang.org
✔ GeoRust - https://github.com/georust

Fire and ice - 3d visualisation of seismic activity during the 2014-2015 Bárðarbunga volcanic eruption

Tryggvi Hjörvar

Icelandic Meteorological Office

Abstract

On August 16th 2014 an unusual burst of seismic activity began under the volcano Bárðarbunga, in north-western Vatnajökull glacier in Iceland. Over the course of the next few weeks, scientists at the Icelandic Meteorological Office (IMO) closely followed the propagation of the activity to the east and north of the glacier, using a network of seismometers and high-precision GPS measurement stations. On August 31st an eruption started on a fault north of the main volcano, eventually spewing 85 square kilometers of lava onto the surface before its end on January 27th 2015. The eruption is believed to be the start of an active period in the Bárðarbunga volcanic system, likely to last up to a decade.

At the time of the activity start there were no compelling visual tools to present what was going on in the area to the public, beyond traditional two-dimensional maps placing earthquake epicenters on the surface.

Using the three.js javascript library and open geospatial data from the IMO and the National Land Survey of Iceland, the IMO IT-team quickly developed a 3-dimensional, interactive model of the seismic activity, showing earthquake hypocenters in relation to the surface terrain. The model is accessible to the public on the IMO website and updated with the most recent seismic data within minutes.

The new tool has also proved to be of great help to the scientists monitoring the activity in real-time, giving them a comprehensive spatial and temporal overview of the events as they continue to unfold.

3D Application for Animal Tracking

J. Sala[1], I. Polo[1], A. López[1], L. Romero[1], A. Duplaa[2], G. Weller[2], D. Stakem[3]

[1] Altamira Information, Corsega, 381, 08037 Barcelona, Spain

[2] CLS, 11 Rue Hermès, 31520 Ramonville-Saint-Agne, France

[3] CLS America, Inc. 4300 Forbes Blvd. Suite 110 Lanham, MD 20706

Abstract

New visualization trends bring new light to the way environmental monitoring data is exploited today. The presented application is an enhancement of the user experience for the visualization and analysis of the Argos service for animal tracking. Argos is a satellite-based system that collects data from Platform Terminal Transmitters, and distributes sensor and location data in order to understand the distribution of wildlife in a given territory.

Virtual globes constitute a visualization platform of geospatial information projected onto the Earth Digital Elevation Model (DEM) in three dimensions, users interactively rotate, zoom and pan the available data, usually projected on a given background that provides a geographic context. The 3D platform allows access to animal tracking data. It also permits the extraction of specific information, path animation and data download from a dedicated and secure platform, with the objective of facilitating the extensive use of this source of information.

The animal tracking 3D desktop application has been built with the World Wind SDK, an open source virtual globe built in Java/OpenGL and developed by the National Aeronautics and Space Administration (NASA). The user can combine tracking data with layers from any public server implementing the WMS (Web Mapping Service) protocol of the OGC (Open Geospatial Consortium) through a tab menu able to handle the layers and opacity. Animal tracks can be animated individually alongside the elevation profile of the trajectory with the error ellipsoids of each position of the animal. As for the input/output capabilities of the tool, the CSV format is used as default in order to be compatible with existing operational services. KML is also available to enhance interoperability within GIS communities. The application allows the possibility to query directly from the WebServices that hold the animal data through a REST API using SOAP (Simple Object Access Protocol) with authentication.

The new 3D approach to the way data is visualized and the possibility to overlay a wide range of data from open WMS servers, allows users of the service to have a more realistic and integrated vision of their animal positioning data.

Landslide risk forecasting based on Unique Conditions Units

Marco Cibrario[1], Andrea Giuseppe Stralla[1], Valentina Marin[2], Paola Salmona[2], Gerardo Brancucci[2]

[1] Università degli Studi di Genova

[2] Geomorfolab – Laboratorio di geomorfologia applicata Università degli Studi di Genova

Abstract

Geological instability, despite being a natural condition, over the years has possibly been worsened by human activities. In particular, cases are increasing where human lives, properties or natural and cultural heritage are heavily affected.

In such a context, this project aims to provide a contribution to the prevention of landslides. A methodology has been proposed and tested in the Ligurian area of "Chiavarese" (Eastern Liguria, Italy) to statistically identify public heritage and properties exposed to landslide risk. Starting from georeferred landslides reports, besides, it can help to identify those areas where comparable events are more likely to happen.

Altitude, slope and aspect (from a 5 meters DTM), land use, geology and rainfall erosivity have been taken into consideration as landslide predisposing factors and eventually classified. The raster maps obtained for each factor have been crossed in GRASS (r.cross) to calculate landslide susceptibility according with the definition of Unique Condition Units (UCU), land units containing the same combination of all predisposing factors. In order to obtain percentage values of the landslide predisposition of the factors combination, it has been to taken into consideration the Italian Landslide Inventory (IFFI project, by ISPRA) that defines 8 different types of landslides, on both national and regional scale. UCU and landslides inventory map – converted into raster of the same resolution – have been crossed all the pixels with the same combination of predisposing factors have been pointed out. Through the GRASS r.stats command a matrix have been set, that provides the number of UCU in the input landslide type. A simple proportion gives as a result the statistical percentage of the pixel susceptibility to the selected type of landslide.

The proposed methodology is going to be continuously updated and improved by crowd-generated inputs, in particular landslides reporting. Using the Ushahidi platform, a crowdmap accessible through a website, a smartphone app and social media have been built, in order to collect georeferred landslides reports by generic users. Even if it is spatially located in a few UCUs, the triggering event could have a wider spatial distribution that could cause landslides in areas where all the predisposing factors are the same as the one reported. A SpatiaLite database helps to manage all of the data, and through some predefined query it is possible to export and visualize in QGIS those buildings and infrastructures exposed to different risk percentages.

MAppERS for crowdsourcing. Citizens and volounteers as human sensors

Simone Frigerio

Italian National Research Council - Research Institute for Geo-Hydrological Protection, C.so Stati Uniti, 4 - 35127 Padova (ITALY)

Abstract

MappERS[1] (Mobile Application for Emergency Response and Support) is a EU project funded under programme 2013-2015 Humanitarian Aid and Civil Protection, ECHO A5. The project deals with capacity of human sensors towards smartphone application (SA) in the contest of prevention and crisis support for natural hazards. Citizens and volunteers are first actors of civil protection and the crowdsourcing creates local and context-specific strategies of surveillance and territorial management. The SA splits into module MAppERS-V for volunteers, and module MAppERS-C for citizens re-designed according to methodological and logical testing gained during pilot study. People involved in training and piloting fulfills a long term objective of participation and crowdsourcing as actors of prevention of hazards, according to the priorities set by the Hyogo Framework, reducing costs of emergency management and promoting responsibilisation of the population. Training curricula for citizens and volounteers will promote awareness, territorial knowledge and specialized jargon to communicate hazard-relevant information toward smartphone. For a meaningful SA a review of best practices of information scheme is compulsory. An analysis on Graphical User Interface (GUI) adopted within SA for natural hazards offers a baseline of communication scheme for MAppERS frame. Following Quesenbery (2003) the usability as design principle handles five components. The efficiency is the speed with which users achieve their goals, the effectiveness highlights the completeness and accuracy with which users achieve their goals, the engagement reveals the level to which the style of the interface makes using the application a pleasing experience. Furthermore the easiness to learn describe how simple it is to start using the interface, and get a deeper understanding of its capabilities. The error tolerance finally shows how the design prevents users from committing errors, and how it helps in correcting them when they occur. The review strengthened criteria for layout, navigation, accessibility, icon setup and text guidelines design (Fowler & Stanwick, 2004; Extron, 2010; Garrett, 2011; Wong, 2011). The development is proceeding for Android 4.0 onwards, covering 82.6% devices as of today. The Web Application includes a Cloud Architecture, a Relational Data Base and Web services, integrated with a Cloud Notification SDK to send push notification. The Mobile App Android includes an app Skeleton, the communication structure

1 http://www.mappers.eu/

(SMS, Emails and phone calls). The bug-fixing and training will be set during piloting and the crowdsource becomes a continuous source of development, testing and updating each smartphone components. The wireframe contains a screen for registering and login common for both modules. The access and tools are later split for the two modules. MAppERS-C proposes a Personal Flood Plan toolkit. Users are citizens, voluntary registered and able to create own set of prevention measures, sharing geo-located information during critical step before flooding events. MAppERS-V is a Danger Survey toolkit for rescue crew, to collect and organize real-time and standardized information for damages during crisis. The first draft of modules provides options to customize dropdown lists and sliding tools for end-users, enlarging the quality of the application based on local experts and witnesses. The expected results are: a) an easy-to-use interface for “human-data” in crisis management, b) a maximised utility of peer-produced data gathering, c) the development of human resources as technical support d) a self-based awareness improvement.

References

- ✔ Quesenbery, W. (2003). The Five Dimensions of Usability, in Albers, M. J. and Mazur, B. (Eds.), Content and Complexity: Information Design in Technical Communication, Lawrence Erlbaum Associates, Mahwah, NJ, 2003.
- ✔ Fowler S. & Stanwick V. (2004). Web Application Design Handbook – Best Practices for Web-Based Software, Elsevier, San Francisco, 2004
- ✔ Extron Electronics (2010). The Extron Guide to Graphical User Interface Design 2010, online at: http://www.extron.com/company/article.aspx?id=guiresources
- ✔ Garrett, J. J. (2011). The Elements of User Experience: User-Centered Design for the Web and Beyond, 2nd Edition, New Riders, Berkeley CA, 2011.
- ✔ Wong, B. (2011). Color blindness, Nature Methods, 8.6: 441.

Collecting and Processing Land Surveyors' Sensor Data

Zoltan Siki

Budapest University of Technology and Economics

Abstract

Our open source project, called Ulyxes was redesigned and rewritten in Python recently. The aim of our OS project is to collect data from robotic total stations, GNSS receivers, Web cameras without human interaction. Collected data are uploaded to PostGIS database and published on the Internet using OpenLayers and MapServer.

For data processing the SurveyingCalculation QGIS plug-in was developed. Special surveying calculations as traversing, network adjustment, co-ordinate transformations can be solved by the Python plug-in. Network adjustment uses GNU Gama OS project.

The developed tools are used for educational purposes at the university and they were tested in real live situations. Students are involved in the development.

The state of the project and some application will be presented. Other sensors are planned to integrate.

References

- ✔ GeoCom Reference Manual TPS 1100 – Version 1.07 (2002)
  Leica Geosystems AG, Heerbrugg, Switzerland
- ✔ Ulyxes Web site - http://www.agt.bme.hu/ulyxes

vOL3 – OpenLayers 3 for Vaadin 7

Stypinski Martin

Direct Mail Company

Abstract

Most web based GIS applications today are written in a mix of languages such as PHP/JS or Phyton/JS, etc. Many frameworks exist to support the development team and make the programming cycle as short as possible. But there is still a downside of choosing a client and a server side technology on its own. The problem starts with knowing at least two technologies to the fullest and there is no end in sight. Even server / client side communication can bring up a few issues and make developers life really troublesome.

With Vaadin (https://vaadin.com/home) these problems are gone, somehow to a certain point. Development starts straight in Java and as it is one of the bigger Java frameworks in the game, support is quite strong and guaranteed. With only one language to develop, the main asset of development is posed on 'making a nice application' rather than 'bolting things together and make it work'. Vaadin somehow enables the focus on the important things instead the technology. The software itself runs on an Apache Tomcat Server and uses Googles GWT technology to render all displaying webpages.

To make use of Vaadin as a framework for GIS internet application there is still a missing piece. As OpenLayers 3 (http://openlayers.org/) has grown to a very popular and strong Java Script library it was a very obvious step to use this technology. VOL3 is nothing more than a wrapper class for all the strong and well working functions of OpenLayers 3 to make them usefull in Vaadin 7.

This framework combination is particularly intressting for bigger software projects or as a GIS enabeling piece for already existing projects. With the new vol3 library we have finally brought OpenLayers 3 to the Vaadin world.

New QGIS functionality for Pros

Sandro Mani

Sourcepole AG

Abstract

Among the constant stream of improvements planned for upcoming QGIS releases, two exciting features include support for curved geometries as well as the introduction of two C++ plugins to detect and automatically fix geometry errors.
Curved geometries have long been a missed feature in FOSSGIS Desktop solutions, with such geometries usually ending up being segmented on import. A rewrite of the QGIS Geometry core now allows for native support of a number of curved geometry types, such as CircularString, CompoundCurve, CurvePolygon, etc., in addition to the traditionally supported Point, Line and Polygon geometries. As part of the redesign, proper support for M and Z coordinate values was also implemented for all supported types.
Geometry errors can easily sneak into large datasets, either because of inexact data acquistion, but also due to gradual loss of precision when importing, exporting and converting the datasets to different formats. Applying analysis operations (such as unions, intersections and differences) to such datasets can easily produce many tiny fragmens (so-called sliver polygons), or even fail outright due to invalid geometries. Manually detecting and fixing such issues can be very time consuming. To assist users confronted with such problems, two C++ plugins have been developed.
The Geometry Checker plugin allows the user to test a dataset for geometry and topology issues (such as duplicate nodes, overlaps, gaps, etc), presenting a list of detected faults. For each error type, the plugin offers one or more methods to automatically fix the issue.
The Geometry Snapper plugin on the other hand is designed for situations where geometries of a foreground layer (for instance, buildings) should be aligned with those of a background layer (for instance, parcels of land). This plugin ensures that all nodes of the foreground layer are snapped to the closest appropriate node of the background layer (within a user specified distance), adding and removing missing resp. superflous nodes as necessary.
This talk will focus on these two topics, discussing the benefits and limitations of the introduced functionalities as well as giving some insight into the technical background and implementation challenges.

References

✔ www.qgis.org

A RESTful API for linking geodata

Francesco Bartoli

Geobeyond Srl

Abstract

Publishing open data is a trend movement but still nowadays geographical information are often released as shapefile despite this common format isn't recommended for such scope.We have evolved GeoNode, a spatial data infrastructure for publishing open geodata through standard OGC Web Services, with a RESTful API to model such resources to the semantic interoperability. This allow to publish interlinked shapefiles as triple stores and search them with GeoSPARQL queries from a Virtuoso backend.

References

- ✔ CKAN – http://ckan.org
- ✔ GeoNode – http://geonode.org
- ✔ Virtuoso – http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/
- ✔ GeoLinkeData code repository – https://github.com/geobeyond/GeoLinkeData

Fast Cache, Fresh data. Can we have it all?

Thomas Ellett von Brasch

Kartverket

Abstract

Most national mapping authorities aim to provide a broad range of authoritative, up-to-date, easily accessable cached basemap services in accordance with industry standards. It is the adherence to these goals that set our cache services apart from the likes of Google and Bing, where a broad range of maps and update frequency have been neglected in order to acheive lightning fast, reliable, single services. Our responsibilities as a national provider of authoritative data require us to go further.

The problem for national mapping authorities is that the consumers now demand the best of both worlds. When accessing our services, they not only assume an authoritatve date source that is correct at the time of viewing, but also demand performance on a par with that offered by the market leaders. It’s telling that both the private and public sectors find it extremely difficult to meet these expectiations; superfast, reliable cache services with data that is as close to real-time accurate as possible.

This presentation will look at the Norwegian Mapping Authorities project to first find a solution to these probems, and then implement it through a completely opensource infrastructure. The project first came about as a result of increased user feedback concerning the cache services following the release of a new national mapping client, Norgeskart.no. We found that, primarily, the complaints were about the *stability* of the cache speed, not about the top speed itself. They encountered this problem because of the inherent issues in providing fresh data, namely an incomplete cache. The consumers expected our cache service quality indicators (speed, stability, uptime) to mirror those of the perceived commercial alternatives, and, due to various reasons, they were not.

The initial primary goal of the project was ‘to increase the speed stability of our cache services, creating an infrastructure that enables them to run as close to top speed 100% of the time during normal conditions’. But through an analysis of the main issues, it was clear that to meet the primary goal, we would need to solve 2 other central problems; incremental updating of the caches and preseeding the caches through all the zoom levels. Our solution was actually implemented from 6 tasks, but these all stemmed from solving the 2 central problems. The upgraded infrastructure still uses many of the same components as the old (PostGIS, Mapserver, geowebcache, linux) but these are now utilised in a different way and the infrastrucure on the whole is configured very differently. We hope that our solution can act as an example for those national mapping authorities facing the same issues that we have faced.

Development of Environmental Impact Assessment Tools Using Mobile Devices for Location-Based Services

SeungWoo Son, JeongHo Yoon

Korea Environment Institute

Abstract

A new mobile application system built on location-based services (LBS) was developed to overcome the limitation of EIASS (Environmental Impact Assessment Supporting System) that stores big data for environmental impact assessments because it was not possible to update ground-truth data collected from field surveys in a timely fashion. The new system allows users to directly transfer geographical coordinates, photos, text, etc. from fields to the main server of EIASS via mobile devices.

References

- ✔ Liu, J., Chen, R., Pei, L., Guinness, R., & Kuusniemi, H. (2012). A Hybrid Smartphone Indoor Positioning Soluiton for Mobile LBS. *Sensors,* 12(12), 17208-17233. doi : 10.3390/s121217208
- ✔ Environmental Impact Assessment Supporting System. (2010). EIASS. Retrieved March 3, 2015, from http://eiass.go.kr

Detection of potential updates of autoritative spatial databases by fusion of Volunteered Geographical Information from different sources

Stefan Ivanovic[1], Ana-Maria Olteanu-Raimond[1], Sébastien Mustière[1], Thomas Devogele[2]

[1] COGIT laboratory, IGN France, 73 Avenue de Paris, 94165 Saint-Mandé

[2] Université François Rabelais de Tours, 60 rue du Plat D'Etain, 37000 Tours

Abstract

Nowadays, needs for very up to date referential spatial data increase significantly. Thus, a continuous update of authoritative spatial databases becomes highly demanding task in both aspects, technical and financial. In the same time, alternative sources of spatial data, such as Volunteered Geographical Information – VGI (Goodchild, 2007) seems to be suitable solution. This data is easy available and is being collected in almost every moment somewhere in the world.

The main objective of our research is proposing a method for identifying potential updates in authoritative spatial databases using VGI data, more precisely GPS tracks. We identified walkway and tractor as very challenging types of roads for continuous update due to their intermittent nature (e.g. they appear and disappear very often) and various landscape (e.g. forest, high mountains, seashore, etc.). Even though, these types of roads are not of the highest priority for a national mapping agency, they are still very important for production of touristic maps and for other different applications such as defense, sport activities, etc. That is why we have focused on GPS traces obtained in sport activities.

To detect potential update, links between similar features need to be defined. This step consists in applying a data matching algorithm in order to match VGI and authoritative data. Then, the question of VGI tracks quality arises. Furthermore, VGI traces are collected without any specified procedures, less or inexistent metadata, usually by low class GPS devices. Hence, heterogeneity of data is very high as well as spatial inaccuracy. In this work we focus on examination of data quality, especially on its spatial and temporal aspects.

First, we present an overview of VGI data sources (websites) and the heterogeneities that characterize them. In terms of data, we can rely on spatiotemporal data (i.e. coordinates and sometimes elevation and timestamps) as well as on a variety of descriptive information in text format such as: type of activity, difficulty, trace description etc.

Second, providing a comprehensive analysis of elements which affect GPS data quality is necessary. Sources of errors related to technical aspect of GPS data

collection are partially important for our work. Since we use data obtained by low class GPS receivers, which positional accuracy is at meter level, we are not concerned about the sources that affect the accuracy at sub-meter level. Therefore, our attention is directed to identifying and classifying sources of errors according to which extent they affect positional accuracy of GPS tracks.
Finally, we are interested in evaluation of data quality by analyzing VGI data itself, without comparing it to referential data. Thus, we tend to obtain the more statistical indicators of data quality that we can, such as indicators of: spatial dispersion, precision, reliability, correlation between data etc. As a result, a process of automatic collection of GPS traces from web-sites and storing them into PostgreSQL database was created. Evaluation of data quality is conducted by using an open source platform GeOxygene[1], developed by COGIT laboratory.

References

- ✔ Goodchild, M.F. (2007). Citizens as sensors: The world of volunteered geography. GeoJournal, 69 (4), 211–221. doi:10.1007/s10708-007-9111-y.

1 Geoxygene: http://oxygene-project.sourceforge.net/

pycsw project status report

Angelos Tzotsos[1], Tom Kralidis[2] , Jeff McKenna[1]

[1] Open Source Geospatial Foundation

[2] Meteorological Service of Canada

Abstract

pycsw is an OGC CSW server implementation written in Python and has recently been accepted as an official OSGeo project. pycsw implements clause 10 (HTTP protocol binding - Catalogue Services for the Web, CSW) of the OpenGIS Catalogue Service Implementation Specification, version 2.0.2. Started in 2010 (more formally announced in 2011), pycsw allows for the publishing and discovery of geospatial metadata, providing a standards-based metadata and catalogue component of spatial data infrastructures. The project is certified OGC Compliant, and is an OGC Reference Implementation. pycsw is Open Source, released under an MIT license, and runs on all major platforms (Windows, Linux, Mac OS X). The project currently powers numerous high profile activities such as US data.gov/geoplatform.gov, Integrated Ocean Observing System (IOOS), US National Geothermal Data System (NGDS), National Oceanic and Atmospheric Administration (NOAA), and the WMO World Ozone and Ultraviolet Radiation Data Centre (WOUDC).

This session starts with a status report of the pycsw project, followed by an open question/answer session to give a chance to users to interact with members of the pycsw project team. We will go over the main features and enhancements made to the software in the last year (which includes the many new features of the pycsw 1.10 release), as well as the current and future direction of the project, the organization of the project and the role of the Project Steering Committee (PSC). Finally we will discuss contribution opportunities for interested developers and users.

Project Based Learning To Establish A Global Earthquake Forecast System

Ron Fortunato[1], Levi Purdy[2]

[1] Trillium Learning

[2] Kodiak Island Borough School District

Abstract

The State of Alaska engaged Trillium Learning to establish an *open model* for teaching high school students. This model has generated leading-edge research results involving earthquake precursor signals in one of the world's most seismically active areas. As part of establishing a Global Earthquake Forecast System (GEFS), the students acquire earthquake precursor data from sophisticated sensor arrays and develop real-time analysis of that data and then use NASA World Wind for visualization of that data. The *open model* project-based-learning incorporates real-time, real-world projects into the students' daily curriculum. This allows them to learn in a research and development environment.

The America Bridge Project team in Kodiak, Alaska is also developing their own ground station technology platforms to acquire continuous magnetic field information from remote locations. These data acquisition stations continuously transmit live data to servers, which the students have configured with multiple algorithms to recognize earthquake precursor magnetic field anomalies and then visualize this data in NASA World Wind. The data visualization represent the magnetic anomalies as 3-axes vectors. This will provide advanced warning for the location in the Earth's crust where tectonic plate stress is increasing.

The students are designing and building the GEFS ground stations to extend the sensor network so that can it be readily located anywhere on Earth. They are also designing a CubeSat (micro-satellite) version. Expansion of this continuous monitoring of the Earth's magnetic field will increase the area of coverage and thereby greatly enhance Alaska's Global Earthquake Forecast System, which can be the basis for a global earthquake forecast system.

The GEFS CubeSat platform and mobile unit will have a circuit board populated with sensors, including WiFi, GPS, gyroscope, accelerometer, barometer, compass, temperature and magnetometer. The portable ground stations, the GEFS Mobile units, will be placed at various locations throughout Alaska to acquire continuous data for a Global Earthquake Forecast System. Both the CubeSat and mobile unit design will be open platforms able to serve any type of sensor data.

NASA World Wind
Visualization Technology for Spatial Data

Patrick J. Hogan, Tom Gaskins, Paul [David] Collins
NASA Ames Research Center, Moffett Field, California, USA

Abstract

How does NASA World Wind facilitate innovative solutions in the FOSS4G world? By providing spatial information in its native context, we can accelerate our ability to understand that information. NASA World Wind for Java, Android, iOS and now JavaScript for use in web browsers, provides an open source virtual globe platform based on open standards. The possibilities for data processing, data analysis and contextual-based sharing of geospatial intelligence are all greatly enhanced by an open visualization platform for geospatial information.

Why World Wind? Over twelve years ago NASA World Wind began as a single program with specific functionality, to deliver NASA content. But as the possibilities for virtual globe technology became more apparent, NASA found that while enabling a new class of information technology, we were also getting in the way. Today World Wind is constructed entirely of modular componentry.

World Wind technology can be a part of any application. Or it can be extended with additional functionalities by application developers. World Wind makes it possible to include virtual globe visualization in support of any objective. As open source, the world community can collectively collaborate in advancing this technology, and thereby continually benefit from optimization and increased functionality of this open source infrastructure.

Open Source + Open Standards = Accelerated Solutions. NASA World Wind is NASA Open Source software. This means that the source code is fully accessible for anyone to freely use, even in association with proprietary technology.

World Wind Community. World Wind is used by many of the world's space agencies (NASA, ESA, JAXA, DLR, et al), various national governments and agencies such as the UN and NATO, along with several Fortune 100 companies, and many in academia, especially for scientific research. The World Wind user community is essentially composed of anyone who needs to manage spatial data in highly customized ways. World Wind was awarded NASA Software of the Year in 2009. And to a great degree, the world-class quality to its code base, is the result of the interaction by a sophisticated international cadre of users via the forum at www.WorldWindCentral.com. This website gives the world community a mutually supportive environment for application development using World Wind technology. Additionally, there are excellent World Wind tutorials online, generously provided by third parties.

ZOO-Project 1.5.0: news about the Open WPS Platform

Bozon Nicolas[1], Fenoy Gérald[1], Raghavan Venkatesh[2]

[1] GeoLabs SARL, FR

[2] Osaka City University, JP

Abstract

ZOO-Project is an Open Source Implementation of the OGC Web Processing Service (WPS), it was released under a MIT/X-11 style license and is currently in incubation at OSGeo. It provides a WPS compliant developer-friendly framework to easilly create and chain WPS Web services.This talk give a brief overview of the platform and summarize new capabilities and enhancement available in the 1.5.0 release. A brief introduction to WPS and a summary of the Open Source project history with its direct link with FOSS4G will be presented. Then an overview of the ZOO-Project will serve to introduce new functionalities and concepts available in the 1.5.0 release and highlight their interrests for applications developpers and users. Then, examples of concrete services chain use will illustrate the way ZOO-Project can be used to build complete applications in a flexible way by using the service chain concept, creating new service by implementing intelligent chain of service through ZOO-API but also by taking advantage of the publication using OGC standards. Various use of OSGeo softwares, such as GDAL, GRASS GIS, OrfeoToolbox, GEOS and pgRouting, as WPS services through the ZOO-Project will be illustrated by applications presentation.

GeoSmartCity: Open geo-data for innovative services and user applications towards Smart Cities

Lluís Vicens[1], Josep Sitjar[1] ,Gemma Boix[1], PiergiorgGio Cipriano[2], Roderic Molina[3]

[1] University of Girona, SIGTE,

[2] Sinergis,

[3] GISIG

Abstract

Smart City management requires integration of geographic data from many and heterogeneous sources, spanning from pan-European data sets (as the ones from the Public Sector Information and the INSPIRE Directives open data infrastructures) to local data with “home-made” semantics. In order to analyze and visualize geographic information (GI) through these data sets, it is necessary to integrate the data in terms of formats, access protocols, transformation and coordinate reference system, data harmonization.

The ICT-PSP European project GeoSmartCity establishes a cross-platform, able to publish open GI and to provide specialized services based on open standards services protocols. Starting by the availability of the open GI through open standards, the platform, or "Hub", gives the possibility to integrate them with other public/private data in order to design the specialized services needed to implement the two addressed Smart City scenarios: Green Energy and Underground.

Exploitation of heterogeneous (open) GI data will be possible thanks to the connection of different consolidated standards (linked data, INSPIRE, Sensor data, GNSS), allowing open cross-sector interoperability between different data providers and domains and the consequent creation of a wide range of user-driven applications.

The potentiality of the toolkit will be demonstrated through the development of 11 operative and re-usable pilot cases These pilot cases will demonstrate the possibility to apply the GeoSmartCity Hub and its specialized services in different areas and municipalities, guaranteeing the exploitability and the long term viability of the proposed solution, in new cities and sectors. The GeoSmartCity "Hub" will be community software and the business model will be based on development of extensions, applications and services. The non-proprietary platform will also foster increased emphasis on standardization and interoperability of detailed, local data that do not conform to generic standards for European data sets.

During the first year of the project, multiple tasks have been conducted. Until now, most of the work has focused on defining the system architecture as well as the definition of data and infrastructure management functions. The tangible results of this first year of the project are a set of deliverables and reports including the analisys of the uses cases and requirements collected, and a report on GeoSmartCity data consisting in an inventory of the GI data being used in the different pilots, not only in terms of their content, but also in terms of their structure, existing level of interoperability and accessibility. Moreover, an exhaustive analysis and collection of GeoSmartCity "Hub" technical requirements and Client technical requirements have been done.

Current work is focused on the definition of the Underground and the Green Energy scenarios data model, needed to harmonize the overall heterogeneous spatial datasets to be further managed by the "Hub", the definition and creation of specialized services to be used by the different scenarios and the set up the "Hub" architecture, hosting these data and specialized services.

Dissemination and training activities are also a very important task on the core of the project, in order to make available the know-how, the documentation and the expertise acquired within the project.

SHOGun as WPS-Client:
Orchestrating Geodata-Services, Users and Geo-Applications

Till Adams[1]

[1] terrestris GmbH & Co KG

Abstract

SHOGun[1] is an Open Source Webmapping Framework which consists of a server-side middleware and various client-side webmapping and administration clients.
The middleware is built out of the frameworks JAVA Spring and JAVA Hibernate, the clients are built with OpenLayers, GeoExt and ExtJS. There are Graphical User Interfaces for administrating users and groups, webmapping-services and also webmapping applications. SHOGun uses GeoServer for publishing, securing and styling geodata-services such as WMS and WFS.
SHOGun was originally developed for bigger organisations with technical administrative duties such as water supply and distribution management or soil protection for the federal state of Rhineland Palatinate.
Whereas terrestris as a technical service provider is able to support their customers with technical knowledge, there is no expertise in these fields. So terrestris may support by developping standard GIS-tools and services, but for developing technical specialized tools there is a big need for explaining by experts.
In order to enable the customer to develop their own experts-tools, a WPS-client was integrated into SHOGun. As WPS-Server the GeoServer-WPS extension is used, although theoretically SHOGun supports any WPS-Server.
Tools develped by anybody could now be integrated into GeoServer-WPS and then published into webmapping-clients within SHOGun and so also dedicated to users and user-groups.
The talk gives a brief overview about SHOGun and WPS and then demonstrates how any WPS-Service could be integrated into any webmapping-client solely by configuration.

1 Acronym for **S**pring, **H**ibernate, **O**penLayers, **G**eoExt and others.

Geo-SEE Institute activities for using FOSS through the Quantum GIS training courses

Bashkim Idrizi[1], Radife Neziri[2]

[1] State University of Tetova

[2] Geo-SEE Institute

Abstract

Beginning from September 2013, Geo-SEE Institute started with professional training courses for using free and open source software (FOSS) in a Geographic Information (GI) field, as well increasing understanding and awareness for the benefits of using FOSS in public institutions, private sector, NGO's and research/educational institutions.

From the list of FOSS for GI, Quantum GIS has been selected as one of more utilized software. Process of software selection has been powered by survey for needs, current conditions, annual budget, organizational tradition, laws and regulations for geospatial data, as well GI demand and supply. Some of governmental organizations, local government sectors, educational institutions, private companies and NGO's as identified stakeholders have been asked to fill the survey questionnaire.

First professional training courses have been started in Macedonia. In cooperation with local partners in Albania and Kosovo, Gjeo Vjosa from Tirana and Geo Institute Kosova from Prishtina, professional training courses extended its geographic area in the Republic of Kosova and the Republic of Albania.

Professional training course on Quantum GIS is designed in three levels, systemized hierarchically from the process of data developing up to the web GIS, in Albanian and Macedonian language, while the training programme is available in English language also. Lectures duration is 20 academic hours per level, by using the latest version of Quantum GIS, actually version 2.8.1.

Actually the process of evaluation is being powered in order to get feedback from former training participants. Results will be used for analyzing the correlation of market requirements with training programme and obtained knowledge in three levels of training courses, in order to analyze potential changes/updates of training programme and training materials.

Assessing social vulnerability to earthquake hazard: from statistical to spatial analysis

Ivan Frigerio[1], Daniele Strigaro[1], Matteo Mattavelli[1], Silvia Mugnano[2]
Mattia De Amicis[1]

[1] Geomatic Laboratory, DISAT Department, University of Milano-Bicocca

[2] Sociology and Social Research Department, University of Milano-Bicocca

Abstract

In the framework of classical natural hazard, multi-hazard and risk assessment the concept of vulnerability is referred to the fraction of the total risk value that could be loss after a specific adverse event (Mazzocchi et al., 2009). However, over the last few years the term 'Vulnerability' has been frequently cited in scientific literature in regard to different context, focusing particularly to social-economic aspects that influence societal conditions such as exposure, sensitivity, coping, adaptive capacity and social capital (Adger, 2006; Gallopin, 2006).

Indeed, the natural hazards does not have a random effect on the local community and generally the most affected groups are the more vulnerable ones, already marginalized by class, race, ethnicity and gender (Blaikie et al., 1994). Therefore, natural hazards can be more or less devastating according to vulnerability, which depends on the time and place where the event happens and the socioeconomic conditions of the population affected (Cutter et. al., 2003).

Within this framework, the main purpose of this work is to assess social vulnerability indicator (SVindex) toward earthquake hazard for Italian country in order to identify hot-spot areas: zones with high seismic levels and at the same time high social vulnerability levels. The methodological approach consist in four different steps:

1. the first is the development of a spatial database using Spatialite with different tables coming from 2001 ISTAT population census. The database is used to increase data management and to create several socioeconomic indicators (e.g. age, education, gender) as proxy variables that better explain the Italian population socio-economic conditions that influence the capacity of a community to prepare for, respond to, and recover from hazards and disasters;
2. the second step is to apply on the proxy variables a multivariate statistical analysis. Through different **R** packages (e.g. vegan, princomp, outliers) principal component analysis (PCA) is performed to reduce dataset. At the end of the PCA, the interpretation of the component matrix generated 4 main factors that explain the relationship between

variables unfolding the 74.6 % of the variance in the entire dataset. These components are interpreted as the follow: **age**, **employment**, **education** and **anthropization**;

3. the third step consist in the aggregation of the 4 components through an additive model to create SVIndex and mapping it using QGIS software;
4. the last step is to produce an exposure map, combining SVIndex with earthquake hazard map, developed by INGV (National Institute of Geophysics and Volcanology) (Gruppo di Lavoro MPS, 2004) to assess the spatial relationship between social vulnerability and seismic hazards.

The exposure map can be integrate in the emergency plans in order to better allocate resources such as people, materials and financial founds, in response to improve emergency management against disaster events.
This study allows also to identify appropriate cost-effective risk reduction measures to be implemented at a local level; to help territorial planners in managing the relationship between natural processes and human communities using FOSS softwares. Finally, the lack of social studies in the hazard and risk assessment highlights the need to better integrate social science research concerning social vulnerability into emergency and risk management decision-making.

References

- ✔ Marzocchi, W., Mastellone, M.L., Di Ruocco, A., Novelli, P., Romeo, E., Gasparini, P. (2009). Principles of multi-risk assessment: interaction amongst natural and man-induced risks (Project report). Office for Official Publications of the European Communities, Luxembourg.
- ✔ Adger, W. N. (2006). Vulnerability. Global environmental change, 16(3), 268-281.
- ✔ Gallopín, G. C. (2006). Linkages between vulnerability, resilience, and adaptive capacity. Global environmental change, 16(3), 293-303.
- ✔ Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (2014). At Risk II-: Natural Hazards, People's Vulnerability and Disasters. Routledge.
- ✔ Cutter, S., Boruff, B., & Shirley, L. (2003). Social Vulnerability to Environmental Hazards. Social Science Quarterly, 242-261
- ✔ Gruppo di Lavoro MPS (2004). Redazione della mappa di pericolosità sismica prevista dalla Ordinanza PCM del 20 marzo 2003 n. 3274, All. 1. Rapporto conclusivo per il Dipartimento della Protezione Civile, aprile 2004, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Milano-Roma, Italy, 163 pp. Retrieved from http://zonesismiche.mi.ingv.it/

WebGL technology for three-dimensional geological data visualization using three.js JavaScript library

Cesare Comina[1], Andrea Filipello[1], Giuseppe Mandrone[1] , Luciano Masciocco[1], Daniele Strigaro[2]

[1] Earth Sciences Department, University of Turin

[2] Geomatic Laboratory, University of Milano - Bicocca

Abstract

Geological investigations for natural hazard assessment or mining prospecting are based on outcrop survey and on underground data (geophysical survey, boreholes; Guglielmetti et al, 2013). The goal of geological study is to build a model including topography, stratigraphy and geological structure distributions.

Usually 3D geological models are based on few data if compared with the extension of the study area. The available geological data are sparse and a major interpretation it's often necessary. The three-dimensional visualization allows to help geological interpretation about this lack of data. In this study the plugin Qgis2threejs is used to represent boreholes, geophysical surveys and monitoring system (piezometers and inclinometers) carried out to study a landslide that affects a wide slope, placed in the Monferrato Hills, in North-West of Italy. This plugin exports terrain data, map canvas image and vector data to a web browser using the WebGL technology (Krooks et al., 2014).

Boreholes data are plotted as box objects by specifying depth and height for each geological unit in the attribute table. The boreholes are categorized according to the stratigraphic column. Piezometers and inclinometers are plotted like boreholes, but only with indication of depths for water table or measured displacements.

Geophysical data consist on 2D electrical resistivity image of subsurface. During data acquisition, the position of every electrodes is measured with a GPS. After processing data, the inverted true resistivity 2D section is georeferenced as points in a vector layer with the X, Y, Z coordinates and the resistivity values in the attribute table. Geophysical data are shown using a graduated colors scale according to the distribution of resistivity values.

High resolution Digital Terrain Model (5 meter), provided by the Piedmont Administration under the terms of Creative Commons public license 2.5, is used as topography.

Finally, a new surface model of the area is computed using photogrammetry. A digital camera (Pentax WG-3), connected to a helium balloon and flying at an approximate height of 40 meters, has taken about 350 pictures, using a time-lapse of 10 second.

IGN'S open source photogrammetric suite Apero-MicMac is used to do a bundle adjustment with ground control points (Deseilligny et al., 2011). The workflow consists in different steps (Deseilligny, 2013): automated tie point extraction; bundle adjustment for camera parameters derivation; transformation image relative orientations into an appropriate coordinate system; dense image matching for surface reconstruction; DSM and orthoimages generation. Resulting DSM and orthoimages obtained are used directly into QGIS, while the resultant point clouds are imported in CloudCompare for visualization, scaling, mesh generation and measurements.
Qgis2threejs has proved a tool to easily generate intuitive and suggestive visualizations of geodata in 3D. Another important feature is the possibility to share our visualization, just copying the html and associated Javascript files onto a webserver. One of main existing limits is inability to make any measurement of distance, area or volume directly on 3D.

References

- ✔ Krooks, A., Kahkonen, J., Lehto, L., Latvala, P., Karjalainen, M., & Honkavaara, E. (2014). WebGL Visualisation of 3D Environmental Models Based on Finnish Open Geospatial Data Sets. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1*, 163-169.
- ✔ Deseilligny, M. Pierrot, Clery, I. (2011). Apero, an open source bundle adjusment software for automatic calibration and orientation of set of images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2011, 38: 5.
- ✔ Deseilligny, M. Pierrot (2013). MicMac, Apero, Pastis and Other Beverages in a Nutshell. *Documentation de la suite Micmac.*
- ✔ Guglielmetti L., Mandrone G., Schil E., Comina C., Negro F. (2013) *3D geological modeling as a tool for geothermal prospection in the argentera massif (South-western Alps)*, EGEC 2013, EGC, Pisa, pp. 1- 8.

Improving public health delivery in northern Nigeria using open source technologies

Sonoiki Dami, Kazeem Owolabi[2]

eHealth Africa

Abstract

Access to adequate health care in developing countries is highly dependent on having the right information. This could range from planning how to effectively distribute drugs to health facilities or understanding the network signal strength of around health facilities and deploying the appropriate mobile phone network to enable uninterrupted information sharing.

Nigeria is at the center of polio eradication in Northern Nigeria and since 2014, West Africa has been combating the Ebola outbreak. eHealth Africa is an NGO based in Nigeria that focused on these areas by creating effective ways of implementing reliable health information management systems, and has a strong emphasis on utilizing GIS and geospatial solutions.

FOSS4G solutions play a substantial role in the various health projects that eHealth Africa is assigned to. The organization has a team of about twenty GIS specialists in place to provide mapping, data collection, and spatial analytics. To support the polio eradication activities, the northern half of Nigeria has been mapped and systems have been set-up to follow vaccination teams. To support the Ebola outbreak response teams, mobile equipment is being used to perform case surveillance, and mapping applications are developed in order to provide accurate insight for relief workers.

FOSS4G software used is supporting altogether their GIS service by giving access to up-to-date data for diverse internal and external client’s need. The main advantages of this model are flexibility in serving data / developing solutions, interoperability with both FOSS and proprietary software, open to multiple sources of data (open data, API, government, private companies, local authorities), web distribution of service in different locations, full integration of different types of data for cloud deployment environment.

PyWPS 4 - The new Python implementation of the Web Processing Standard 1.0 and 2.0

Luís de Sousa[1], Calin Ciociu[2], Ulrich Leopold[1]

[1] Luxembourg Intitute of Science and Technology (contact: luis.desousa@list.lu)

[2] University of Luxembourg

Abstract

PyWPS[1] is an open source, light-weight implementation of the Web Processing Service (WPS) standard, issued by the Open Geospatial Consortium (OGC). It allows users to code geo-spatial analysis processes with the Python programming language and publish them over the internet through a CGI server such as Apache. The option for the Python language makes the usage of geo-processing libraries, such as GRASS[2], seamless and easy to learn.

The development of PyWPS started almost a decade ago, its general philosophy and the tools on which it relies are starting to appear dated today. The coming release of PyWPS (version 4) is a complete re-write of the code, taking advantage of modern libraries and tools. PyWPS 4 uses the Werkzeug utility library[3] to expose user processes to the internet through the WSGI 1.0[4] standard. PyWPS 4 also relies on Flask[5], a light-weight web application framework, for the core its HTTP request-response processing. The Tox automation library[6] is used to manage the virtual running environment and streamline the test suite. In this release, PyWPS is relying on the Multiprocessing library for asynchronous process execution.

At this stage PyWPS 4 is already fully compliant with version 1.0 of the WPS standard. Compliance with WPS 2.0 is targeted for the next iteration (PyWPS 4.1), which will require a process queuing and message passing mechanism. Each process run must be traceable and open to asynchronous messaging, in order to deal with the new requests specific by the standard: PAUSE, CANCEL and RESUME.

Even if it already allows the coding of processes and their exposure to the web, further goals have been set for PyWPS 4 to augment its functionality. An

1 http://pywps.wald.intevation.org/
2 http://grass.osgeo.org/
3 http://werkzeug.pocoo.org/
4 http://wsgi.readthedocs.org/
5 http://flask.pocoo.org/
6 https://testrun.org/tox/latest/

administrative REST interface is projected to facilitate the publication of new processes and eventually allow request submission. The automatic publication of Complex outputs (i.e., geo-spatial data) using OGC services enabled third party software (e.g. MapServer) is another goal. A further release (PyWPS 4.2) will target more advanced features such as transactional WPS.

PyWPS 4 is now undergoing a testing phase to which the community is invited to contribute[7]. The enrichment of the test suite will also attest the interoperability with other OGC services (WFS and WCS) and the integration with geo-processing tools such as GDAL/OGR[8] and GRASS.

With version 4 PyWPS is now officially in the incubation process towards accreditation as an OSGeo project.

PyWPS 4 is being developed with funding from the MUSIC Interreg IVB project[9] and from the Google Summer of Code 2015 at the Luxembourg Institute of Science and Technology (LIST).

7 The code is kept at a public repository: https://github.com/jachym/pywps-4
8 http://www.gdal.org/
9 http://www.themusicproject.eu/

Smart City Platform - An OGC based decision support platform for Smart City Planning to stimulate Open Data in Urban areas

Matteo De Stefano, Luís de Sousa, Ulrich Leopold, Christian Braun, Rui Martins

Luxembourg Institute of Science and Technology (contact: matteo.destefano@list.lu)

Abstract

The Smart City Platform was developed to provide a common approach to urban energy planning for different partner cities in the Interreg IVB MUSIC project[1]. The system was designed to manage both input data and processing algorithms from multiple sources, providing geo-spatial analysis outputs relevant in the project context.

To meet the project requirements, the iGUESS (integrated Geospatial Urban Decision Support System) software, based on open source technologies, was developed (de Sousa, 2012). It is a web-based, spatially enabled interface for decision support, relying on a distributed and service-oriented architecture. It focuses on integration and interoperability, achieved by a systematic adoption of the OGC open standards. In particular, IGUESS uses the WFS and WCS standards to access and publish geo-spatial distributed datasets, the WPS standard to interact with remote geo-spatial processes, and the WMS standard to visualize data and results on a Web GIS. iGUESS is an open source product itself[2].

Beyond the output of the processing algorithms, the system provides various outputs suitable for decision support, such as urban planning scenarios and modelling tools to asses the impacts of CO2 emissions mitigation actions in space. IGUESS offers an integrated interface and approach to the urban planning process: data collection, data processing, modelling, decision assessment and presentation of results.

The option for a standard-compliant and service-based architecture resulted in a flexible, modular framework, independent of the underlying models and data it accesses. The distributed context which shaped its development makes the current iGUESS system relatively easy to extend to new contexts, such as new cities requiring an integrated tool for decision support on urban planning or other geo-spatial dependent planning activities.

While the platform was initially developed around energy planning models, the

1 http://www.themusicproject.eu/
2 Available at the public repository: https://github.com/ERIN-LIST/iguess

system can easily be adapted to new planning and decision support contexts, simply by registering new WPS processes and datasets from any available source. The system was succesfully applied in the Last Mile Logistics project[3], and is being tested on other domains, such as hydrology management. The system can be considered as a broad purpose platform, supporting smart cities planning and decision making, that makes the most of geo-spatial data and standards.

References

de Sousa, L., Eykamp, C., Leopold, U., Baume, O. & Braun, C. (2012). iGUESS - A web based system integrating Urban Energy Planning and Assessment Modelling for multi-scale spatial decision making. *Procedings of the International Congress on Environmental Modelling and Software (iEMSs)*.

3 http://www.lamiloproject.eu/

Real-time projection visualisation with Indicatrix Mapper QGIS Plugin

Ervin Wirth[1], Péter Kun[2]

[1] PhD student, Department of Photogrammetry and Geoinformatics, Hungary, Budapest University of Technology and Economics

[2] GIS developer, Institute of Geodesy, Cartography and Remote Sensing Hungary, Budapest

Abstract

The Indicatrix Mapper plugin created by Wirth & Kun presents the tiss-indicatrix. The tiss-indicatrix term is derived from the Tissot-ellipses, it is a quick-and-dirty Tissot-indicatrix realization introduced by Szabó & Wirth. The tiss-indicatrix uses circles of constant radius instead of the original infinitesimal circles (Tissot-indicatrix). (Goldberg & Gott, 2007)

The plugin projects the tiss-circles from a reference sphere to a selected projection. The software uses On The Fly (OTF) transformation method. Then we can study the distortions of the circles in a blink. The QGIS contains approximately 2 700 categorized projections.

On Figure 1 the reader can see the exported QGIS layers as KML (Keyhole Markup Language) files imported to Google Eath Pro.

The Figure 2 shows the mentioned layers in QGIS Wien software enviroment with thematic worldmap about countries in the background.

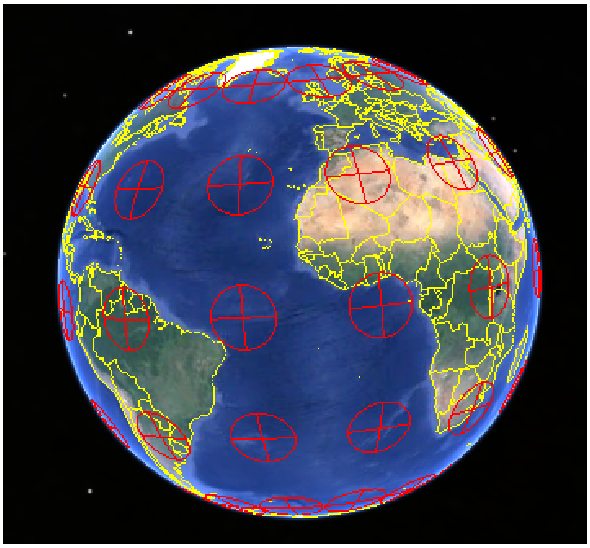


Figure 1 : the 800 km tiss circles on a reference quadric (from Google Earth Pro)

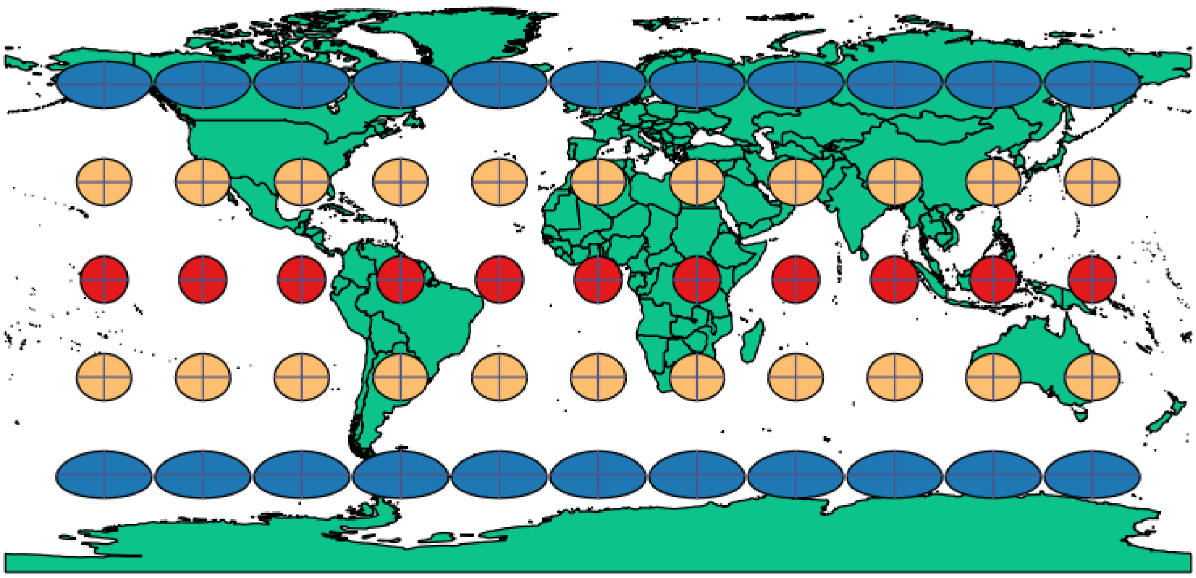


Figure 2 : tiss circle and line layers in QGIS 2.8 in WGS84 global geographic coordinate system (EPSG:4326)

Then by opening CRS (Coordinate System) settings we can transform our layer on the fly (OTF) with the support of PROJ4 library. We can do investigations on any projections from the popup list in a blink by selecting. Figure 3 and 4

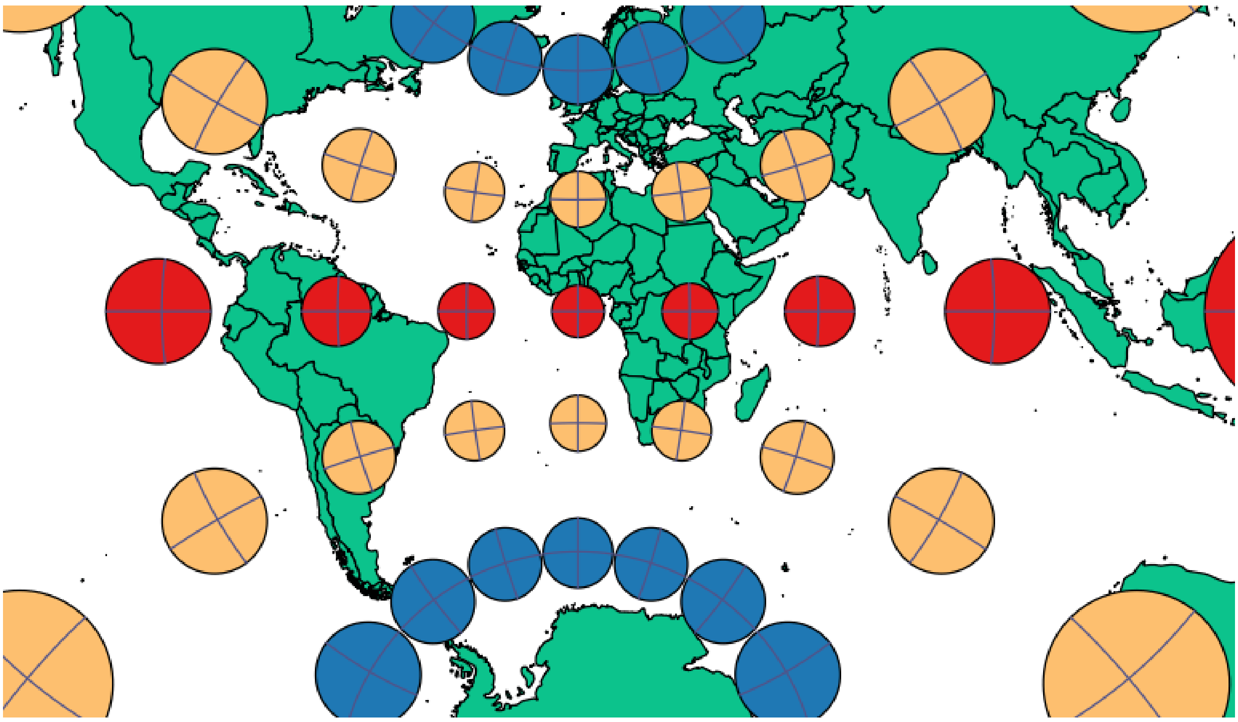


Figure 3 : after OTF in a stereographic projection (Sphere Stereographic, EPSG:53026)

In Figure 3 the reader can see that the transformations of the circles are still circles, thus stereographic is a conformal projection, meaning that it preserves angles.

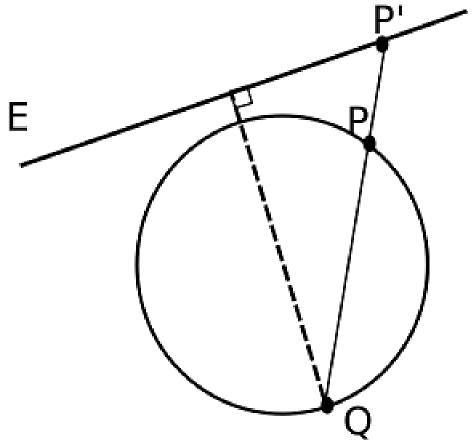


Figure 4 : stereographic projection of a sphere from a point Q onto the plane E, shown here in cross section

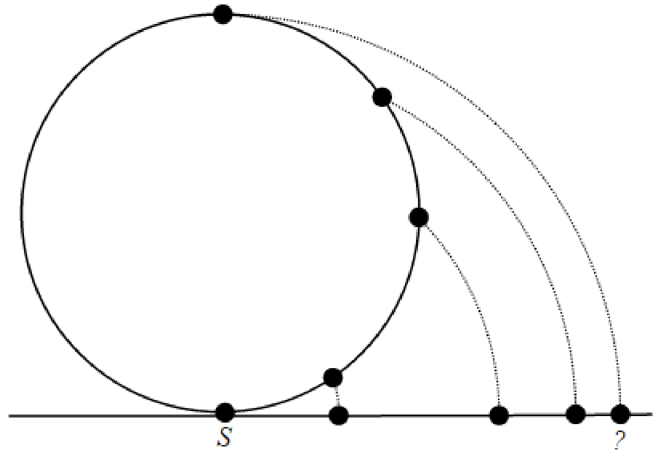


Figure 5 : a cross sectional view of the sphere and a plane tangent to it at S. Each point on the sphere (except the antipode) is projected to the plane along a circular arc centered at the point of tangency between the sphere and plane.

Figure 4 illustrates the procedure of projecting happened in Figure 3. After switching the QGIS project coordinate system on the fly to a Lambert azimuthal equal-area projection (Figure 5) loads the map of Figure 6.

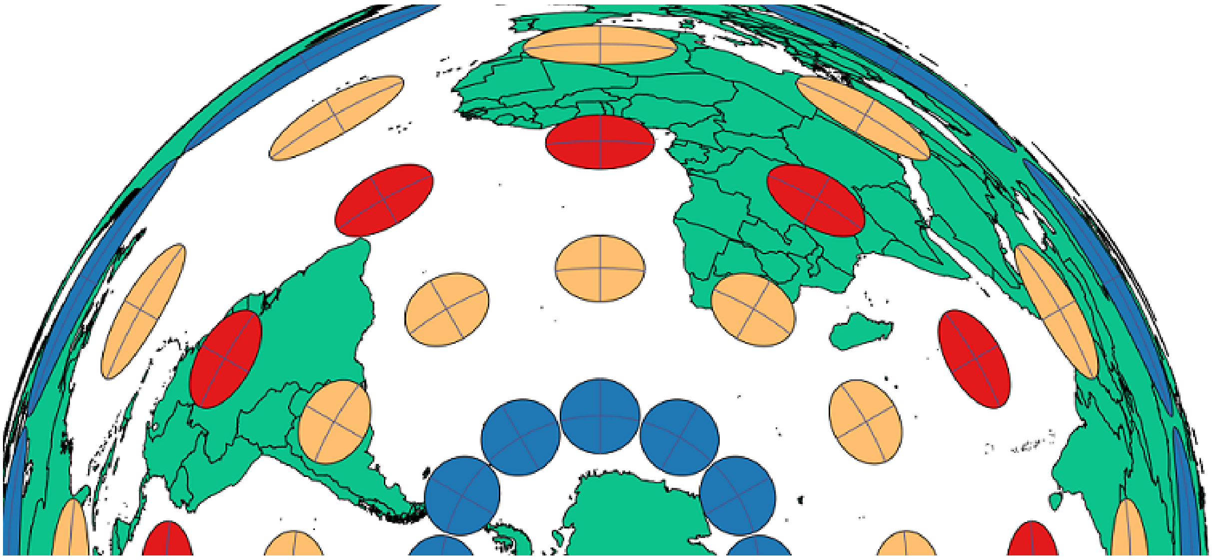
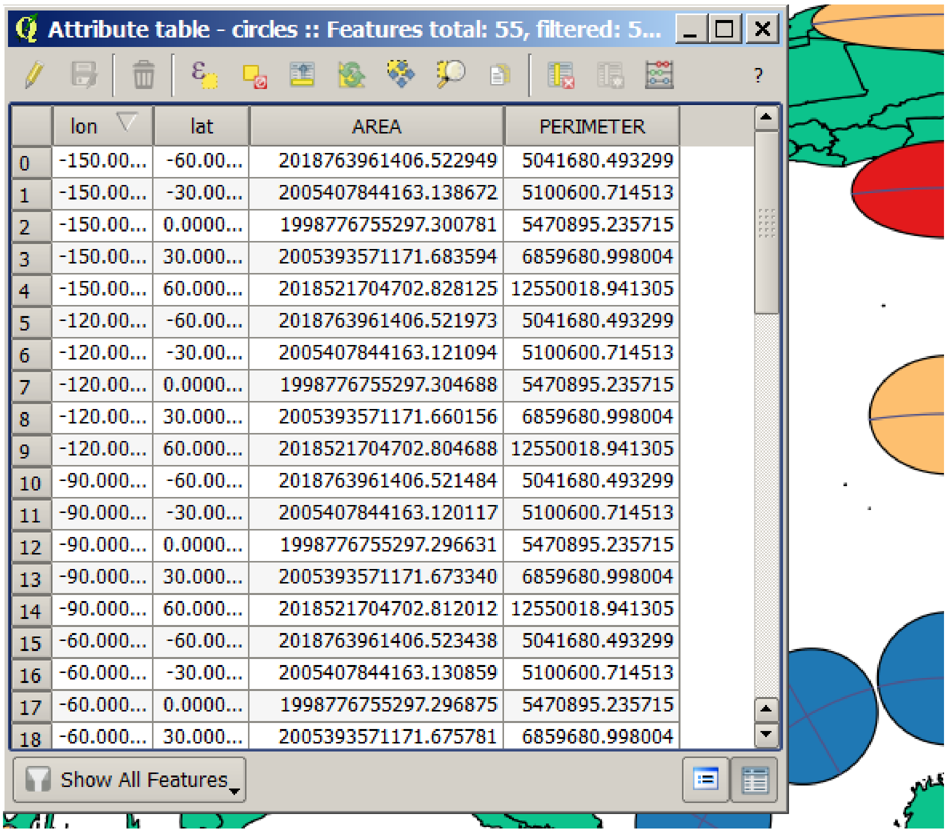


Figure 6 : South Pole Lambert Azimuthal Equal-Area Projection in QGIS (EPSG:102020)

Finally we can calculate the areas of polygons in the given project coordinate system to proof it's equal-area being.



Attribute table - circles :: Features total: 55, filtered: 5...

	lon	lat	AREA	PERIMETER
0	-150.00...	-60.00...	2018763961406.522949	5041680.493299
1	-150.00...	-30.00...	2005407844163.138672	5100600.714513
2	-150.00...	0.0000...	1998776755297.300781	5470895.235715
3	-150.00...	30.000...	2005393571171.683594	6859680.998004
4	-150.00...	60.000...	2018521704702.828125	12550018.941305
5	-120.00...	-60.00...	2018763961406.521973	5041680.493299
6	-120.00...	-30.00...	2005407844163.121094	5100600.714513
7	-120.00...	0.0000...	1998776755297.304688	5470895.235715
8	-120.00...	30.000...	2005393571171.660156	6859680.998004
9	-120.00...	60.000...	2018521704702.804688	12550018.941305
10	-90.000...	-60.00...	2018763961406.521484	5041680.493299
11	-90.000...	-30.00...	2005407844163.120117	5100600.714513
12	-90.000...	0.0000...	1998776755297.296631	5470895.235715
13	-90.000...	30.000...	2005393571171.673340	6859680.998004
14	-90.000...	60.000...	2018521704702.812012	12550018.941305
15	-60.000...	-60.00...	2018763961406.523438	5041680.493299
16	-60.000...	-30.00...	2005407844163.130859	5100600.714513
17	-60.000...	0.0000...	1998776755297.296875	5470895.235715
18	-60.000...	30.000...	2005393571171.675781	6859680.998004

Figure 7 : after calculating geometric indices (area, perimeter). The attribute table convinces us about keeping area after projection, because the perimeter varies besides area isn't.

As conclusion we can say that QGIS Indicatrix Mapper plugin can give a new approach in map projection education instead of representing neverending formulas.

References

✔ Goldberg & Gott, Flexion and Skewness in Map Projections of the Earth, 2007, www.physics.drexel.edu/~goldberg/projections/goldberg_gott.pdf

MapServer Project Status Report

Thomas Bonfort[1], Stephan Meißl[2]

[1] thomas.bonfort@gmail.com, Terriscope EURL, France

[2] stephan.meissl@eox.at, EOT IT Services GmbH, Austria

Abstract

2015 is again a big year for the MapServer project with the finally forthcoming release of the 7.0 major version. This presentation highlights the new features included in this version, like WFS 2.0 for Inspire, UTFGrids, or heatmaps, as well as a recap of the main features added in recent releases including the addition of MapCache and TinyOWS. It further shows the current and future directions of the project and discusses contribution opportunities for interested developers and users.

After the status report of the MapServer project there will be the opportunity for users to interact with members of the MapServer project team in an open question/answer session. Don’t miss this chance to meet and chat face-to-face with members of the MapServer project team!

OPENGIS.ch Qfield - A native user interface for mobile touch devices.

Marco Bernasocchi[1], Matthias Kuhn[1]

[1] OPENGIS.ch

Abstract

The ubiquity of mobile devices has seen a massive increase in the last years. With more than 2 billion mobile devices before 2014 [0] and a constantly growing market hand held devices also gain importance in business and work.

Intuitive interaction,low weight, integrated positioning and other sensors and a (comparatively) low price make the acquisition of mobile devices easier than ever and a pleasure to use.

Thanks to its multi-platform nature (Windows, Mac, Linux, Android) and its broad feature-set (Desktop, Server, Web-Client) QGIS offers a broad range of applicability in different places in the stack of geospatial software. The complement with a native touch user interface that integrates nicely with this software suite adds an additional important piece to tackle everyday challenges.

Based on the experience gained in the development of QGIS for Android we have realized what is required for mobile apps. More than that we have realized what to absolutely avoid: Complexity, small UI elements and project definition work on mobile devices.
Our paradigm in the development for the new UI therefore is: “Less is more”.

With predefined modes (Data acquisition, data validation, measurements...) and clear user interface elements the user can therefore concentrate on his work in every situation.
To further ease the work we are developing a new offline synchronization tool that allows a seamless data exchange between your mobile devices and the available infrastructure.

Demo und further information: http://www.opengis.ch/download/qgis/

References

✔ [0] http://www.gartner.com/newsroom/id/2645115

Geographic information for the management of flood risk insurance

Franco Guzzetti[1], Alice Pasquinelli[1], Paolo Viskanic[2]
[1]Dip. ABC Politecnico di Milano – Via Ponzio, 31 – 20133 Milano
[2]R3 GIS S.r.l. – Via Kravogl, 2 – 39012 Merano (Bz)

Abstract

This paper describes a research project aimed at introducing geographic information in the management of flood insurance policies through GIS technologies. The main scope is the development of a toolbox to take geographic information into account in the calculation of flood insurance premiums.

Several factors influence real estate flood vulnerability: building's exposure depends on territorial aspects, like the presence of flood areas, the elevation profile, the position of the building compared to hazard zones, but it is also determined by the construction characteristics of the building itself.

In a pilot project workflows were set up to manage the whole underwriting process, following a rational methodology and taking into account all variables influencing building flood vulnerability. The developed tools are all based on free and open source software, server-side as a WebGIS tool and API and client-side as an Android application.

Through a more accurate evaluation of the risk, the insurer is allowed to better manage its risk exposure and this way guarantee solvency in the case of a flood event and on the other hand be more competitive on the market.

Keywords

Open source software, hydraulic risk, GIS, flood insurance

1. Introduction

In Italy the demand for flood insurance comes mainly from companies, to insure their economic activities. In the past the Government has contributed to cover damage costs caused by natural calamities. The tendency however is to encourage owners to ensure their properties, since the Government is not in the position to cover all damages, especially in an environment where extreme weather situations tend to become more frequent due to climate change.

As a consequence insurance companies offer currently this kind of coverage only to selected costumers, with important insured values through tailor-made policies. However due to the changes in Government policy, insurers will soon be required to extent this coverage to the mass market. Since late 2013 the Italian Government stated that costs due to natural disaster cannot be borne exclusively by the state anymore and the orientation is to introduce soon a public-private cost sharing system involving the insurance industry. The new challenge for insures is to provide flood coverage policies with prices that allow, on the one hand, to guarantee solvency in case of event (as required by European Directive 2009/138/CE "Solvency II"), and on the other hand to be

competitive in the market.
Given this environment, the introduction of geographic information in the risk assumption process can be an opportunity to assess objectively the exposure of assets, define a risk-based competitive pricing, and improve the safety of the portfolio by balancing the real exposure.
For these reasons at the beginning of 2014 major Italian insurance company started a pilot project with the Politecnico di Milano for the creation of a thematic GIS on hydrological risk. A research group was created, composed by insurers, the University and a software company specialized in geospatial management tools, with the aim of testing the use of geospatial technologies in the evaluation of assets flood vulnerability and the correct estimation of policy premiums. The project was organized in three parts:

1. definition of parameters that have to be taken into account for the correct estimation of insurance premiums
2. data retrieval, analysis and standardization;
3. creation of a prototypal web/mobile application to be used by insurers when evaluating the vulnerability of an asset.

2. Risk assessment methodology

Several factors influence flood vulnerability of buildings and policy premiums should reflect the degree of vulnerability observed for each property: for this reason it is important to clearly define which aspects to consider and which data have to be collected in order to proceed with a fair and comprehensive risk-based pricing for flood insurance.

Figure 1: Italian Basin Authorities.

First of all the information about risk sources has to be considered. In Italy Water Basin Authorities are the administrative bodies responsible for water management and hazard area identification. Each authority have deliberative and financial autonomy and develops a Basin Plan to enforce its policies. In these plans hazard areas are mapped and classified but there are no common rules for the hydraulic modeling of the river network and for the classification of risk areas. Consequently flood risk area in Water Basin Plans are defined without a standard and without common rules for a coherent and homogenous mapping at national level.

As required by the European Flood Directive 2007/60/CE a national standardized flood risk database is being developed and is expected for June 2015. However when this project started this unified and standardized national map was missing. Therefore in order to obtain a usable country-wide data layer, a detail study of the methodologies applied by each Water Basin Authority and a standardization of these maps was necessary.

The second step in order to assess building flood vulnerability is the overlay of the insured asset with the risk map. This is a quite logical and easy operation with GIS tools, however the success depends on the accuracy of the asset position. While in territorial planning and management sector the geographic location of items, like hazard areas, is a common practice, building positions are normally expressed through an address. Moreover, in environments without a culture of geographic data, addresses are often unstructured, partial and out of date.

The transition from the address to the coordinate pair for each building requires a cultural shift towards a more geographic approach: while addresses can be unstructured, incorrect and change over time, a coordinate pair is unique worldwide and doesn't undergo any modification. This transition can be operated automatically, geocoding existing addresses or, for new acquisitions identifying a property directly from a map or in the field through a GPS. Geocoding can be subject to errors, depending on the geocoding engine used and on the quality of addresses, but is the only method available for historic databases. The exact positioning in the field by GPS can be an option for new acquisitions, especially for tailor-made high value policies.

The third step has to do with physical aspects, like ground height compared to the possible flood height, altitude, distance from to the nearest river. These and other parameters may affect building exposure to flood. All these information are easily computable using a Digital Terrain Model, which are not available as open data in all regions. In any case, when available this data is useful to enrich the analysis on building position and its interaction with risk areas.

A last parameter which influences the exposure of buildings is the way they are constructed: buildings located in risk areas may register different levels of damage depending on construction characteristics. Normally, the most affected part in case of flood are underground and ground floor, while higher floors remain undamaged; in some cases the ground floor might not be vulnerable since the building is constructed on piers for example. Building exposure can also be reduced thanks to active and passive flood proofing measures like sealing material, shields for openings, backflow valves, inflatables barriers, or protecting the insured assets from water. The survey of such information requires an on-site inspection by insurance agents that might be associated to the GPS survey of the position. As previously mentioned this option involves costs and time that have to be commensurate to the actual need: a preliminary comparison between building position (and so its insured value) and hazard areas will suggest whether or not to proceed with an on-site survey. If the property is located in a safe place no additional investigations are required; on the other hand, if the building falls within a hazard zone, this option should be taken into account, especially in case of high value assets.

All factors above influence building flood vulnerability and have an effect on the insurance premiums This step by step analysis determines the methodological procedure to follow for the general assessment of building vulnerability:

- the collection of spatial data about territorial hydraulic hazard zones,
- the definition of building position and the overlay with the hazard areas,
- the collection of detailed information on the characteristics of properties.

The pilot project developed software tools to carry out all these steps when subscribing a new insurance policy.

3. Data collection, standardization and processing

As previously mentioned geographical information about risk sources and building position are required for a correct assessment of building flood vulnerability. As this information is not immediately available in a structured and standardized format, the following paragraphs describe the analysis and standardization operations conducted on hydraulic hazard areas and the georeferencing process carried out on a significant sample of insurance policies.

3.1 Hazard Areas data

The general mapping of flood hazard areas in Italy at the beginning of the project here presented (December 2013) consists of many tiles composing a national puzzle: every Water Basin Authority produced its own Basin Plan without sharing classification methodology. Therefore the classification of hazard areas differs in each watershed. Since these tiles were generated independently the assembly process requires some operations to create a nationwide map: standardization and edge-matching operations were necessary.
Within the adaptation process related to the Flood Directive, the Ministry of Environment released common rules defining a proper procedure to move from an heterogeneous classification to a unified legend. Basin Authorities are required by June 2015 to convert their classification following these rules and, when possible, to increase the level of detail. Due to the timing of the project the team had to proceed and carry out the standardization on its own.
Generally speaking hazard areas in Basin Plans are identified through hydraulic models which simulate floods in terms of recurrence and water load considering existing water control devices: as a consequence some plans identify hazard areas predicting both these factor, also focusing on critical parts of the water network, while some other provide just the return period information. In some other plans no indication about the computation of hazard areas are mentioned. Moreover, some Basin Authorities compute hydraulic analysis just for the main rivers, while others extended it to the whole hydrological network. It is important to consider the level of detail distinguishing Italian Water Basin Plans: territories may appear free from dangers when the mapping of hazard areas is not complete or rigorous, but this does not mean that in real world those territories are safe.
Hazard areas shapefiles were retrieved from each Basin Authority: in some cases data were easily downloadable from official websites while in other cases formal request to the administrations were necessary to access this information. Even if the contents of Hydraulic Plans are formally public, their accessibility is not immediate, reflecting the current “openness” of spatial data in Italian public administration. Data manipulation consisted in harmonisation, following ministerial directives, and geometrical geo-processing aimed to remove overlaps and gaps between hazard areas.

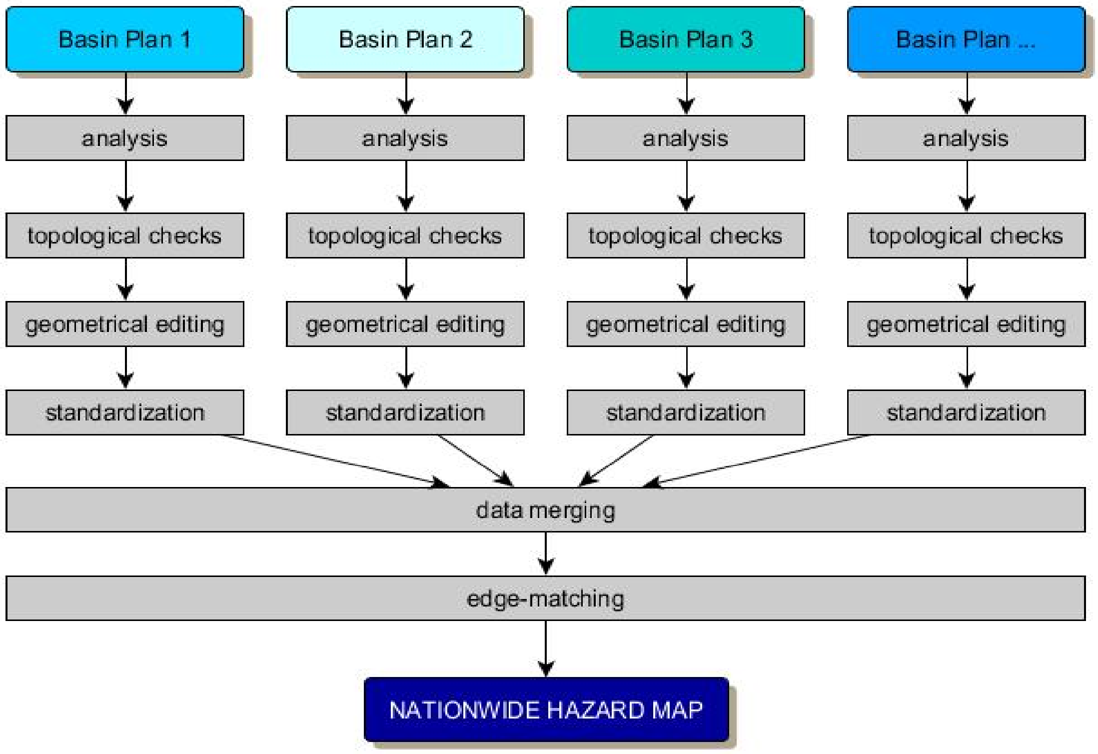


Figure 2: Hazard Areas data processing and standardization.

The outcome is a four class classification of each watershed territory:

- H0: no hazard areas;
- H1: low hazard areas (floods are expected to occur on a less than 200 years frequency)
- H2 medium hazard areas (floods are expected to occur on a 100-200 years frequency)
- H3 high hazard areas (floods are expected to occur on a higher than 50 years frequency)

In order to keep track of the different levels of detail encountered in each Basin Plan an extra table was created relating the general classification attributed (H1, H2, H3) to its original classification, thus allowing to refer anytime to the original information on recurrence and hydraulic heads.

With reference to the INSPIRE directive it is interesting to point out that data specifications related to the topic of floods are provided by the working groups on Hydrography and on Natural Risk Zones: standard rules can so be applied both for base information (like water bodies, water network, catchments) and for hazard areas that identifies parts of territories threatened by floods. Hopefully, as the deadline for the unification of the flood risk maps at national level is approaching, international shared rules will be adopted also by the public administrations involved in data production.

3.2 Building geocoding

Building's position with respect to hazard areas is an essential factor in its flood vulnerability assessment. Defining coordinate pairs for buildings starting from historical datasets is the other main issue to deal with. Normally building location is expressed through addresses instead of coordinates but thanks to common geocoding tools the transition from a textual address to a coordinate pair can be processed automatically. This as long as addresses are up to date and constructed in a standard format.

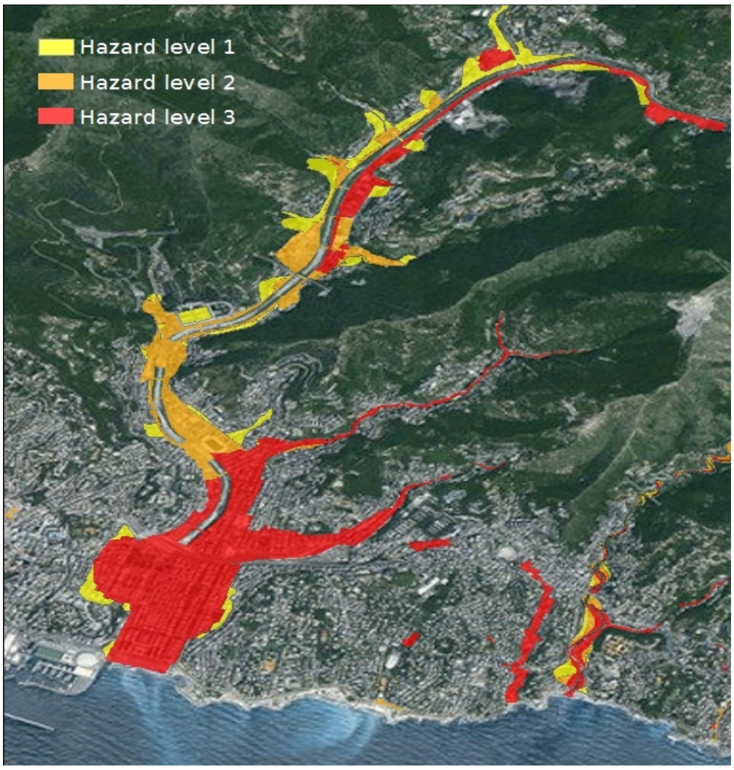


Figure 3: Hazard areas in Genova city.

For testing purposes the insurance company provided the research group with a significant dataset on insurance policies (more than 300,000 records spread all over Italy). Unfortunately the composition of addresses didn't follow a rigorous structure: toponyms were indicated differently (e.g. street vs st.) and some values were missing or out of date. The first operation on this dataset was address normalization, creating a new textual attribute containing address strings as follow:

ADDRESS: "street name, number, postal code, city, country"

After that, strings produced were submitted to three of the most common geocoding engines (Here.com by Nokia, Google and MapQuest) and feedbacks were evaluated in order to identify the more performant tools: preliminary tests were conducted on three regions (Piemonte, Veneto and Calabria) considering a sample of about 35.000 items and once the best geocoder was identified geographical position were generated for the whole portfolio.

As response to the address submission geocoders return a table where coordinate pair for each tuple is identified, together with an indication of the accuracy of geocoding:

- House Number: the geographic position was detected precisely identifying a specific house number in a street;
- Street name: the geographic position refers to the centroid of a street;
- City: the geographic position refers to the centroid of a city;
- POI: the geographic position refers to a point of interest (POI).

Here.com turned out to be the most accurate geocoder with 81% of high quality results; google also obtained good coordinates while Mapquest results were of less accuracy. The following table summarizes the results indicating for each geocoding engine the accuracy of geocoding on a sample of 37.000 addresses.

Address	House number	%	Street name	%	City	%	No response	%
Nokia here.com	29.763	81%	4.767	13%	2.110	6%	172	1%
Google	21.462	68%	3.137	10%	7.112	23%	211	1%
Mapquest	279	1%	15.081	41%	21.062	58%	972	3%

Table 1: Confrontation among geocoders' positioning results.

In addition, some testing on the relative and absolute level of accuracy were carried out: the former type of test was aimed to measure the variation (in meters) between positions located by Nokia and Google considering only tuples with good results; the latter focused on the confrontation between geocoded positions and addresses acquired through GPS survey. Tests results are reported in detail in Guzzetti et al. (2014). This analysis demonstrated how automatic geocoding of buildings is suitable only to have an approximate location of properties and on site survey with a GPS device is required at least for high value policies.

4. Software design

The computation of policy premiums is normally related to statistical analysis conducted on historical data: prices for a certain insurance product are defined based on the refund trend in time. The introduction of geographic information represents a real innovation in the Italian insurance environment and due to the lack of expertise in this field agents need to be supported by technology: technology has to guide operators in all steps of the acquisition process, working automatically on the geographic part and computing all variables involved in the building vulnerability assessment, producing a final feedback on the insurance premium. For this reason software features were designed taking into account the Risk Assessment Methodology presented in chapter 2.
The software tools developed can be divided into three components:

- a webgis framework to view and analyze the insurance portfolio
- an API to calculate parameters for the premium of new insurance policies
- A mobile app to collect all relevant information in the field and communicate with the API to evaluate all parameters.

All these components were developed with Free and Open Source software.

4.1 WebGIS

The WebGIS server is based on the framework FreeGIS.net, which contains following components:

- CentOS operating system
- PostgreSQL database with PostGIS spatial component

- Apache Web Server
- Mapserver, Mapcache e TinyOWS
- php 5.4
- Gisclient author
- FreeGIS.net Web Client

On top of the FreeGIS.net stack, we developed a management software to allow the insert, modification and analysis of insurance policies. This software is composed of a WebGIS interface to view and analyze the whole portfolio of insurance policies, a mobile app to allow collection of all the necessary information for a new policy and a management software to allow entry and management of policies through a webGIS interface.

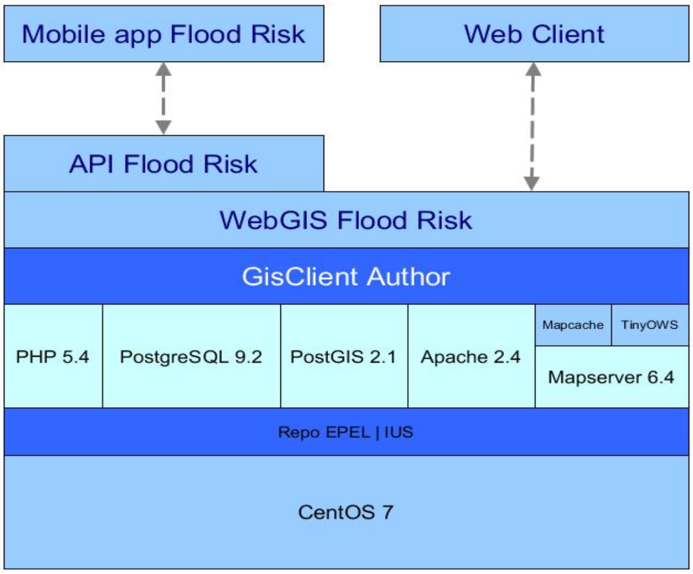


Figure 4: Architectural schema of the WebGIS application.

The WebGIS application has also a web interface which allows registered users to analyze the data and create reports. In particular the WebGIS interface has following analysis functions:

- display of all policies geolocated onto a map
- possibility to select policies based on the Flood Risk areas
- possibility to produce statistics on how many policies fall into each hazard area
- Possibility to overlay areas of damage and policies to understand how many policies are affected by a flood event.

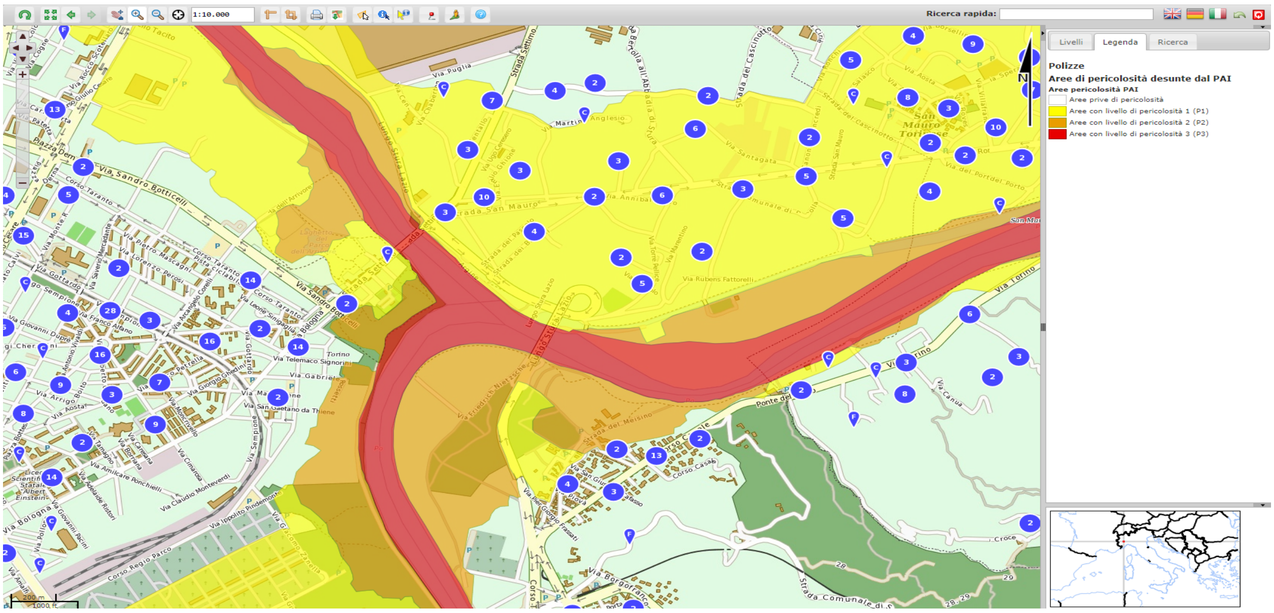


Figure 5: WebGIS interface to analyse position of insured assets compared to flood hazard.

4.2 API

In order to allow the Insurance company to take into account the flood risk areas, altitude, distance from rivers and other information when subscribing a new policy, the existing software tools need to be able to retrieve this information from the GIS Server. This is done through a REST Webservice, where a json request generates a query in the database and responds with the required information.

It is through this API that external databases or the mobile app can communicate to the GIS engine, transmit the position and obtain the information on hazard areas or other geographical layers. In fact also the mobile app communicates with the server through this channel. The communication is based on rest service over https and uses json as data format.

4.3 Mobile App

The mobile app is a prototype which was developed to test the acquisition of a new insurance policy in the field. The mobile app was developed using AngularJS and Bootstrap and compiled for Android with PhoneGap. It allows to enter the data of the policy in the field, obtain the GPS position and enter some information on the site survey. Once this information has been entered, the data is transmitted through the API to the server, overlayed with the hazard area and analyzed to calculate the insurance premium.

The whole process is now being integrated in the IT System of the insurance company.

Figure 6: Mobile interface to record data of field survey and position of insured asset using a smarphone.

4.4 New policies' acquisition process

As stated at the beginning of chapter 4, the technology described has been developed with the scope of providing support to insurers throughout the whole underwriting process of new flood policies, guiding the operators in the collection of all information required for the estimation of policy price. The web interface and the mobile application allow to register in the policies' database (hosted in the webGIS) all data regarding the new policy, both from the insurers' office and from a remote position during an on-site survey. The choice between these two options depends mainly on the value of the new policy to be subscribed: in case of a standard policy the acquisition process may take place from the insurer's office, without any particular inspections; on the contrary, in case of high value requiring a taylor made policy the insurer would be interested in having more detailed data on the property's level of exposure and the mobile application will guide the operator during an on-site survey.

In the case of a standard policy the insurance agent can access the policies' database on the webGIS through an internet browser and start creating a new policy item, submitting all data required (customer name/surname, policy ID, policy value, property address…) to the system: the GIS will acquire the property position through to the geocoder and overlay it with the hazard areas database. The system will calculate the policy price based on the hazard level detected for the property and provide the insurer with final price for the new policy.

In case of a taylor made policy the insurance agent would be required to go on site and collect detailed data regarding building's position and features and submit this information to the system through the mobile application using a smartphone. After preliminary information is registered (customer name/surname, policy ID, policy value, property address...) the position is detected using the GPS device of the smartphone: this operation will guarantee a more accurate property location compared to the geocoded position obtained from the address. As position is acquired, the operator is required to collect information related to the building structure (building period, materials, underground floors...) and to the existing flood proofing devices. Once all this information is gathered it is submitted to the system that will interface it with hazard areas: the hazard level detected, the building's features and flood proofing measures are parameters that affect building exposure and the combination of these parameters will be weighted through a coefficient. The system estimates this coefficient and applies it to the standard policy price, providing the insurance agent with the final policy premium for the new flood policy.

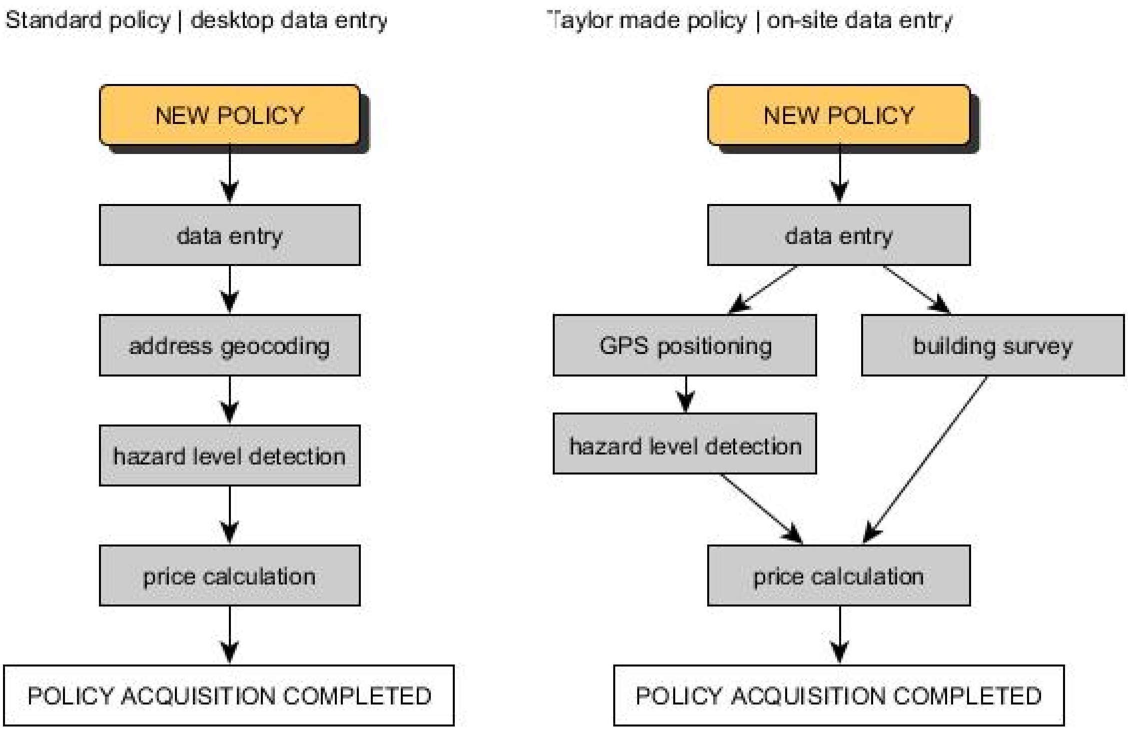


Figure 7: New policies' acquisition flowchart.

5. Conclusions

The result of this project shows that the use of geographic information in a sector which has not made use of it so far can be very productive and bring consistent added value.

The use of open source software in a team outside the insurance company has allowed more flexibility and to reach good results in a reasonable period of time. Now the challenge is to transfer all these tools into the IT structure of the client, which uses to date mostly proprietary technologies.

The results of this pilot project allowed also to unveil the great potential of geographic information in the insurance sector. In fact there are a number of fields where this technology can be applied and used to improve workflows. Currently tests are being made on the earthquake risk, on the estimation of real dam-

ages and on the use of meteorological data. In post-event applications remote sensing may be use to investigate the real impact of a flood on assets: Lidar data surveyed within one day after a flood will provide the company with information about the extent and the heigh reached by the flood as muddy water (differently from clear water) reflect laser pulse, allowing a first general assessment of damages. Also from a damage prevention point of view earth observation may help: the agreement with weather forecasting companies will allow the insurance company to provide meteorological alert services to the insured customers, in order to make them aware of the of the arriving bad weather and allowing them to activate protection devices they have available.
One of the main obstacles encountered in this pilot project is the missing availability of homogeneous open data on flood hazard. A lot of time had to be spent on collecting the Water Basin Plans from the different authorities and transform them in a standard nationwide format. Hopefully this problem will soon be overcome, once the updated and standardized Hazard Maps will be published at national level.

References

- ✔ D'Alesio, F., Hopfgartner, P., & Viskanic, P. (2013). FreeGIS.net una soluzione per la pubblicazione e gestione dei dati territoriale conforme agli standard OGC ed INSPIRE, Atti della 17a Conferenza Nazionale ASITA, Riva del Garda.
- ✔ European Directive 2009/138/CE "Solvency II".
- ✔ European Directive 2007/60/EC "Flood Directive".
- ✔ Guzzetti, F., Pasquinelli, A., Privitera, A., Ronconi, M. (2014). Test metrico sulla ricerca automatica della posizione degli indirizzi, Atti della 18ª Conferenza Nazionale ASITA, Firenze.
- ✔ Legge 18 maggio 1989 n. 183, Norme per il riassetto organizzativo e funzionale della difesa del suolo.
- ✔ Ministero dell'Ambiente, della Tutela del Territorio e del Mare, Indirizzi operativi per l'attuazione della direttiva 2007/60/ce relativa alla valutazione ed alla gestione dei rischi da alluvioni con riferimento alla predisposizione delle mappe della pericolosità e del rischio di alluvioni, 2013.
- ✔ Thompson, S. (2013). Be insured with risk mapping, in *Geospatial World Magazine.*
- ✔ Unesco (1972). *Report of consultative meeting of experts on the statistical study of natural hazards and their consequences*, document SC/WS/500.
- ✔ The new GisClient (www.gisclient.org)
- ✔ FreeGIS.net Geo Information for the public sector (www.freegis.net)

Vector tiles for the Swiss Federal Geoportal

Ingensand Jens, Marion Nappez[1], Cédric Moullet, Loïc Gasser[2] ,Sarah Composto[1]

[1] GIS Lab, University of Applied Sciences Western Switzerland. Route de Cheseaux 1, CH-1401 Yverdon-les-Bains
{jens.ingensand, marion.nappez, sarah.composto}@heig-vd.ch

[2] Swiss Federal Office of Topography. Seftigenstrasse 264, CH-3084 Wabern
{cedric.moullet, loic.gasser}@swisstopo.ch

Abstract

Vector tiling aims at cutting vector data into smaller entities. It has several advantages for the development of web-mapping systems, such as the possibilities to apply different styles, to access attributes or to render 3D data. In this paper we investigate several parameters that need to be considered for the generation of vector tiles. Furthermore we present several options that will be tested in order to establish vector tile services for future versions of the Swiss Federal Geoportal.

Keywords

Vector Tiling, Web GIS, WebMapping, Web Services, Geoportal, Vector Data

1 Introduction

The principle of vector tiling is similar to raster data tiling where large raster data sets are tiled into smaller pieces and stored in hierarchical structures, either in databases or in file systems. Raster tiling allows for an efficient consumption of raster data through a network connection, for instance using the standardized protocol WMTS (OpenGIS Web Map Tile Service, www.opengeospatial.org) implementation standard.

Within the context of web-mapping systems vector data, as compared to raster data, has several advantages – first, vector data allows for more flexibility for rendering maps, such as the possibility to apply different styles or rendering 3D data. Second, it is possible to transfer not only geometries, but also an entities attributes. A third advantage is that, depending on the data layer, the features' attributes, the feature density and the level of detail, data can be compressed at a higher level. A fourth advantage is the fact that data can be reused and transformed on the client side; for instance using coordinate transformation, spatial analysis, and so forth.
However transferring vector data from a server also bears certain risks; e.g. the fact that data can be illegally downloaded and reused for other uses that were not intended by the administrators of a server infrastructure.

The idea of tiled vector data services is to combine both the advantages of vector data with the advantages of tiled raster data services.

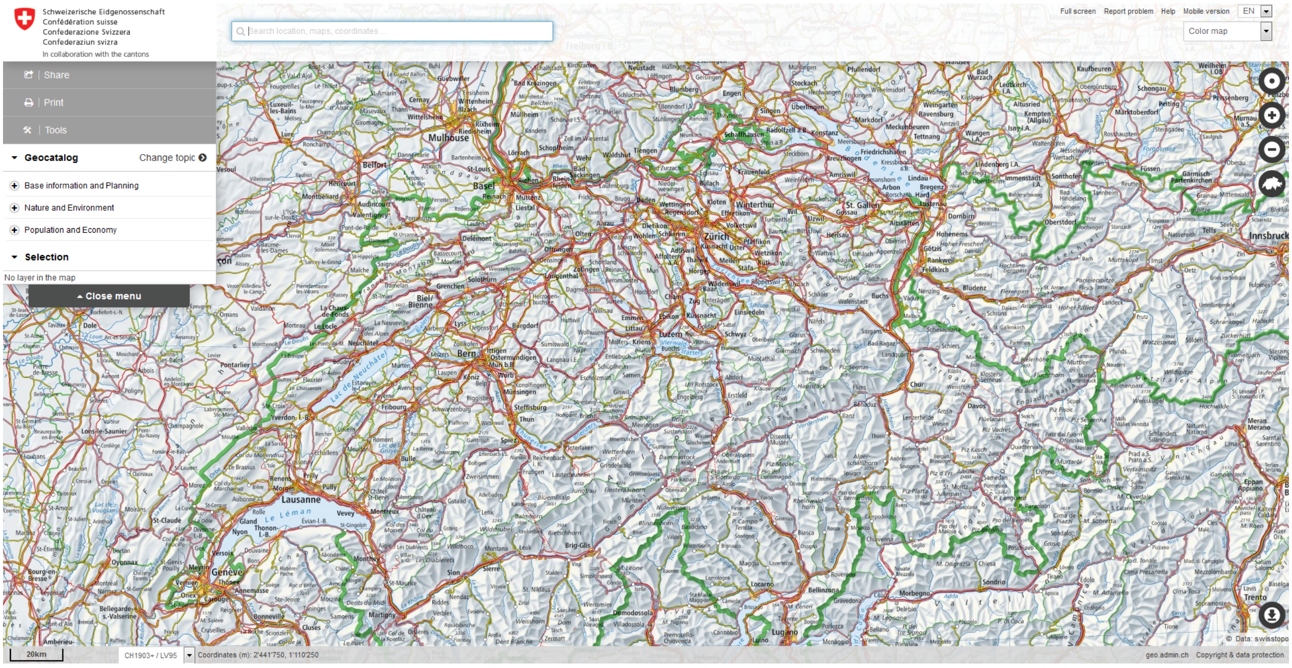


Figure 1: Screenshot of the Swiss Federal Geoportal map.geo.admin.ch

Today the Swiss Federal Geoportal *map.geo.admin.ch* is the official geoportal of the Swiss state, serving almost 400 data layers (see Figure 1) (Moullet 2015). Map.geo.admin.ch is based on the open-source framework MapFish (mapfish.org) which itself consists of several Python and Javascript libraries such as Pylons (www.pylonsproject.org), OpenLayers (www.openlayers.org), ExtJS (www.sencha.com/products/extjs) and GeoExt (geoext.org). The geoportal uses Amazons EC2 and S3 (Amazon 2015) cloud computing infrastructure as a service provider. A majority of the available data layers are generated using WMTS web services; some layers are available as WMS (Web Map Service, www.opengeospatial.org) services. About 2'500'000'000 WMTS unique raster tiles are stored in the cloud infrastructure (Oesch 2014).

One problem in the context of vector tiling is the absence of standards. There are standards for the transfer of vector data such as WFS (Web Feature Service, www.opengeospatial.org) and standards for serving tiled raster data such as WMTS, but today there is no standard for tiled vector data. Possible reasons for this lack of standards might be found in the complexity of vector tiling.

Within the context of the Swiss Federal Geoportal, the goals of the project GeoTile are:

- to investigate existing vector tile service solutions and to analyze which are the different parameters that need to be taken into account for the implementation of tiled vector services and for the development of

clients that need to assemble data layers from cut vector data
- to analyze which of the identified parameters and existing solutions would be usable for the Swiss Federal Geoportal
- to test different parameters in order to identify the best solution for tiled vector services

In the following sections we will illustrate different solutions and concepts for generating, serving and consuming vector tiles. Due to the vast scope of vector tiling, only the most important approaches are cited. Thereafter we will discuss which of these solutions will be tested in the Swiss Federal Geoportal infrastructure in order to identify the solution that is best adapted for future versions of the geoportal.

2 Different approaches to vector tiling

2.1 Regular versus irregular tiling

A first question in the context of vector tiling is whether a vector layer should be cut into pieces of regular size (in terms of equally-sized and shaped bounding-boxes, see Figure 2), whether a vector layer should be cut into pieces with a similar payload or if each object (such a polygon or a point) should be stored in a tile of its own (see Figure 2)

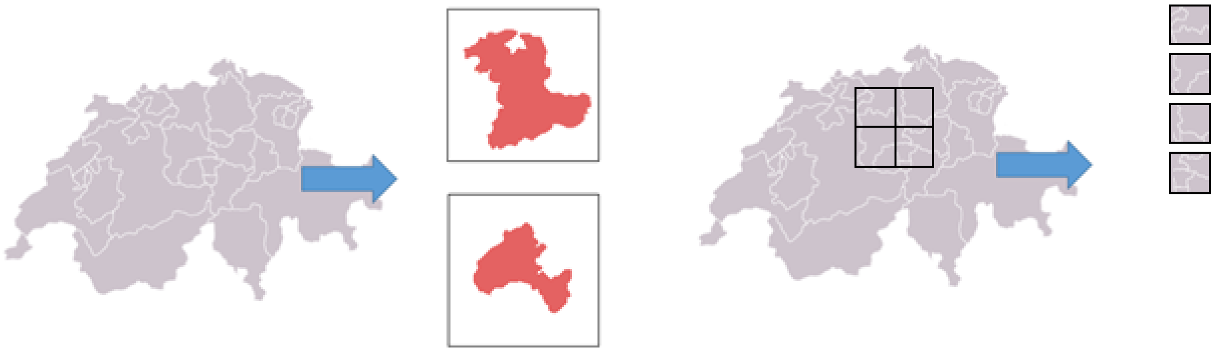


Figure 2: Examples for irregular and regular vector tiling.
Irregular tiling: each object is a tile of its own (left)
Regular tiling: a square bounding-box is applied to cut a layer into equally-sized and shaped tiles (right)

Today several applications such as Google Maps (maps.google.com, Google Inc, Mountain View (CA), USA), GISCloud (www.giscloud.com, London, UK) or Polymaps (polymaps.org) have adopted the concept of regular vector tiling.

One notable example for irregular tiles has been described by Dufilie and Grinstein (2014). In their approach shapes are cut using the TileSplit algorithm into pieces of equal payload. The results are both an overview layer that indicates the positions of the tiles and the tiles themselves. In order to generate tiles for different levels of detail, the authors used Visvalingam's algorithm (Visvlingam and Whyatt 1992).

2.2 Formats for storing and transferring vector tiles

Similar to raster tiles, vector tiles can be stored in database structures or file systems. Two common formats for transferring vector data in a web context are XML-based formats such as GML and JSON-based formats such as GeoJSON and TopoJSON. Both types of formats have advantages and disadvantages. XML-based formats are extensible and offer several possibilities to verify data. The disadvantage of XML-based formats is the fact that much storage is used for tags. JSON files on the other hand need less space and JSON files are easily parsed using Javascript (JSON 2015)

2.3 Generalization and simplification of vector data

Depending on the zoom level, data needs to be simplified and generalized in order to maintain readability and payload. Several methods and algorithms could be used within the context of tiled vector data.

One possibility to generalize vector data is the GAP (Generalized Area Partitioning) tree method which focuses on topological area data. This method identifies the importance of areas. Depending on the level of detail less important areas are removed (Van Oosterom 1993). Tinghua and Van Oosterom 2002 have extended this method in order to include bridge areas between aggregated polygons.

Another algorithm described by Visvlingam and Whyatt (2002) focuses on the importance of a line's edges. In this method an edges neighbors are used to calculate the area of a triangle. Depending of the calculated area, an edge is considered more or less important. If the line needs to be simplified, less important edges are removed. This method can be used both for linear features as well as for area features.

The Douglas-Peucker (Douglas and Peucker1973) algorithm is perhaps the most widely used algorithm for generalization in GIS software. The algorithm uses distances between the points of a line to determine which points can be omitted.

For point features, there are basically two possibilities to reduce the number of points: Either points can be grouped into clusters, for instance using the K-Means clustering algorithm (MacQueen 1967), or thinned such as described by for instance Moenning and Dodgson (2003). In both cases the loss of information (e.g. in terms of attributes) is an important subject to consider.

2.4 Attribute handling

If geometric entities are cut into several pieces, one important question is how attributes are handled. One solution to this problem is to replicate the attributes to all pieces of the cut vector feature and thus to increase the overall payload. Another possibility is to attach the attributes to one of the pieces and to enable clients to reassemble geometries and attributes. A third solution is to transfer attribute data separately and independently from the geometries.

2.5 Client-side object assembly

A challenge in the context of vector tiling is the assembly of cut entities into the original vector features. The main question is how a client can know that all pieces of one entity have been transferred. The simplest method is to check all tiles that have been received for e.g. the existence of pieces of one specific entity. The drawback of this method is the time that is necessary to analyze all tiles. Another method has been described by Nordan (2012) – in order to locate pieces of tiled vector features, pointers can be included in the pieces, indicating in which directions a client needs to look for more pieces.

3 Identification of the best approach

In order to find the method that is best adapted for the Swiss Federal Geoportal, a testbed has been deployed. This testbed includes a copy of several data layers stored in PostgreSQL/PostGIS databases. Python scripts will be used to connect to these data sources and to produce vector tiles. The following options have been chosen:

- Only tiles with equally shaped bounding-boxes will be generated (regular tiling). The reasons for this decision are that the effort to generate irregularly shaped tiles appears to be greater than for generating regularly shaped tiles; second OpenLayers, the Javascript library that is used in the geoportal already has a built in support for regularly shaped raster-tiles (WMTS). Thereby an extension of this library (in order to enable the consumption of vector tiles) would be easier to accomplish. Furthermore an idea is to use the WMTS tile-indexation for the generation of the tiles (e.g. a vector tile would be addressed using a similar URL as a raster tile) to simplify the assembly of both raster- and vector tiles in a client.
- Several algorithms for the generalization of vector tiles will be tested.
- All tiles at all levels of detail will be generated at once. Thereafter we will analyze the payload, the stored features, etc. with statistical methods.
- Only JSON-based formats will be considered. The main reason is that JSON-based formats are lighter than XML-based formats in terms of payload and data processing.
- Attributes will be handled in separate files, independently from the vector tiles. This has the advantage that no redundant data will be generated.

The main objective of our approach is to find a solution that balances flexibility, performance and complexity. Important parameters that will be considered are the payload of the generated tiles as well as the time it takes to generate the tiles for a given layer. Moreover the generated tiles will be used as a repository for the development of the Javascript client.

4 Discussion and conclusions

Vector tiling paired with webGL-enabled Javascript libraries permits to create high-performance web-mapping applications. Today mostly proprietary solutions such as Google Maps are using vector tiles. We can state that there is a lack of standards for the generation and consumption of vector tiles. Vector tiling however is a complex subject as there are many different possibilities to implement such services depending on the context of the application. In Google Maps for instance there is a limited number of data layers available, however these data layers cover the surface of the entire planet. The Swiss Geoportal on the other hand offers hundreds of data layers – from point feature layers containing all addresses in the whole of Switzerland to polygon layers with all of Switzerland's buildings.

Our approach to tiled vector data refers to existing standards such as the WMTS tile indexing and JSON-based file formats. Moreover we will use a testbed that enables empirical testing of different alternatives.

Since this project is an ongoing project, at the time of writing no empirical results are available.

References

Amazon (2015): AWS Case Study: Swisstopo. Amazon web services portal: http://aws.amazon.com/fr/solutions/case-studies/swisstopo/ (9.3.2015)

Douglas, D., Peucker T. (1973): Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. In: The Canadian Cartographer. (10) 2. 1973, pages 112–122.

Dufilie and Grinstein (2014): Feathered Tiles with Uniform Payload Size for Progressive Transmission of Vector Data. In: Pfoser and Li (eds): Web and Wireless Geographical Information Systems. 13th International Symposium, W2GIS 2014, Seoul, South Korea, May 29-30, 2014. Proceedings. Pages 19-35 Springer

JSON (2015): JSON: The Fat-Free Alternative to XML. http://www.json.org/xml.html (9.3.2015)

MacQueen, J.B. (1967): Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297

Moenning, C. and Dodgson, N.A (2003) A new point cloud simplification algorithm. Proceedings of the 3d IASTED International Conference on Visualization, Imaging and Image Processing (VIIP 2003), 8-10 September 2003, Benalmádena, Spain

Moullet, C. (2015): Géoportail de la confédération map.geo.admin.ch: contenu et évolution technologique. Journée de veille technologique IGSO, 28.1.2015, Yverdon-les-Bains, Suisse

Nordan, R.P. (2012) An Investigation of Potential Methods for Topology Preservation in Interactive Vector Tile Map Applications. Trondheim University, Norway.

Oesch, D. (2014): The Swiss Federal Geoportal www.geo.admin.ch - more than maps. Geospatial World Forum 2014 8.5.2014 Geneva.

Tinghua, A. and Van Oosterom, P (2002): GAP-Tree Extensions Based on Skeletons. In: Proceedings of the 10th International Symposium on Spatial Data Handling. July 9-12 2002, Ottawa, Canada

Van Oosterom, P (1993). The GAP-tree, an approach to “On-the-Fly” Map generalization of an Area Partitioning. In: Proceedings of the GISDATA Specialist Meeting of Generalization. Compiègne, France, 15-19 December 1993

Visvalingam, M. and Whyatt, J.D. (1992): Line Generalisation by Repeated Elimination of the Smalles Area. Cartographic Journal., 30 (1) pages 46 - 51

A GIS tool for reduction day precipitation to subday

Martin Landa[1], Petr Kavka[2]

[1] Department of Geomatics, Faculty of Civil Engineering,
Czech Technical University in Prague, Czech Republic

[2] Department of Irrigation, Drainage and Landscape Engineering, Faculty of Civil Engineering,
Czech Technical University in Prague, Czech Republic

Abstract

Subday design precipitation series are important for hydrological modelling of soil erosion problems in a small catchment scale when designing common measures for promoting water retention, landscape drainage systems, etc. The project with the name „Variability of Short-term Precipitation and Runoff in Small Czech Drainage Basins and its Influence on Water Resources Management“ is based on the precipitation scenarios analysis from observed data of point gauging stations and radar data in terms of events' return period, rainfall total amount, internal intensity distribution and spatial distribution over the area of the Czech Republic. The project has been launched in April 2014.

In first phase has been implemented automatization of well-known method which is based on reduction of 24 hours design precipitation to shorter time. GIS is used for spatial supervised classification of point values of specified repetition periods (2, 10, 20, 50 a 100 years) in the area of the Czech Republic. This work presents a tool for reduction day precipitation which has been implemented in GRASS GIS. The goal is to provide a tool to the public as a stand-alone program and later also as a web geoprocessing service. The tool uses methods of zonal statistics to compute average values of design 24 hours precipitation for a selected area or for a spot. This value is reduced to the chosen length design rain for selected period of repetition.

The result provided by this tool will be used as a reference value for subsequent analysis in this project.

This work has been supported by the research project QJ1520265 - „Variability of Short-term Precipitation and Runoff in Small Czech Drainage Basins and its Influence on Water Resources Management“.

References

- Torfs, P. J. J. F., Horacek, S., & Rakovec, O. (2010). Chimera, a framework for conceptual modelling using analytical sensitives to parameters: implementation and application to PDM model.
- Sokol, Z., & Bližňák, V. (2009). Areal distribution and precipitation–altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year. *Atmospheric Research*, *94*(4), 652-662.

FOSS4G
EUROPE
COMO 2015