SITOGEO: A geographic database used for GIS applications

M. Bisson, A. Fornaciai and F. Mazzarini

Istituto Nazionale di Geofisica e Vulcanologia - Via della Faggiola 32, 56126 Pisa, Italy

Summary. — This contribution aims to present the geographic database "SITO-GEO" developed with GIS technology. This database manages data of different nature, source and resolution (land images, digital elevation model, cartographic maps and vector data) covering the whole of Italy. In this paper we show that our database can be used for the assessment, management and mitigation of risks regarding the preservation of cultural heritage and as valid support for the remote-sensing technologies.

PACS 91.10.Da – Cartography.
PACS 93.85.Bc – Computational methods and data processing, data acquisition and storage.
PACS 93.30.Ge – Europe.

1. – Introduction

Geographic Information Systems (GIS) allow to store, visualize, manipulate and analyze huge and various volumes of different data types (spatial, non-spatial, simplex, complex, quantitative and qualitative) providing a variety of modelling techniques. For these characteristics the GIS applications range from simple geographic databases where different informative layers can be visualized in overlain and queried, to complex spatial decision-making systems used for analyses, management and mitigation of natural disasters [1].

This paper describes in detail the framework and the data of the Geographic Information System "SITOGEO" related to the Italian territory and shows two specific and different applications of such GIS. The first application regards the assessment, management and mitigation of natural risks for cultural heritage. In this case SITOGEO is used as a basis for the creation of a census of cultural heritage at a municipal level and successively as a tool for the assessment of hazard and risk maps in case of natural disaster (such as flooding events). The second application shows as SITOGEO is a useful tool to correct geometrically satellite images. It can be a valid support for analyses of remote-sensing data sets.
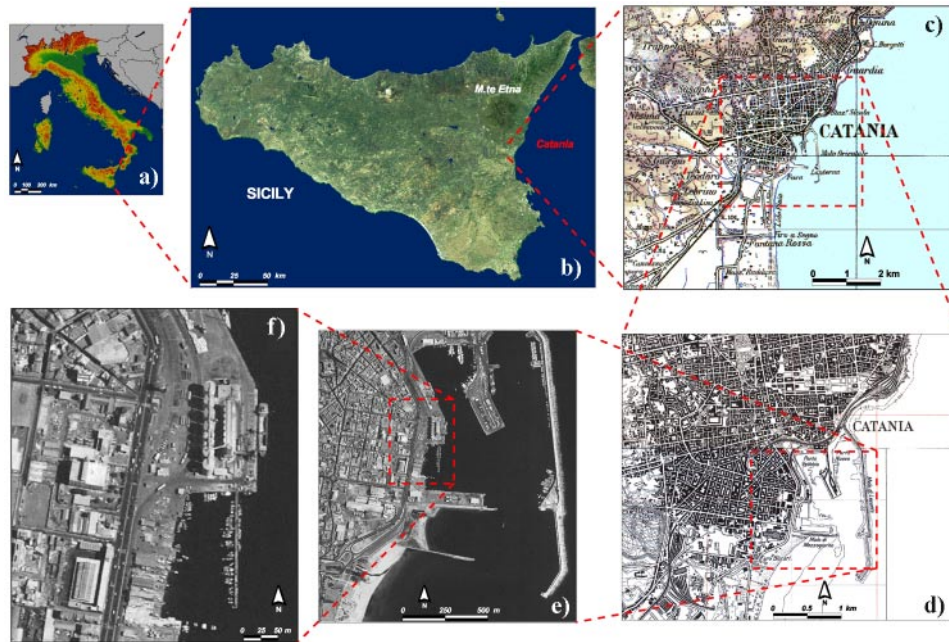
Fig. 1. – The image data type in SITOGEO: a) DEM of Italy with step of 240 m, b) LANDSAT 5 TM image of Sicily, c) IGM Cartographic Map at a 1:100000 scale: Catania Area, d) IGM Cartographic Map at a 1:40000 scale: City of Catania, e): Aima Orthophoto: Catania seaport at a 1:20000 scale, f): Zoom of Aima orthophoto: wharf with a wet dock in the Catania seaport at a 1:5000 scale.

2. – Data available in SITOGEO

SITOGEO (Sistema Informativo Territoriale per la Gestione delle Emergenze e Organizzazione della pianificazione) is a GIS application created in collaboration with the Dipartimento di Protezione Civile of Italy [2] in order to provide a useful tool for assessing, managing and mitigating natural risks (*e.g.*, earthquakes, volcanic events and hydrogeological phenomena) and was realized with the GIS software ArcView 3.3 [3]. The data available in SITOGEO are geographic and cover the entire territory of Italy ($301680 \, km^2$). They come from various sources and are characterized by different resolutions. In order to preserve the metric information and ensure territorial continuity, all data are organized and georeferenced in three geodetic cartographic systems (Datum, Projection, Projection zone): Rome 40, GB, zones west and east; ED50, UTM, zones 32 and 33; WGS84, UTM, zones 32 and 33. The data base of SITOGEO consists of raster, vector and alphanumeric data.

2˙1. *Raster data*. – *Raster data* are constituted by images of different nature and source (fig. 1). They are distinguished in four different typologies: digital elevation model, satellite images, raster IGM, airborne images.

1) *Digital elevation model.* In SITOGEO two images of digital elevation model of Italy were available: the shaded relief and the coloured altimetric image (fig. 1a). The first is an image with 256 gray levels obtained shading the elevation matrix with GIS

techniques. This type of elaboration allows to visualize the topography of territory and to derive several information useful for scientific studies in morphologic and tectonic fields. In fact it possible to recognize faults, landslides, analyze the drainage network, and many others [4, 5]. The second image is obtained combining the first with the information of elevation coloured for altimetric range. Both image types are memorized in TIFF format. The elevation matrix used has a resolution of 240 m and was produced by Carozzo *et al.* [6] as a mean-height archive derived from 1:25000 scale contour lines of IGM.

2) *Six satellite image mosaics.* Each image is a mosaic of Landsat 5 TM images; it was stored as a RGB natural colour composite in TIFF format and with a geometric resolution of 30 m. These Landsat images provide approximate land use and land cover information for all of Italy (fig. 1b).

3) *Raster IGM* (Istituto Geografico Militare) are digital cartographic maps at nominal scales of 1:25000, 1:50000, 1:100000 and 1:250000. The small scale 1:100000 and 1:250000 maps are recorded in TIFF format (CCITT Group 4, Packbits) and have a pixel resolution of 10 and 25 meters, respectively (figs. 1c and d). The medium scale 1:25000 and 1:50000 maps are in BMP format and show a pixel resolution of 2.5 and 5 m, respectively. In these raster maps the various informative levels (*e.g.*, built-up areas, orography, hydrography, roads) can be visualized individually since they are recorded in different bits. The images at scale of 1:50000 and 1:250000 have a colormap describing 16 informative levels, the images at scale of 1:25000 and 1:100000 have a colormap showing 256 informative levels. On the whole the IGM database is composed by about 9000 images for a total of 300 Gigabyte. For each group of maps it is possible to visualize the graphic file (reference frame) linked directly with the respective images.

4) *Airborne imagery.* AIMA (Azienda di Stato per gli Interventi nel Mercato Agricolo) images are orthophotos characterized by 256 grey levels and a ground resolution of 1 m. This resolution allows the identification of detailed features such as roofs, bridges, road intersections, sidewalks, trees, etc. (fig. 1f). The orthophotos database consists of about 30000 images saved in ERMapper's compressed wavelet (ECW) format for a total of 150 Gigabyte. The spatial coverage of every single image is 6.0×7.4 km but considering the zones in overlapping each orthophoto provides univocal information only for an area of 4.5×4.5 km. The acquisition of original photograms was realized by the AIMA and took place in 1997/98 at altitudes below 2000 m; the photograms were taken at a scale of 1:40000 and ensure a planimetric tolerance less than 4 m.

2˙2. *Vector and alphanumerical data.* – The *vector data* consist of 14 different graphic layers or geographic themes (see table I) describing information on administrative boundaries, road and rail networks, hydrography and toponymy. Each layer is memorized in a data format called shapefile storing the geometric and attribute information of spatial features. A shapefile consists of a main file, index file and a database file. The main file (.shp) describes the geometry and the spatial coordinates of features, the database file (.dbf) is a table cointaining the attributes of each feature memorized in the corresponding record and the index file (.shx) links each attributes record to the associated shape in a relationship one to one. So each graphic layer can be visualized and queried. An example of layer is the municipality boundaries shapefile consisted of 512 features connected to the ISTAT 95 alphanumerical database storing several geographic information and population statistics: for each municipality it is possible to know the name, the extension (km^2), the length of boundary (km), the ISTAT code, the administrative divisions (province and region belong to), the total number of inhabitants, the inhabitants density (number/km^2), the number of inhabitants for range of age, the number of women, the

TABLE I. – *The vector data in SITOGEO.*

Layer	Description	Graphic feature
Regioni.shp	Regional administrative boundaries	Polygon
Province.shp	Provincial administrative boundaries	Polygon
Comuni.shp	Municipality administrative boundaries	Polygon
Prefissi.shp	Teleselective codes	Polygon
Cap.shp	Post Code	Polygon
Z.franche.shp	Tax-free zones	Polygon
Capoluoghi.shp	Regional capital town	Point
Città.shp	Town	Point
Città_poly_.shp	Town (area)	Polygon
Ferrovie.shp	Railway lines	Line
Strade.shp	Roads	Line
Fiumi.shp	Rivers	Point
Laghi.shp	Lakes	Polygon
Toponimi.shp	Toponymy (about 800000 toponyms)	Point

number of men, the number of family and many others information.

All vector themes are at a nominal scale of 1:100000 except the layer of toponymy. It consists of 800000 georeferenced place names derived from IGM cartography at a 1:25000 scale and memorized in a point graphic format; this layer allows to identify various features such as cities, rivers, mountains, hills, small towns and many others.

3. – The tools of SITOGEO

The SITOGEO application was realized implementing the ArcView 3.3 Graphic User Interface. The basic functions like displaying views with multiple map layers, moving and zooming throughout a map, drawing graphic elements and descriptive texts, identifying and selecting features with queries or with spatial relationship, are enhanced with additional software facilities such as new menus, tools and buttons. These new extensions were written in Avenue, ArcView's built in programming language, and allow two different operative modes: automatic and interactive.

In automatic mode the application loads and visualizes data according to the display scale (table II) and geographic area. Using only the "*pan*", "*zoom out*" and "*zoom in*" tools we can quickly view numerous information on any part of the territory either at a glance or in great detail. In this way it is possible to fly ideally over the whole national

TABLE II. – *The Image data in SITOGEO.*

Data	Display scale
Landsat 5 TM	1:200000–1:1000000
Raster IGM 1:250000	1:100000–1:200000
Raster IGM 1:100000	1:40000–1:100000
Raster IGM 1:50000	1:20000–1:40000
Raster IGM 1:25000	1:10000–1:20000
AIMA Image	< 1:10000

territory at different scales, from an overall view of the country to close up views where details are clearly visible (*e.g.*, buildings, squares, monuments, roads, sidewalks).

In interactive mode, the application activates new buttons and new pop menus (coloured in red or blue); the most frequently used are "*Fm_view*", "*Index image*", "*Go xy*", *Go lat long*" and "*Find toponym*". *Fm_view* is a pop up menu for choosing the type of data to visualize irrespective of the display scale. The *index image* tool loads or unloads raster data for a selected area, allowing the overlay of other images with informative levels in transparency. The "*go xy*" and "*go lat long*" buttons localize a geographic point according to user specified coordinates (respectively projected and nonprojected); this point is represented by a clearly marked red circle on the viewed data. Lastly, the "*Find toponym*" button is very useful for rapid positioning on the territory. In the querying mode (for the entire database, a category, or a geographic area), when one types the name of a toponym the system shows a list of names satisfying the request. After selecting the desired target, the application localizes a yellow point (which label identifies the selected name) and visualizes a suitable image as background. In this way it is possible to quickly pan the territory and rapidly identify natural and manmade features, whose relative coordinates are automatically displayed. An example of "*Find toponym*", where the searched element is the mouth of the Arno River (Tuscany), is described in fig. 2.

4. – Application of SITOGEO to the cultural heritage: The centre of Florence

SITOGEO can be a valid support to create a census of cultural-heritage assets of our country. The realization of such census is the first step to assess risk and hazard map for cultural heritage in case of natural disastrous (*i.e.* earthquakes, volcanic eruptions). The particular application proposed here regards the cultural assets of the city of Florence and the areas that could be exposed to disastrous flooding events like the overflow of Arno river in the November 1966 [7].

Working in interactive mode and with a suitable scale of visualization (1:3000-1:4000), SITOGEO allowed to recognize easily different details (historical buildings, monuments, squares, bridges, etc.) moving quickly throughout high-resolution images (AIMA ortophotos). Each structure was identified by a point graphic element and several related information was memorized in the associated table. Querying this table it is possible to know the geographical location of each monument (x and y are expressed in metric coordinates ED50–UTM–Zone 32), the respective name and the typology it belongs to. All cultural assets identified in the centre of Florence are stored in a graphic layer. Figure 3, coming from a SITOGEO view, shows this layer overlaid on the AIMA ortophotos. The layer identifies and localizes 93 cultural assets in Florence centre; the different symbols correspond to different typologies of monuments (*e.g.*, church, chapel, cathedral, museum, temple, baptistery).

In order to assess a hazard map for cultural heritage, the following step has been to relate this census to the areas prone to possible overflow of Arno river. Using the Aima images as main cartographic reference base, fig. 4 shows the polygonal graphic theme defining the areas affected by overflows of Arno river [13]. According to the level reached by water (h) we can recognize three zones characterized respectively by high, intermediate and low hazard: zone A: ($h_{\text{water}} > 4\,\text{m}$), zone B ($1\,\text{m} < h_{\text{water}} < 4\,\text{m}$) and zone C ($h_{\text{water}} < 1\,\text{m}$). The zoom proposed in fig. 5 visualizes an area near the centre of Florence ($\sim 1\,\text{km}^2$ wide) where the zones A, B, C and several monuments as Battistero, Duomo, Galleria degli Uffizi and Palazzo Medici are well visible. More exactly we can distinguish 36 monuments: 9 localized in area A, 24 in area B and 3 in area C. Displaying
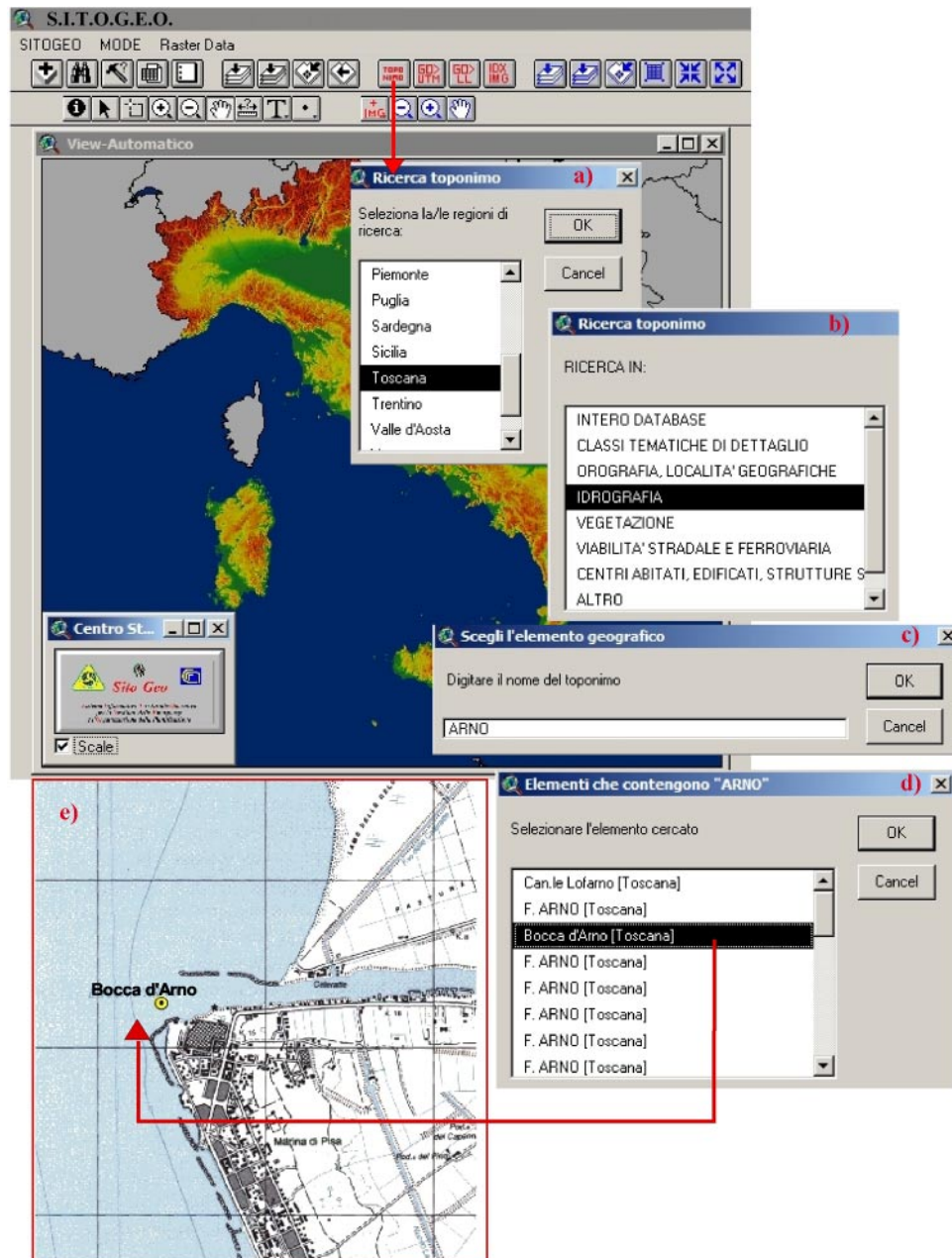
Fig. 2. – Locating the mouth of the Arno River through the "*Find toponym*" option. The illustrated steps are the following: selection of the study area (a), category selection, *e.g.*, hydrography (b), typing the name of the river (c) and final selection from the list (d), display of the toponym on a suitable image (e).
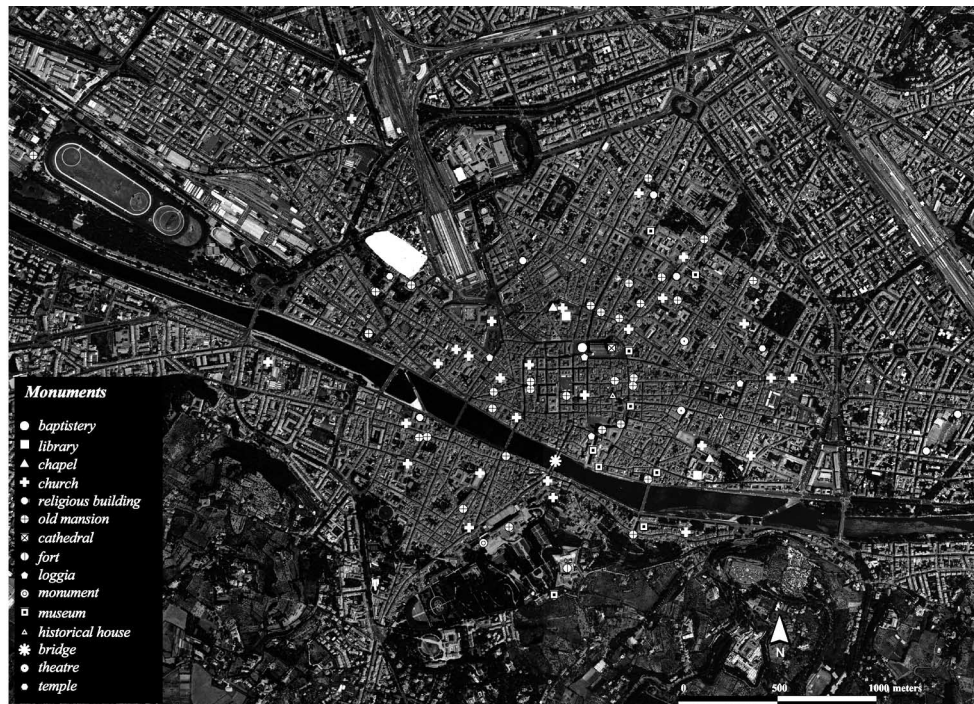
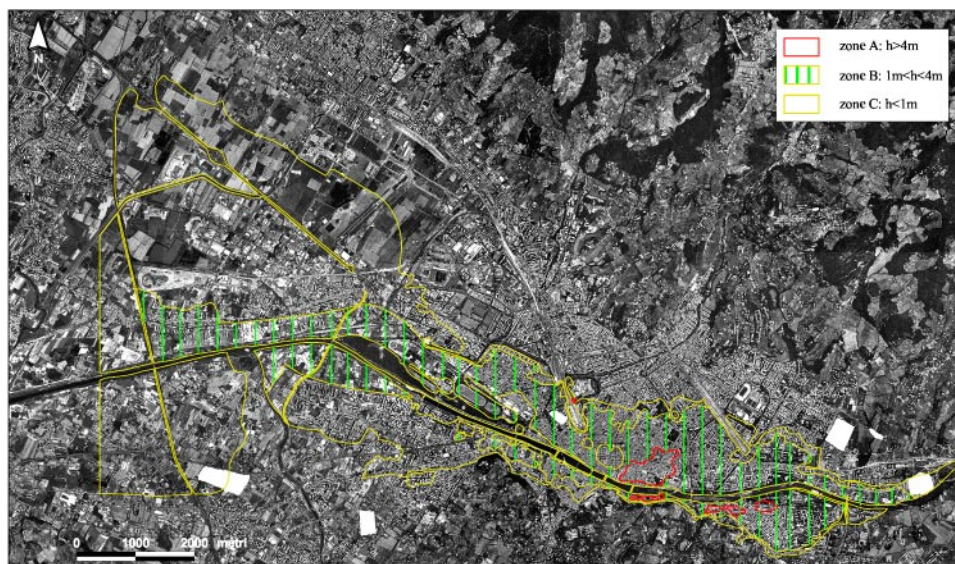Fig. 3. – The census of cultural assets in Florence Centre.

Fig. 4. – The areas affected by a possible overflow of Arno river (h indicates the level reached by water).

Fig. 5. – The monuments and the risk zones.

a multiple layer map, very simple but fundamental operation in a GIS system, allowed not only to localize the structures respect to areas interested by a possible overflow but, most of all, to individualize rapidly the assets that could be subjected to mayor and minor risk. In our specific case, the cultural sites subjected to mayor hazard and therefore interested by a more prompt intervention of protection are localized in zone A (*i.e.* Biblioteca Nazionale, Teatro Verdi, Casa del Buonarroti and others). The case study on Florence is a very simple example but this type of application, extended to the entire territory of Italy allows to drawing up at national scale hazard maps for the cultural heritage in case of natural calamity.

5. – Application of SITOGEO for the geometric correction of Landsat 7 ETM+ imagery

The integration of remote sensing and GIS is widespread in a broad range of environmental studies [8, 9]. Owing to the geometric distortion caused by the sensor and platform orientation and terrain elevation, it is possible to obtain a planimetrically correct image for mapping land features. An image that has been orthorectified has an
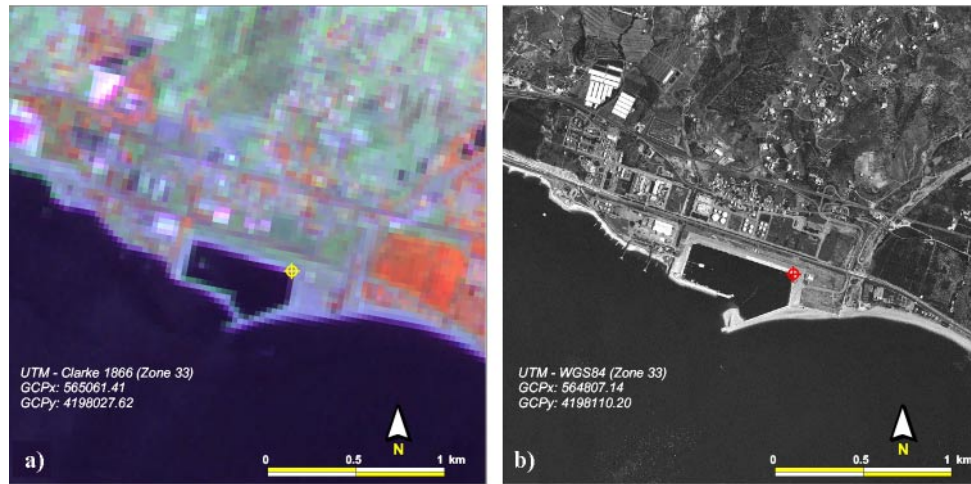
Fig. 6. – Porto of Saline—Montebello Ionico (RC). GCP source in LANDSAT 7 image (a) and GCP reference in AIMA orthophoto (b).

accurate geometry, and can be used as a base map for new data sets or for updating existing databases. In this application we show that SITOGEO can assist the geometric correction of a series of satellite images using the AIMA orthoimages of Italy with a high geometric resolution, a digital elevation model with a cell size of 30 m (DIGITALIA Project [10, 11]) and a Landsat Geometric Model of ERDAS Image 8.6 software [12].

Landsat 7 ETM+ scenes, covering the entire territory of Italy, were acquired for the period February 15, 2001 to November 10, 2002 in 8-bit integer CEOS format referred to the UTM Clarke 1866 System.

Geometric correction of satellite images involves modelling the relationship between the image and ground coordinate systems. Orthorectification of the image requires several steps: identification of ground control within the image, collection of Ground Control Points (GCPs), development of an orthorectification model, orthorectification of the image using the model and accuracy checks.

Source Ground Control Points were determined using ERDAS software, while reference Ground Control Points were identified through SITOGEO, using Aima orthophotos. Their high resolution allows the identification of detailed features. About ten control points with as uniform distribution as possible over the entire scene were selected (fig. 6). The RMS error for GCPs derived from multispectral data and panchromatic band ranges between 0.0 and 3.5 pixels with mean value of 1.0.

Misregistration between the orthorectified Landsat image and the digital orthorectified airborne image was quantified calculating the total RMS error for each Landsat image and the RMS for each point. The accuracy of the procedure was also checked visually. The selected features usually have high contrast and a constant spectral definition over time (*e.g.*, road Intersection; bridges; artificial coastlines and shorelines).

The reference GCPs were collected through SITOGEO using the interactive tools to load and automatically visualize the suitable AIMA images at the appropriate scale. The points were acquired in georeferenced digital format (vector file) and linked to an information table (*e.g.*, *x*-coordinate, *y*-coordinate, Aima name, Description of GCP, etc.) that can be updated automatically or by the user. The Landsat image GCPs

were subsequently linked to the corresponding AIMA GCPs by setting the reference coordinates.

Orthorectification was accepted if the total RMS error was less than 2 pixels. Planimetric errors of the order of 60 m can be easily tolerated because both structural measurements and mapping usually concern features of several kilometres in length [13, 14]; in any case the value is close to the resolution of the Landsat image.

Lastly, the panchromatic, multispectral and thermal bands of each Landsat scene were resampled, respectively, to 15, 30 and 60 meters using the nearest-neighbour resampling method.

Eighty-five Landsat ETM 7+ images (multispectral and panchromatic data) covering the entire territory of Italy were orthorectified using 861 GCPs collected by SITOGEO. Moreover, the SITOGEO application was used to archive the GCPs in an interactive database and to attribute features to each point. The searchable database contains the following information on each GCP: a unique identification number (ID); the Landsat numerical identifier for the GCP (IDL); the orbital path (PATH); the row identifier (ROW); a Landsat name; the Cartographic Reference System of the orthorectified image (Datum, Projection, and Zone); x and y coordinates of reference GCPs (UTM-WGS 84, zone 32/33), x and y coordinates of the source GCPs (UTM-Clarke 1866 zone 32/33); x and y residual and RMS values of multispectral imagery (in pixels); x and y residual and RMS values of the panchromatic band (in pixels); the AIMA name; the region, province, municipality, locality, and description of features used as reference GCPs and the scene (path and row).

The collected GCPs have elevation ranging from 0 m to 1701 m a.s.l. Most points lie on the plain, which has been significantly modified by human activity (313 of 816 GCPs are 0 to 2 m a.s.l.). The most commonly selected features are wharves with wet docks (206 GCPs), buildings (128 GCPs); this database can be updated and queried in either geographic or standard mode and can be used for the geometric correction of new data.

6. – Conclusion

GIS systems are able to store, update, manipulate, display and analyze several forms of spatial information, so the application fields range from natural resource studies to land management. The paper presents a geographic database, "SITOGEO", covering the entire territory of Italy with data of different nature and source (digital elevation model, satellite and airborne images, cartographic digital maps, vector data and alphanumeric data about administrative information). This database is not only a simple static archive storing territorial information but, most of all, an interactive system that can be updated implementing new data (*i.e.* a new digital elevation model of Italy with higher resolution; the airborne images of all national territory stored as 24 bit in RGB format; the Landsat ETM+ geometrically corrected).

We have shown the potential value of SITOGEO presenting two specific applications. The first regards the use of SITOGEO to preserve the cultural heritage; it allows to create a detailed census of cultural assets at a municipal level and to draw up hazard maps in case of natural disaster. The second shows how the SITOGEO and remote-sensing technologies complement each other; SITOGEO permits the acquisition and the storage of a large number of GCP used to orthorectify the satellite images, and moreover the geometrically corrected images can be updated and stored in SITOGEO.

REFERENCES

[1] PARESCHI M. T., CAVARRA L., FAVALLI M. and GIANNINI F., *Nat. Hazard*, **21** (2000) 361.
[2] CAVALLARO A., GALANTI E., MERIGGI A. and PARESCHI M. T., S.I.T.O.–G.E.O. (2001).
[3] Environmental systems research institute (ESRI) ArcView GIS. ESRI, Redlands, 1996.
[4] JORDAN G., MEIJNINGER B. M. L., VAN HINSBERGEN D. J. J., MEULENKAMP J. E. and VAN DIJK P. M., *Int. J. Appl. Earth Obs. Geoinf.*, **7** (2005) 163.
[5] OGUSCHI T., TATSUTO A. and NOBUHISA M., *Comput. Geosci.*, **29** (2003) 885.
[6] CAROZZO M. T., CHIRENTI A., LUZIO D., FEDERICI G., MARGIOTTA C., QUARTA T., TUNDO A. M. and ZUANNI F., *Boll. Geod. Sci. Affini*, Vol. **XLIV**, 1 (1985) 37.
[7] U.T.P. Protezione Civile del Comune di Firenze – S.I.T.M.e.t S.P.A. Carta delle Alluvioni del 1966, 2001.
[8] BISSON M., FAVALLI M., FORNACIAI A., MAZZARINI F., ISOLA I., ZANCHETTA G. and PARESCHI M. T., *Int. J. Appl. Earth Obs. Geoinf.*, **7** (2005) 217.
[9] VIEDMA O., MELIÀ J., SEGARRA D. and GARCIA-HARO J., *Remote Sens. Environ.*, **61** (1997) 383.
[10] AGRICOLA B., PARESCHI M. T., VENDITTI A., FACCIORUSSO M., FAVALLI M. and BATTISTINI A., in *Proceeding Sixth National Conference ASITA, Perugia 5–8 November* 2002, pp. 71–76.
[11] FAVALLI M. and PARESCHI M. T., *J. Geophys. Res.*, DOI: 10.1029/2004JF000150 (2004).
[12] ERDAS, 1999, ERDAS Field Guide, Fifth Edition, Atlanta, GA.
[13] JOHNSON C. A. and HARRISON C. G. A., *Remote Sens. Environ.*, **28** (1989) 273.
[14] TIBALDI A. and FERRARI M., *Geocarto Int.*, **1** (1988) 3.