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.

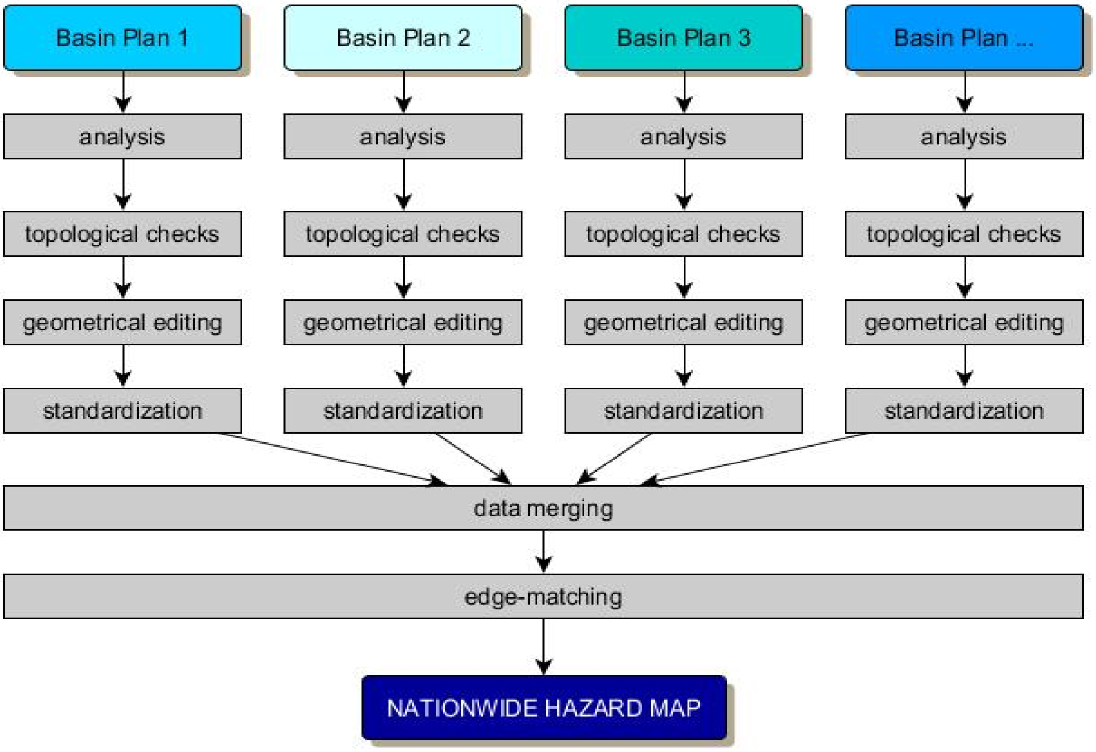

Figure 2: Hazard Areas data processing and standardization.

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.

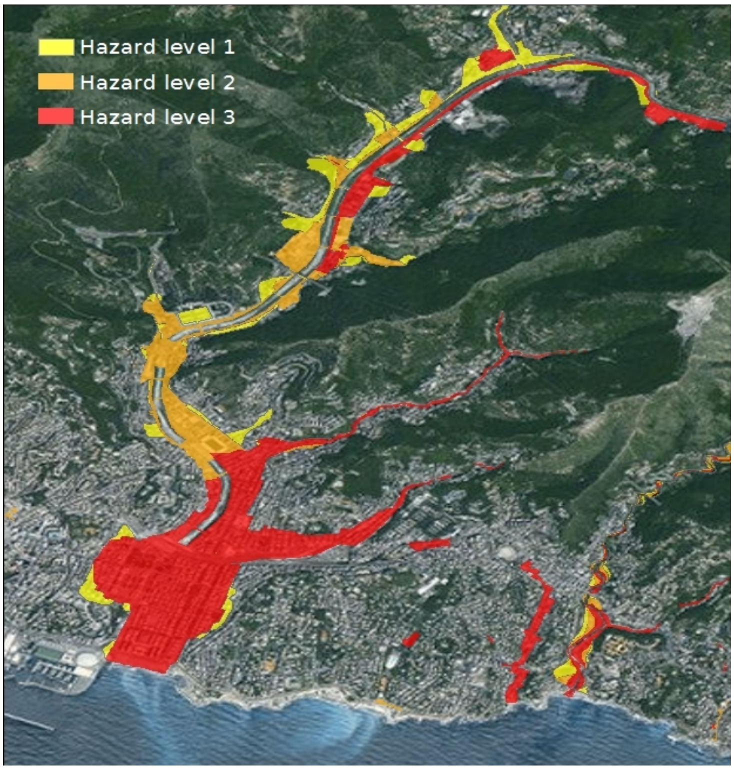

Figure 3: Hazard areas in Genova city.

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.

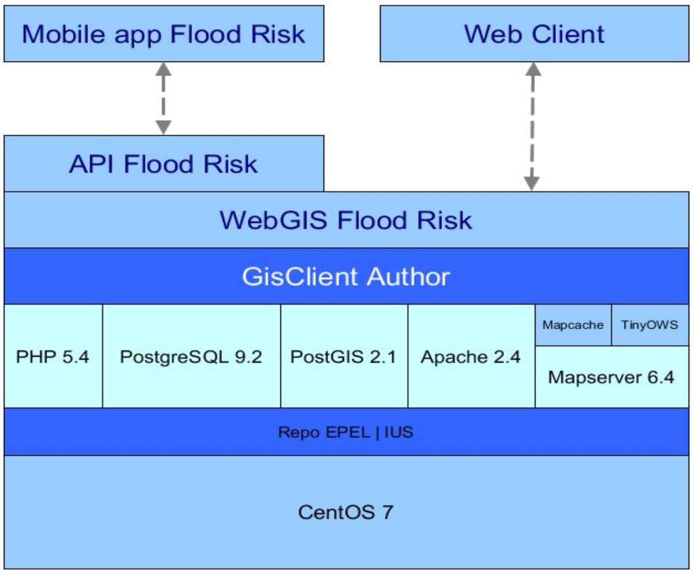

Figure 4: Architectural schema of the WebGIS application.

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.

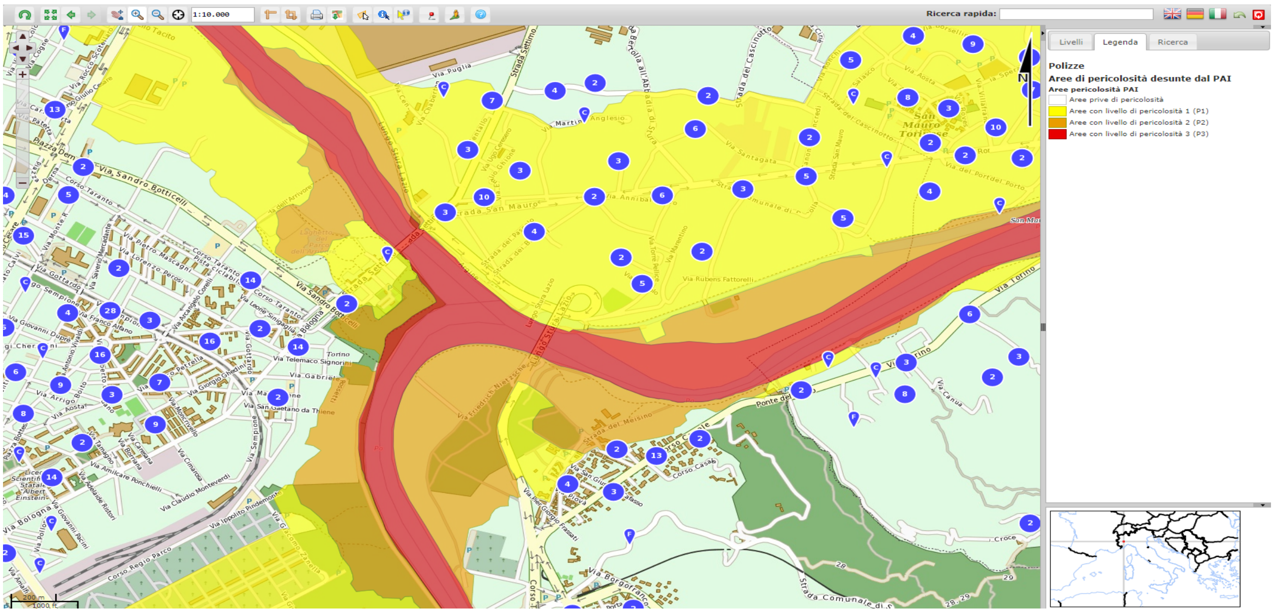

Figure 5: WebGIS interface to analyse position of insured assets compared to flood hazard.

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.

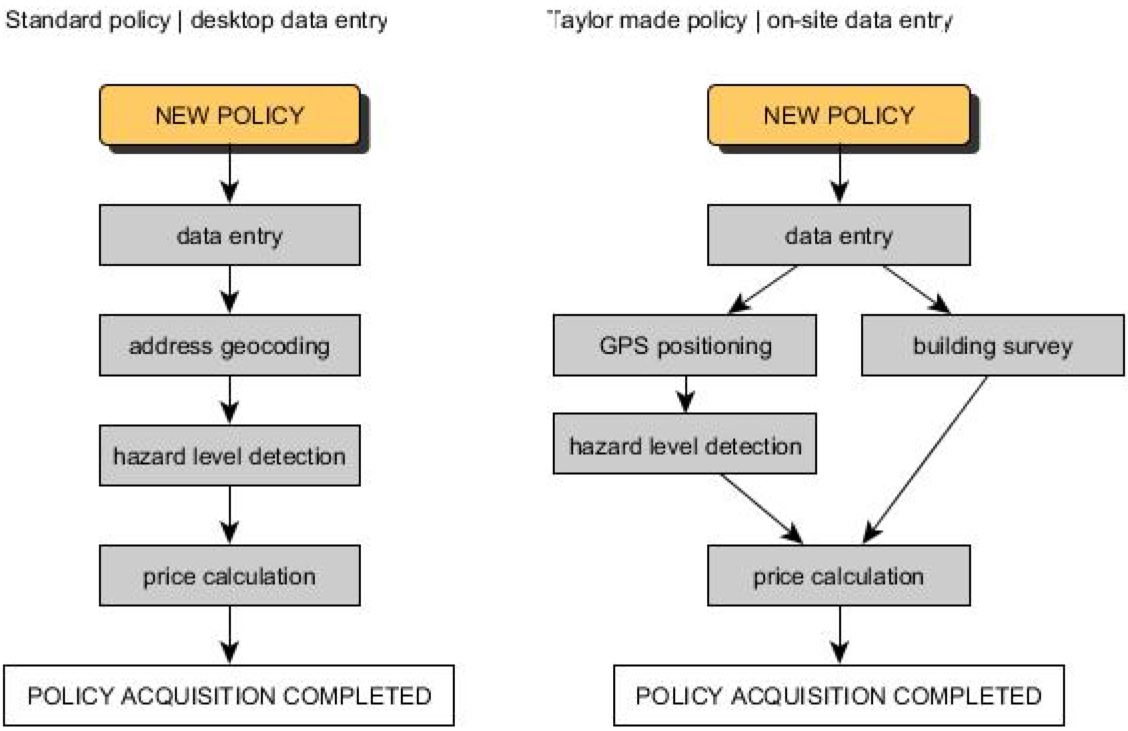

Figure 7: New policies' acquisition flowchart.

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)