Paolo Ardissone – Francesca Balossi Restelli – Emanuela Borgia – Emanuele Brienza

Web-GIS Solutions for the Analysis and Valorisation of Archaeological Sites in the Mediterranean Basin

Abstract: The Department of Archaeological Science of the Sapienza Università di Roma and the Land, Environment and Geo-Engineering Department of the Politecnico di Torino are involved in a research project (ARCHEOMEDSAT), with the aim of defining the architecture and components of an information system suitable for the management of excavation data and the provision of innovative services to the visitors of archaeological sites in the Mediterranean area. Open source software and web accessibility are considered a priority for the fulfilment of such an aim. This system also has to be designed to facilitate research activities in these sites and the elaboration of the data retrieved. It should become a fundamental tool for archaeologists and other researchers, and permit complex queries and multifactorial analyses of stratigraphic and architectural data, and that of the single objects found. Here too, web accessibility has proven useful for a multidisciplinary international collaboration between researchers.

Introduction

Seven different Italian universities are at present involved in the ARCHEOMEDSAT project (http://www.ricercaitaliana.it/firb/dettaglio_firb-RBNE03YP9J.htm; http://topografia.unica.it/index.php?option=com_content&task=section&id=9&Itemid=63 [28 Nov 2007]) funded by the Italian Ministry of University and Research, with the aim of creating a system for the management, interpretation and realisation of all data obtained from extensive archaeological excavations. This system is also intended to support research activities and to facilitate their management and logistics. This not only involves the excavation of the archaeological sites, but also surveys of the territory, restoration and conservation interventions and access by non-specialists to the sites and data. To achieve the goal of this research project, the first step is to organise the information obtained from the archaeological investigations into a data model, and to define rules, conventions and file formats that can be integrated and managed in a GIS environment. This is in fact the role of our research group: to conceive and set up GIS solutions that can be used in different archaeological excavations.

The Choice of the Archaeological Sites

Three sites, which are representative of the different situations in the Mediterranean area and are all being excavated by groups from the Sapienza University (http://w3.uniroma1.it/archeologia/Ricerche%20e%20scavi/Ricerche.htm [28 Nov 2007]), have been chosen as samples on which to build and test our system: Arslantepe (http://w3.uniroma1.it/arslantepe [28 Nov 2007]), a long-established settlement in Eastern Turkey, occupied in prehistoric as well as more recent times; the north-eastern slopes of the Palatine Hill and Meta Sudans, part of the ancient city of Rome; and Elaiussa Sebaste, a Hellenistic, Roman and Byzantine city on the Mediterranean coast of Southern Turkey. The three sites were chosen because of their different geographical settings and layouts, and historical context, in order to have a wider view of the exportability and communicability of the system that was being built.

Arslantepe is a tell, a typical Near Eastern mound, made of more than 30 m of superimposed mud brick architecture and archaeological occupation levels, spanning from the 5 millennium BC to the Roman and Byzantine periods. The stratigraphy of Near Eastern mounds is very complex and particular, because of the common practice of cutting terraces into earlier occupations, levelling out or moving the terrain to prepare for new buildings. The most impressive finds on the site are those of a Palatial complex of more than 2000 m^2 dating back to 3000 BC: it was on this monument that the first data formalisation has started. Excavation of this site started in 1961 and is now directed by Prof. Marcella Frangipane (http://misart.it/; http://www.ricercaitaliana.it/firb/dettaglio_firb-RBNE03YP9J.htm [28 Nov 2007]).

The Palatine and Meta Sudans, in the centre of both ancient and modern Rome, were subject to later rebuildings, changes, obliterations and superimpositions, which makes the management and organisation of research and exhibition of the results difficult. Buildings and interventions in the area date back from at least the 6th century BC (an ancient road network) to the 4th century AD. Amongst the many important finds during the excavations, which started in 2001 under the direction of Prof. Clementina Panella, are a sanctuary, in use from the 6th century to Nero's fire of 64 AD, part of the Domus Aurea, and the Meta Sudans fountain itself, which was used with different re-buildings from the Augustean period to the age of Constantine.

Elaiussa Sebaste, where the excavations started in 1995 under the direction of Prof. Eugenia Equini Schneider, is a 23 ha wide city, with two ports, and was first settled on a promontory and later extended into the mainland. So far, living quarters, activity areas (cisterns, wells, kilns), monumental public buildings and necropolises have been brought to light. A huge Byzantine palace, a small basilica, a bath, and some private houses were excavated on the promontory. The Roman monumental area on the mainland includes a theatre, a commercial agora, a public bath building and a huge Corinthian temple which rises on the hill to the south of the city. This site is a case of extensive territorial excavation, in which complex geological and topographical information has to be matched with archaeological and historical data.

Field Documentation Procedures and their Compatibility with a GIS Structure

Distinct methods for the acquisition and organisation of data had been adopted at the three sites as a consequence of the specific necessities and differences between the areas. Whilst the Palatine and Elaiussa excavation system has a common basis, that of Arslantepe is completely different from the other two. This is due to the distinct character of the archaeological finds. The common "stratigraphic unit" forms are not used in Arslantepe, but the format of the documentation entirely depends on the interpretation of the finds, as made directly on the field. Houses are recorded with a house number, rooms with a room number, hearths within a room have the room and hearth number, etc. This creates a kind of pyramidal structure, with the largest entities at the top and the single finds, contained in these entities, at the bottom (for example, you would record, with distinct and specific forms: a building, its rooms, their walls, floors, layers filling the rooms, objects in the layers and on the floors, etc).

The more widely-used system of "stratigraphic units" is instead used in the other two sites, though with some distinctions in vocabularies or definitions stemming from the differences between the two sites. The finds are treated in a similar manner: each category, such as pottery, coins, bone objects, metals, glass, architectural elements, inscriptions, is recorded into specific forms which are directly connected to their context of retrieval. This makes the objects the central element of the GIS system adopted in Rome and in Elaiussa Sebaste, with a sort of radial organization of all the data related to these.

In the long formalisation process of the excavation information, which was carried out in order to implement a GIS oriented data structure, the Arslantepe documentation system proved to be particularly "GIS friendly"; information, in fact, is already greatly formalised by the recording system itself. The construction of the database was more complicated, however, because many more tables and vocabularies had to be built.

Only in the case of the Palatine excavation had a GIS system, running on ESRI software, already been constructed, whilst in the other two cases, work had to start from the very beginning with the formalisation and organisation of data (site reports, forms, survey data, architectural plans, photos, etc.). Creating terminology and formalising the data was not only useful in making it compatible with GIS computer technology, but it has offered a critical methodological evaluation of the documentation systems, which led to the consequent modifications of some specific matters.

However, not all datasets can be reduced to or summarised by a preset term or sentence. In this, archaeology still shows its humanistic aspect. Archaeologists need text and memo entries in which information that is still uninterpretable or incomplete can be collected, something which is disliked by computer specialists. Even though the data in these records might not be used immediately for any sort of analysis, it is always possible that subsequent discoveries will be able to offer an explanation. Therefore, in constructing the database, such entries still remain, albeit greatly reduced in number. In view of future developments, an archaeological GIS cannot be a closed structure, but must be expandable and infinitely modifiable.

The differences in recording systems for one site and the diversity of the finds and thus of the set vocabularies for the other two, led us to build three distinct GIS structures using file formats which are easily exchanged across different platforms.

Data Management Policy

When designing the GIS data model, three main problems had to be faced: firstly, the different approaches, investigation methodologies and procedures for data collection adopted by each of the archaeological groups involved had to be preserved; secondly, the system had to permit both the direct acquisition of new excavation data from the field, coded according to the project requirements, and the *ex post* reorganisation of datasets obtained from studies performed by different specialists, or in the past; thirdly, security, conservation and portability of the stored data and of the system itself had to be guaranteed.

Notwithstanding the diversities between the three sites and their documentation procedures, we initially tried to conceive a unique GIS data structure shared by the three sites. We soon realised, though, that such a task was impossible as the necessary homogenisation of the data would have invariably resulted in a great loss of precious information. From a theoretical point of view too, we concluded that a single, universal recording system was not desirable: different historical periods, geographical regions, geological conformations, and thus archaeological contexts, even though excavated in the same manner and with the same scientific archaeological methods, have their own specificities, historical and anthropological issues, and we therefore agreed that such distinct archaeological contexts as those of the three sites here sampled should keep their own documentation procedures.

The excavation proper and recording of finds is only the first and in some ways the smallest task of archaeologists, who then go on to analyse and interpret the data. The role of GIS is that of joining or comparing information coming from these different research activities, and allowing the analysis of the data stored in it. Therefore, the data model had to allow both these levels of implementation and use. The data model was structured so as to preserve the original logic, to ensure accuracy and to state the source of the data. In general, specialists' requirements were mediated with the necessity common to each system of having information organised and coded in an efficient manner. Moreover, unlike "traditional" databases, GIS technologies require alphanumeric and spatial information to be structured and managed together, making the implementation task more complicated. The attributes and the relationships between data have to be matched by the definition of the relationships between information and spatial data.

In summary, therefore, we planned a specific GIS-oriented data model for the two archaeological sites of Arslantepe and Elaiussa Sebaste, adapted to the requirements of the different contexts, and at the same time we designed a common technological approach, aimed at ensuring longevity, preservation, sharing and dissemination of the data. Whilst contents, data formats and structure, and analysis tools were variable for the two sites, analogous file formats, characterised by a high portability, were suggested for the GIS implementation and for the project conservation, sharing and dissemination. A metadata structure was proposed to describe the contents and format of the data and to allow an initial evaluation of the resources stored in the system. In this way, the system is able to create data that is loadable and manageable by different GIS software, both desktop and web-based, both commercial and open source. Consequently, other GIS projects will be able to integrate this data and gain an insight into the strengths and weaknesses of our method. The existing GIS from the Palatine Hill and Meta Sudans excavations was used to test the capacity of communication between the different systems, as well as an aid in the development of the database for Elaiussa Sebaste.

Translation and Reorganization of Data in Digital Format

The documentation acquired during the excavations can be divided into two main classes: alphanumeric data and spatial data. The three archaeological teams have a long tradition in field data acquisition (both for the spatial and the alphanumeric information), by now standardised and well coded. This was of course respected and maintained since it best describes the archaeological contexts investigated. An ER Model, reproducing the logic of the alphanumeric documentation of the paper archives and defining some rules, standards and guidelines for the creation of CAD files, was created. In this way, data has been and can continuously be imported into a digital format designed to aid the implementation of the GIS project. The link between the spatial and alphanumeric components can then be made in a GIS environment, as described below.

Alphanumeric Data

As noted above, the structure of the data collection used by the archaeologists has been followed as far as possible. Translating a data collection system based on a paper archive into a digital information system *(Fig. 1)* requires, as can intuitively be imagined, the definition of several rules, that might be implicit and only partially followed and checked in a paper archive. Classes of entities which can be identified and described with a common set of attributes have thus been initially flagged.

The translation from the paper archive to the database tables was also an opportunity to optimise several aspects of the data collection. Lists of coded attributes were used in the database structure as much as possible. This can help in simplifying the system, where sets of coded attributes common to different tables can limit errors in the data in-put.

In the first phase of the project, Microsoft Access was used to create and test the ER model and to define the coded attributes and vocabularies. Then, tables and records stored in the database were exported to .csv files and in this way were formalised and made available for GIS implementation.

Spatial Data

The graphic documentation produced on the field is traditionally imported into a CAD environment after excavations are complete. The drawings represent the areas, structures, and objects using the descriptive capability of the CAD software, rather than representing reality using symbols, layers, or predefined forms for fixed types of objects. The on-field survey data is not referenced back to the alphanumeric data during data entry.

The computer-aided (CAD) software programs are conceived to help users generate and visualise the graphic representation of the data, without adding additional information. GIS, instead, requires information that is structured and organised according to its meaning and its querying possibilities. Practically speaking, the CAD files created for strictly archaeological purposes are useful for their graphic representations, but are not directly loadable in a GIS environment and they cannot be integrated with the alphanumeric data.

The drawings, maps, surveys and CAD files have to be reconverted following the data model designed for the GIS *(Fig. 2)*. The CAD files will populate the geometry of the entity classes described by the alphanumeric attributes. CAD environments are not usually able to manage more than one attribute for each single entity. Therefore, the alphanumeric attributes are stored and managed in an external file (.csv format), creating a unique key code to link the CAD objects to their attribute records.

The main problem is that the layers inside a CAD file are often composed of points, polylines, closed and open polygons. GIS software is not able to understand and interpret this conflicting information. It can only manage the geometry of the elements if these are described unambiguously as closed polygons, polylines or points. For these reasons, several guidelines were defined and followed in order to

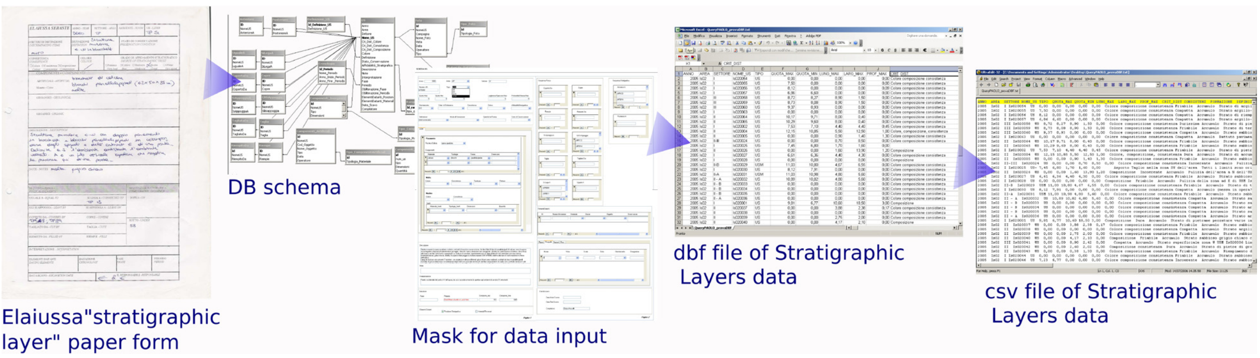

Fig. 1. Alphanumeric data treatment. Examples of paper forms from the excavation of Elaiussa are shown, and their translation to digital formats, first through a common database (Access) – the mask for the data input and the resulting .dbf file are shown here – then to a csv file ready for GIS implementation.

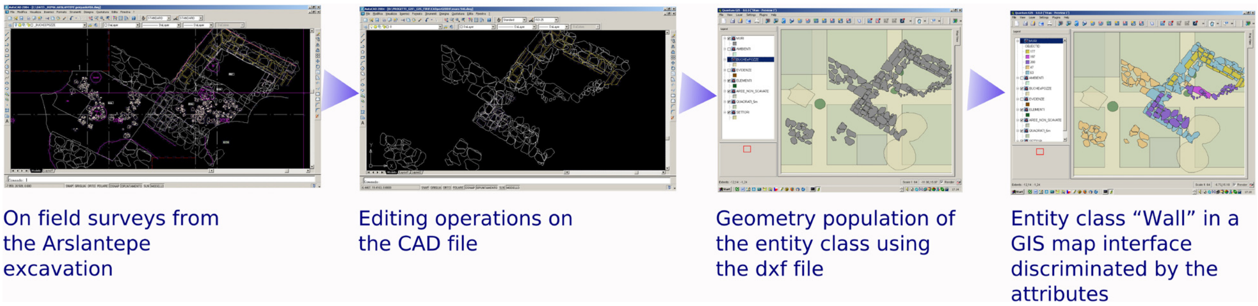

Fig. 2. Spatial data treatment. Archaeological CAD drawings have to be edited to be loadable in a GIS environment. This allows the GIS program to recognise and distinguish all the distinct attributes of the finds and plans drawn.

create CAD files which could be easily imported into a GIS environment. This problem was faced when converting already acquired data, but above all when planning the acquisition of new data in future campaigns.

The conversion of the CAD data, in order to facilitate its communication with GIS, requires long and complex editing operations. Entities have to be singled out and reorganised in new specific layers. Areas have to be represented by closed polygons. Lines have to be continuous polylines, and the geometrical information of the entities has to be coherent with the class they are going to populate. This operation can only be done manually. Automatic procedures, as the programming of a batch processing tool, were taken into consideration, but it was soon clear that the manual interpretation of the survey outputs by specialists was the only adequate and effective procedure.

Creation of a Supporting Management System Based on a Web Architecture

A flexible usage of the GIS and database had to be guaranteed, both in the data collection phase and in the completion of the system, in order to allow it to be updated at various phases, possibly by different people in different places or even simultaneously. A supporting management system is thus now being created based on a web architecture.

The first and undeniable advantage of using Internet technology to publish GIS data is the simplicity and economy of the implementation. Users who access GIS resources shared using a web server and published on a specific web site, no longer need particular hardware configurations or software knowledge: it is sufficient to connect to the WWW and use a traditional web browser.

The system is completely based on open source tools, which are considered to be the best solution to achieve the project goals. In fact, with these technologies it is possible to build complete and flexible solutions, by continuously updating software packages and applications and being sure of the efficiency of the tools, minimizing the set-up and maintenance costs.

The server is based on a Linux Debian operating system. Apache Server, which is the most widely-used and powerful, has been selected as the web server.

All the data is stored in a PostgreSQL DataBase Server combined with its spatial extension PostGIS (http://postgis.refractions.net/; *Fig. 3*). In this way, both geometric and alphanumeric information is stored in a single tool, and as a consequence, the architecture of the system is simple and light.

A specific web application based on PHP scripts is being developed to load the alphanumeric data acquired during the excavation phase into the database: specific forms in a web page are available for a correct and controlled data input. In this way data is immediately digitalised and available in the system and for web realisation; this solution also avoids problems of misalignment and duplication. No plug-ins are needed on the client side. Whereas alphanumeric data are uploaded via web and directly stored in the database, and thus immediately accessible, geometric surveyed data are sent in dxf format to the system. These files are structured following the specific directives: with simple office work, spatial data are inserted manually in the database. At present *dxf2postgis* (http://sourceforge.net/projects/dxf2postgis/), a tool converting DXF files to PostGIS geometry tables, is being tested. Using *dxf2postgis*, a single DXF file is converted to a PostgreSQL-PostGIS SQL script, to create and fill tables. While the connection between features and attributes is guar-

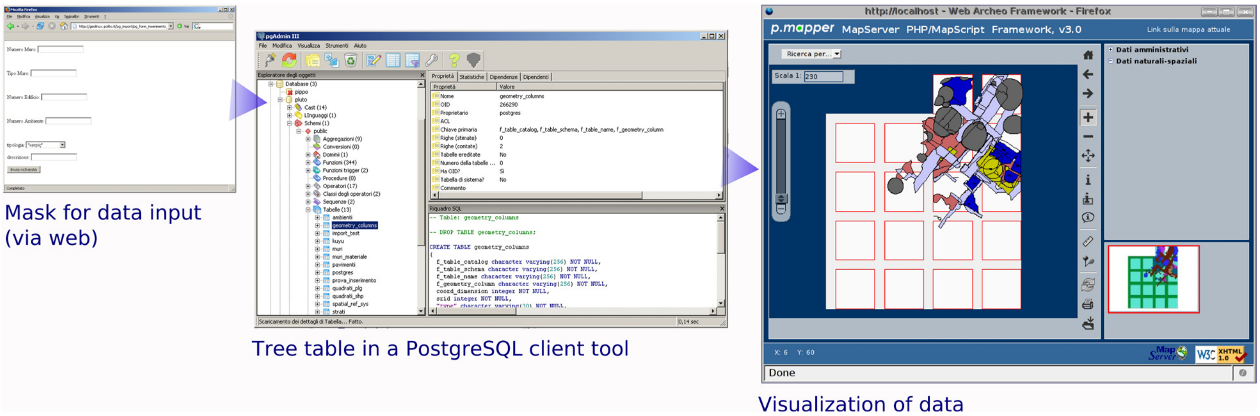

Fig. 3. Examples of data input, managment and visualisation on the PostgreSQL Database and PostGIS.

anteed by the group of structured rules for CAD data acquisition and implementation identified in the first phase of the project, a procedure for checking this has been developed and implemented with a specific tool based on the QGIS (http://qgis.org/) interface. Using QGIS, specialists can access the PostGIS geometry table, check and edit the spatial data and then assign to each entity its unique key code to link it to the related attribute records.

The geometric data is published using the University of Minnesota MapServer (http://mapserver.gis.umn.edu/), which is an open source development environment for building web mapping applications; the web user interface is based on a MapServer client open framework called Pmapper, which takes advantage of Ajax technologies with easy navigation and simple analysis tools.

The web page for data insertion has restricted access and only authorized users can log in and add to the data, while further security settings are adopted in the firewall and Apache configuration on the server.

Each authorized user has their own account on the server and all operations performed are recorded. Suitable security settings are adopted for data storage safety: the hard disks have raid controllers to avoid data loss in case of a disk breakdown.

Concluding Remarks

The use of GIS in archaeology is common nowadays and its value is becoming more and more evident when it comes to storage, management, analysis and sharing of data which otherwise would be accessed with difficulty. It is even more so in the case of the three sites used as case studies in our application, because of the long duration of these field projects and the immense amount of data acquired.

What this project is trying to tackle in greater depth is how to combine these needs with those of publication and communication of the results of this research to the wider public. The choice of working within a web framework seems to be the right one in this case, particularly as it easily permits simultaneous implementation and work on the system. Web access permissions or restrictions should also be analysed in more detail, and a system for the protection of unpublished and unprocessed data created.

The research project also shows how, after an accurate GIS data model definition, open source technologies can be exploited to create low-cost products that can be customized according to individual requirements. Open source also allows the system to be constantly updated without any further expense.

ARCHEOMEDSAT is an on-going project: the coming year will be devoted to a thorough testing of the system aimed at its correct functioning for everyday activities, on field, and in the post-excavation study phase. We shall also upload the system onto the web for an experimental phase with restricted access, during which we shall request feedback from the first users in order to correct potential errors and make it more user-friendly.

Acknowledgements

Our thanks go the coordinator of the ARCHEOMEDSAT project, Prof. Manzoni, for having first formu-

lated the ideas and proposed tackling them. The work of Professors Equini Schneider, Frangipane, and Panella, directors of the archaeological excavations at Elaiussa Sebaste, Arslantepe and Palatine for many years, is not only instructive, but of constant guidance and inspiration, providing the best solutions to many of the methodological and theoretical problems tackled here. ARCHEOMEDSAT involves seven Italian universities (the Universities of Trieste, Pavia, Palermo, Milano, Cagliari, Torino, and Roma), and it is thanks to the cooperation between all that such significant issues can be so broadly and completely analysed.

References

Biswel et al. 1995
S. Biswel / L. Cropper / J. Evans / C. F. Gaffney /P. Leach, GIS and excavation: a cautionary tale from Shepton Mallet. In: G. Lock / Z. Stancic (eds.), Archaeology and geographic information systems: a European perspective (Somerset 1995) 269–286.

Brienza 2005
E. Brienza, Un GIS intra-site per la Valle del Colosseo. Geomedia 1–2005 (Roma 2005).

Equini Schneider 1999
E. Equini Schneider, Elaiussa Sebaste I. Campagne di scavo 1995–1997 (Roma 1999).

Equini Schneider 2003
E. Equini Schneider, Elaiussa Sebaste II. Un porto fra Oriente e Occidente (Roma 2003).

Frangipane 1997
M. Frangipane, A 4th millennium temple/palace complex at Arslantepe-Malatya. North-south relations and the formation of early state societies in the northern regions of Greater Mesopotami. Paléorient 23,1, 1997, 45–73.

Frangipane 2004
M. Frangipane, Alle origini del potere. Arslantepe, la collina dei leoni (Milano 2004).

Frangipane / Palmieri 1988
M. Frangipane / A. Palmieri, A protourban centre of the Late Uruk period. Perspectives on Protourbanization in Eastern Anatolia: Arslantepe (Malatya). An Interim Report on 1975–1983 campaigns. Origini 12,2, 1988, 287–454.

Gillins / Wise 1998
M. Gillins / A. L. Wise (eds.), GIS guide to good practise (Oxford 1998).

Panella 1996
C. Panella (ed.), Meta Sudans I. Un'area sacra in Palatio e la valle del Colosseo prima e dopo Nerone (Roma 1996).

Panella 2001
C. Panella, La valle del Colosseo prima del Colosseo e la Meta Sudans. In: A. La Regina (ed.), Sangue e Arena (Roma 2001) 49–67.

Panella / Zeggio 2004
C. Panella/S. Zeggio, Tra Palatino e valle del Colosseo: nuovi dati. Workshop di Archeologia Classica 1 (Pisa, Rome 2004) 65–87.

Pensabene / Panella 1999
P. Pensabene / C. Panella (eds.), Arco di Costantino. Tra archeologia e archeometria (Roma 1999).

Wheatley / Gillings 2002
D. Wheatley / M. Gillings, Spatial technology and archaeology: the archaeological applications of GIS (London 2002).

Paolo Ardissone
DITAG – Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino
Italy
paolo.ardissone@polito.it

Francesca Balossi Restelli
Emanuela Borgia
Emanuele Brienza

Sapienza Università di Roma
Dipartimento di Sc. Storiche
Arch. e Antrop. dell'Antichità
P.zzle Aldo Moro, 5
00185 Roma
Italy