# Across Space and Time

Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology
Perth, 25-28 March 2013

Edited by Arianna Traviglia

Amsterdam
University
Press

Computer Applications & Quantitative Methods in Archaeology

# Across Space and Time

## Papers from the 41st Conference on Computer Applications and Quantitative Methods in Archaeology

*Perth, 25-28 March 2013*

CAA 2013 PERTH

Edited by
**Arianna Traviglia**

AMSTERDAM UNIVERSITY PRESS

# Estimation of Archaeological Potential with a Page Rank Based Predictive Model: the Urban Area of Pisa

**Nevio Dubbini and Gabriele Gattiglia**
*University of Pisa, Italy*

**Abstract:**
*We present an analysis of multi-faceted, GIS managed data for determining the archaeological potential of the urban area of Pisa (MAPPA 2014a). The identification of relations among finds is a key issue for the data mining in the archaeological interpretation process: spatial and functional organisation levels allow only particular configurations of parameters defining the potential. A PageRank based model is applied, since the criteria for assigning importance to web pages by search engines are also based on relations. The procedure includes the categorisation of finds; the assignment of initial absolute values of potential to the available data, through an automatic procedure; the definition of functional areas (urban, suburban, rural); and the application of a PageRank based algorithm. The results, including the map of archaeological potential, are to be considered as the first steps towards an automatic, formally definable, and repeatable approach to the computation of archaeological potential.*

**Keywords:**
*Predictive Modelling, Archaeological Potential, PageRank, Archaeological GIS, Geomorphology*

## 1. Introduction

The MAPPA project is a research project in which archaeologists, geologists and mathematicians have studied predictive modelling tools applied to the archaeological potential of an urban area. The map of archaeological potential is a technical but above all conceptual development of common archaeological maps. All the information taken from excavations, written sources, archive documents, and aerial and satellite photographs is included in a map of archaeological potential. Moreover, this is a predictive map: the estimation of the probability that certain areas may conceal unknown archaeological remains is achieved by projecting the knowledge regarding neighbouring areas onto them, with a degree of approximation that varies according to the quantity and quality of available data.

In this study we present results obtained by an analysis of multi-faceted, diachronic, GIS managed data for determining the archaeological potential of the urban area of Pisa. The need for dealing with the urban context implied the necessity to consider archaeological finds instead of archaeological sites, and to develop a new predictive model (Dubbini and Gattiglia 2012) that takes also into account the heterogeneity and the complexity of datasets. We considered a number of datasets, concerning different aspects of the problem: archaeological data, buildings archaeological data, historical data, toponymic data, geomorphological data. The archaeological GIS data model was developed to manage heterogeneous data, which represent the urban archaeological complexity. We worked with both topographic (e.g. geomorphologic, hydrographic, toponymic data) and urban data (e.g. archaeological stratifications, buildings, road network, hypotheses of historians and archaeologists). The archaeological data model combined raw data and interpreted data, and range from less synthetic data (i.e. the context level) to more synthetic data. The key unit of the data model was the archaeological intervention, but the model also included the filing of published data, archive data and data resulting from building archaeology, and georeferencing and vectorisation data in order to understand the urban fabric development and the level of architectural heritage preservation. In addition, we considered: (i) the collection of written and published documentary sources with the aim to locate no longer existing place names, production activities, infrastructures and topographic structures; (ii) the computerised acquisition of

*Corresponding author: nevio.dubbini@gmail.com*

historical mapping to trace urban transformation throughout the modern and contemporary ages.

Preliminary to the implementation of the mathematical model and to the creation of the input files, available archaeological data had to be categorised assigning each find to a category, in order to effectively implement the algorithm, and to make the results general enough to be applied also in different contexts (Fabiani and Gattiglia 2012). The archaeological potential represents the possibility that more or less significant archaeological stratification is preserved. The following parameters were identified to estimate the archaeological potential: type of settlement, density of settlement, multi-layering of deposits, removable or non-removable nature of the archaeological deposit, degree of preservation of the deposit, and the depth of the deposit (Anichini et al. 2011).

The aim of this paper is to describe the mathematical model and the algorithm used to estimate archaeological potential, and to present some results about the case study area: the urban area of Pisa. First, we had the problem of properly defining the concept of archaeological potential. To this aim, it was necessary to distinguish between what we called absolute potential and archaeological potential. The absolute potential (i.e. initial conditions) is defined in order to assign a potential value to the available data, while the archaeological potential (i.e. the output of the algorithm) is an estimate of the informative value of the archaeological stratification. Both measure of potentiality were divided into 5 different levels, by partitioning uniformly the numerical value in 5 parts. The division into levels is particularly useful for large-scale urban and landscape planning, and for users to have an overview of the areas with greater or lesser potential. Level 1 refers to (especially extra-urban) areas containing mostly natural stratifications or stratifications reworked by man (e.g. floodplains involving agricultural work), with low information potential. Level 2 refers to contexts similar to the previous one; however, evidence from photointerpretation suggests more marked intervention by man: e.g. traces of agricultural partitioning or road connection networks. Rural structures may also be found, as well as necropolis areas and other deposits with high information potential, but limited diachrony. Level 3 refers to areas characterised by variable information potential, usually restricted to certain historical periods. Level 4 refers to areas with complex stratigraphic deposits and high, averagely developed (in diachronic terms) information potential. Lastly, level 5 potential refers to an area featuring the highest information value and high diachrony of archaeological deposits.

As for the value of absolute potential of finds, it was obtained by summing two different components, concerning the type of archaeological information and the removable or non-removable nature of archaeological finds. The archaeological reasoning behind this rating is given by the necessity to formalise (in our approach, in a binary manner) the amount of archaeological information that each find can produce.

- The type of archaeological information

  To compute the value of this parameter a list of 19 areas of interest were drawn up, corresponding to the main informative fields on which the finds can provide information. The identified areas of interest are: production, building techniques, trade, food, agriculture/breeding, worship, waste management, political/institutional aspects, social and gender aspects, physical anthropology, fauna/flora, geomorphology, viability/transport, health and hygiene, warfare, land management, leisure, tradition, water system. After having defined those areas of interest, categories of finds were assigned a value of absolute potential by summing, for each area of interest, the value *1* if the category provided information on that informative field, and *0* otherwise. For example the category "domus" was given the value *12* since a *domus* can provide information about *12* areas of interests.

- The removable or non-removable nature of the archaeological finds

  This parameter concerns the persistent/transient nature of each archaeological trace. It is assumed that more structured archaeological remains (e.g. a stone building) have higher information potential, since there is a greater chance of identifying those remains in the archaeological deposit. The values assigned to this parameter are 1 for masonry-built items (e.g. 'insula', 'tower house', 'domus', 'palace', 'theatre', 'church', 'prison',' forum', 'church'); 0.5, for items envisaging masonry but which can be less structured (e.g. an adobe building, 'river bank', 'enclosure', 'henhouse', 'pigsty', 'stable', 'hearth'); 0, in the case of items without masonry work (e.g., 'canal', 'reclamation', 'trench', 'agricultural land', 'clearing', 'camp', 'waste dump') (Anichini et al. 2013).

In this way the value of absolute potential can range from 1 to 20. Once the value of absolute potential has been defined, we had a way of transforming the available archaeological data into numerical values, suitable to serve as an input for the mathematical model. The other main ingredient of the model is a way of representing relations among finds. The identification of relations among finds is a key issue for the data mining during the archaeological interpretation process. In urban areas the spatial and the functional organisation provide meaningful information for the automatic extraction of possible configurations of the parameters defining the potential. In other words, depending also on the archaeological period we are considering, it is possible to distinguish parts in which only particular configuration of parameters that define the archaeological potential are feasible, or most probable. So the relations among finds can strengthen or weaken the archaeological potential of the area itself. The authors showed how a PageRank based model (Bini, Dubbini and Steffè 2011; 2012) can be used to assign the archaeological potential: the criteria used for attributing archaeological potential and those used for assigning importance to web pages by search engines are both based on relations. The method, moreover, has revealed to be relevant in the urban context, where, to the knowledge of the authors, no existing predictive model has been applied.

## 2. Problem Definition: Estimation of the Archaeological Potential

All the data listed in this section were managed as different GIS layers. The mathematical analyses were carried on in a matrix ambient (Matlab®), and the matrix produced by the algorithm was converted in an output raster, so readable by a GIS application.

The subsurface was divided into 3-dimensional cells forming layers, each one covering the whole study area. The number of layers is equal to the number of archaeological periods, because this is a straightforward way to distinguish between the relations acting inside the same archaeological period and the relations acting through periods. The 7 periods the archaeological team defined were: Protohistory, Etruscan Period, Roman Period, Late Roman Period, Early Medieval Period, Late Medieval Period, Modern Age, Contemporary Age. We therefore divided the subsurface of the work area in $n = n1 \times n2$ cells for every one of the 7 layers. For each of the $7n$ cells the absolute potential value was given on the basis of available data, as described before. So at the initial stage, the value of each cell is equal to the archaeological absolute potential of that cell. The problem is that of estimating the archaeological potential in every cell. On the basis of the information of different kind resumed from available data, the following input data were created, to be processed by the algorithm. The files on the list below were created for every archaeological period.

- **Certain geolocation data**: certain geolocation data were defined here as the data with known spatial coordinates. Notice that data was considered as certain geolocation data even where the dating was uncertain, that is, what defines the category of certain geolocation data is the quality of the knowledge of the spatial 2-d coordinates. For instance, we considered as certain geolocation data from excavations with known spatial 2-d coordinates, as well as data relating to aerial photography anomalies. Certain geolocation data were organised in a $n1 \times n2$ matrix, whose element $i,j$ represents the absolute potential of that cell given by the certain geolocation data inside the cell.

- **Uncertain geolocation data**: uncertain geolocation data were defined here as the data with unknown spatial coordinates, i.e. data for which we only knew that they are located in a certain spatial region. For instance, we considered as uncertain geolocation data, data from excavations with uncertain spatial 2-d coordinates, as well as data from medieval written sources. The values of absolute potential of uncertain geolocation data are computed by dividing the absolute potential of the category of find by the number of cells included in the probable geolocation of the find. Uncertain geolocation data were organised in a *n1 x n2* matrix, whose element *i,j* represents the absolute potential of that cell given by the uncertain geolocation data inside the cell.

- **Shapes**: from certain geolocation data, the archaeological team tried to give a shape to finds, e.g. to draw a house from a wall or a floor, or the continuation of a street from a piece of it. Each shape was given an empirical accuracy value varying from *1* to *6* (*1* is the value of maximum precision), expressing how much the size and the orientation of the shape can be deduced with precision, on the basis of finds close by and of the geomorphologic datum (Anichini et al. 2013). The values of absolute potential of shapes were computed by dividing the absolute potential of the category of find the shape belongs to, by the precision value of the shape. Shapes data were organised in a *n1 x n2* matrix, whose element *i,j* represents the absolute potential of that cell given by the shapes.

- **Geomorphologic data**: these data were deduced from geological survey, in order to identify, for each archaeological period, the diverse geomorphologic features, which were distinguished in river, floodplain, wetland, marshy area, morphological high. Also the geomorphologic datum was given an absolute potential value, by summing for each geomorphologic feature the absolute potential of all the categories of finds that can be present in that geomorphologic feature. Geomorphologic data were organised in a *n1 x n2* matrix, whose element *i,j* represents the absolute potential of that cell given by the geomorphology.

The input files defined above, and their values of absolute potential, were used in different ways, as will be shown later in the section describing the algorithm. In addition, the algorithm makes use of functional areas, i.e. levels of spatial and functional organisation in which the urban space is organised (e.g. urban, suburban, rural areas). Each urban centre, in each archaeological period, is surrounded by a suburban area, and, more externally, by a rural area. Apart from the different names and functions that those three functional areas assumed in each

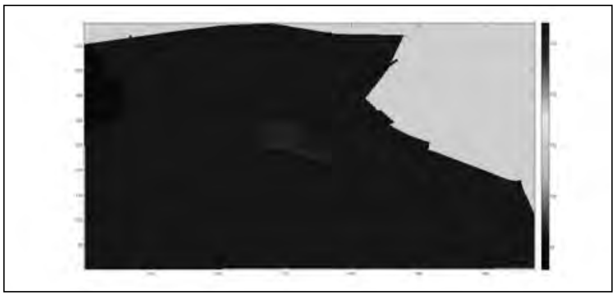

***Figure 1**. The geomorphological datum, as obtained for the Late Medieval Period: the floodplain is in dark red, the morphological high is in red, the river, the wetlands, and the marshy areas are in (different shades of) blue.*

archaeological period, it was useful to define them in a rather general way. The identification of the functional areas is based on many elements, since it depends on the different settlement types, the relationships among them and the environmental context. To limit the subjectivity in defining the functional areas for each archaeological period, we used an automatic procedure to define them. We describe the basic procedure below, disregarding any minor variations due to the particular characteristics of some archaeological period.

- First, each category of finds has been associated with the functional areas it can belong to, and with the functional areas it cannot belong to. On the basis of that two different types of finds were defined: those characterising a functional area (characterising finds), and those not characterising any particular functional area (non-characterising finds).

- For each find characterising a functional area, a *200 m* (*300 m* for the Etruscan period) circular buffer is defined. All the non-characterising finds in that buffer are absorbed in that functional area.

- In a following step, non-characterising elements already included in the buffers in the previous step, generate new buffer of *50%* of the length of the original radius, enlarging in this way the functional area.

- All the non-characterising finds define a *100 m* radius circular buffer. If at least *5* finds are inside that buffer, and each one of them is less than *100 m* far from a find characterising the urban

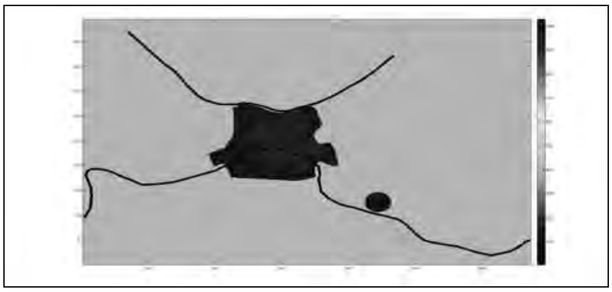

***Figure 2**. The functional areas, as obtained by the described procedure for the Late Medieval Period: the urban area is in red, the suburban area in dark red, the rural area in sky blue.*

functional area, then the 5 finds and their buffer are included in the urban functional area.

- The expansion of each functional area continues until some buffer of a functional area meet a buffer of another functional area, determining in this way a boundary between the two functional areas.
- After all functional areas have been defined, the cells not belonging to any functional area are assigned to the rural functional area, except for those cells in the geomorphologic values of "river", "wetland", "marshy area".
- In order to use functional areas in the algorithm (see 3) , an absolute potential value was defined. In a similar way to what was done with the geomorphologic input, for every functional area the categories of finds that can be inside the area were listed, and their absolute potential values were summed to give the value of the functional area (Anichini et al. 2013).

After the definition of the functional areas, another two input files were created, expressing synchronic and diachronic associations among finds. Those associations were used in the algorithm to enhance the spread of archaeological potential on the basis of the probable presence of "valuable" finds in the surroundings of particular (categories of) finds.

| INPUT | DESCRIPTION | ABSOLUTE POTENTIAL COMPUTATION |
|---|---|---|
| Certain geolocation data | Data with known spatial (but not necessarily temporal) coordinates | The absolute potential value is given by the sum of potentials of certain geolocation finds inside each cell |
| Uncertain geolocation data | Data with unknown spatial coordinates | First divide the absolute potential of the find by the number of cells included in the probable geolocation of the find.The absolute potential value is then given by the sum of potentials of uncertain geolocation finds inside each cell |
| Shapes | Probable shape of certain geolocation finds. An empirical precision value is assigned (from 1 to 6) | First divide the absolute potential of the find the shape belongs to, by the precision value of the shape. The absolute potential value is then given by the sum of potentials of shapes inside each cell |
| Geomorphological data | Deduced from geological survey, to identify geomorphological features | The absolute potential value is assigned by summing the absolute potential of all the finds that can be present in each geomorphological feature |
| Functional areas | Levels of spatial and functional organization, i.e. urban, suburban, rural areas | The computation procedure is described in the paper |
| Synchronic associations | Each category is associated to the most probable categories founded at a distance <50 m, in the same archaeological period | The values of potential are computed by summing the potential values of the categories associated to the finds located in that cell |
| Diachronic associations | Each category is associated to the most probable categories founded at a distance <50 m, in the chronologically previous or following archaeological period | The values of potential are computed by summing the potential values of the categories associated to the finds located in that cell |

***Table 1**. Inputs of the algorithm.*

- As for the synchronic associations, for all categories of finds, depending on the functional areas in which they are located, each category was associated to the most probable categories found at a distance <*50* m. With the expression “most probable” we mean at least in 75% of the attestations. This input file, for each cell, contains the sum of the potential values of the categories associated to the finds located in that cell.

- As for the diachronic associations, for all categories of finds, depending on the functional areas in which they are located, each category was associated to the most probable categories found at a distance <*50* m (in the 2-dimensional spatial coordinates), in the chronologically previous or following archaeological period. This input file, for each cell, contains the sum of the potential values of the categories associated to the find located in that cell.

The size of (square) cells was chosen to be *10 x 10 m*: this size was the outcome of different factors taken into account. On one hand, the archaeological data could be located with precision, but the geomorphologic data could be given a precision no higher than *10 x 10 m*, due to the number of elevation points available for the creation of each historical DEM; on the other hand there is a trade-off between the size of the cells and the total number of cells covering the work area, so that the smaller the size of the cells, the higher the total number of cells covering the area. Therefore the smaller the size of the cells, the more cells there are for which the archaeological potential has to be estimated. For example, in the limiting case for which we would like to estimate the archaeological potential of each point of the study area with arbitrary precision, we would have a finite number of input data, but we should estimate the potential of an infinite number of points. To solve this problem we searched (numerically) the maximum of a function representing the difference between the amount of information given by the available data in the case that cells are considered as “pixels of an image”, and the area of cells divided by the total number of cells. The amount of information given by the available data is computed as the Shannon entropy (Shannon 1948) of a binary matrix where each cell containing input data is given the value *1*, and the other cells are given a *0* value. So the resulting formula is

$$max_{l \in N} \; w[-p_1 log(p_1) - p_0 log(p_0)] - (1-w)\left[\frac{l^2}{n}\right]$$

where $p_1$ and $p_0$ are defined as the relative frequency of cells with data and with no data respectively in the study area, $l$ is the edge of the cells, and $w$ is a parameter. The maximisation, for different values of $w$ near *1/2*, yield the “optimal” size of cells between *10* and *14* m. For this reason we chose the size of cells to be *10* m. This decision was also taken for practical reasons: since the first trials, data were given for 10 m, and so no other smoothing is needed to adapt data for cells of other sizes. In order to perform computations on values of archaeological potential “per unit of area”, the value of absolute potentials of cells was divided by its area, so making computations as independent as possible of the cell size.

## 3. The Mathematical Model

Mathematically, we applied a PageRank based model. The model not only needs the vector of absolute potential, but also a matrix of weights. A general introduction to the PageRank model can be found in Langville and Meyer (2006), while the application of PageRank based techniques to the estimation of the archaeological potential can be found in Bini et al. (2011). It is important to note that the algorithm was applied to each period separately, so that we have a potential map for each archaeological period, which we can “sum” as a last step (Dubbini 2013) .

Before describing how the model works, we describe how the two inputs of the PageRank based model are constructed, the vector *D* representing available data, and the matrix of weights *S*. The vector *D* is a matrix of dimensions *n1 x n2*, reshaped to a column vector of length *n*, and is obtained by the sum of the matrices of absolute potential of certain geolocation data, uncertain geolocation data and shapes.

The matrix of weights *S* is a *n x n* matrix whose element *i,j* represents the weight (value) of the link between the cell *j* and the cell *i*. We now describe how this matrix is computed:

- Each cell with a find distributes its importance to a square mask of cells centred on the cell itself.

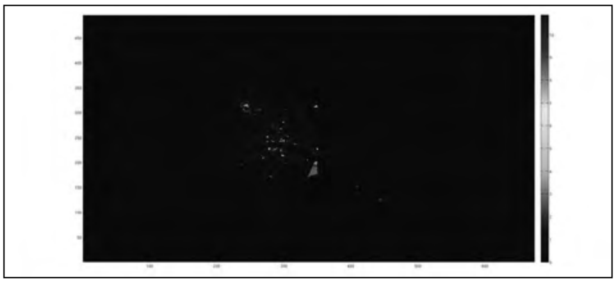

***Figure 3***. *The vector D for the Late Medieval Period, obtained by the sum of the matrices of absolute potential of certain data, uncertain data and shapes.*

The edge of this square is set to be equal to the absolute potential value of the functional area the cell belongs to. We used the values of functional areas in this way since the absolute potential of functional areas was assigned on the basis of the finds that can be present in the functional areas: therefore, in higher valued functional areas, it is justified that the region in which the potential is spread is larger, and vice versa.

- When a mask as in the previous step is constructed, the total weight (i.e. the sum of weights distributed by a cell) inside the mask is given by the value of the functional area plus a value proportional to the sum of synchronic and diachronic associations. We used this value to address the total weight because the "quantity of potential" that a cell can spread is influenced as well by the probability of finding high or low valued finds in the nearby.

- The distribution of weights in the mask around each cell is given by the uniform distribution weighted by the geomorphologic values of the mask around the cell, and weighted by the functional areas values. This is because the geomorphologic datum constitutes a basic influence on the spread or archaeological potential, and because the functional areas values - since they are proportional to the total potential of the finds you can find in the area – can be used to weight the diffusion of potential.

The model uses the data described above, to estimate the archaeological potential. The algorithm is made of a basic procedure, applied repeatedly, consisting in a modification of the standard PageRank, whose output is the vector of the estimated archaeological potential:

1. The vector $D$ representing available data and the matrix of weights $S$ are generated as described above;

2. The following iterations are performed:

*for i = 1,...,1000*

$$A = S \cdot x + \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & \ddots & \vdots \\ 1/n & \cdots & 1/n \end{bmatrix} \cdot x;$$

$$A = (1\text{-}yield) \cdot A + yield \cdot x;$$

$$u = [1\ 1\ \dots\ 1];$$

$$y = rel \cdot (A) + (1\text{-}rel) \cdot [D \cdot (u^T \cdot x)];$$

$$y = \frac{y}{\sum_{i=1}^{n} y_i}$$

$$x = y;$$

*end*

$$D = speed\_up \cdot x + (1\text{-}speed\_up) \cdot D;$$

In these formulas $x$ is a stochastic (i.e. the sum of its component equals 1) random column vector of dimension $n$, used as an initial condition for the application of the iteration described in the "for" cycle (the result of these iterations are independent of the initial condition).

The following parameters are used in the algorithm:

- *Maxit (natural number)* is the number of times the steps 1. and 2. are executed. Each time we

| | |
|---|---|
| Each cell distributes its importance to a square mask of cells | The edge of the square is given by the absolute potential value of the functional area the cell belongs to |
| The sum of weights distributed by a cell | Given by the value of the functional area plus a value proportional to the sum of synchronic and diachronic associations |
| The distribution of weights | Given by the uniform distribution weighted by the geomorphological values of the mask around the cell, and weighted by the functional areas values |

***Table 2***. *Computation of the matrix of weights.*

applied steps 1. and 2. the algorithm makes a step in the prediction of archaeological potential, and after each step the result is taken as the new starting point for the next step. So, the greater *maxit* is, the more the predictions "turn from" the original data.

- *Speed_up (belonging to [0,1])* is the weight expressing the part of the new absolute potential due to the results of the application of steps 1. and 2., and the part due to the absolute potential of the previous step. So, the more *speed_up* approaches *1*, the less the new computation is due to the data from the previous step.

- *Rel (belonging to [0,1])* is the parameter determining how much we take into consideration the potential given by the weight matrix *S* (relations), with respect to how much we take in consideration the potential given by the values of absolute potential. So, the closer *rel* is to *1*, the less the absolute values of the potential are taken in consideration, and the more the matrix of weights *S* (i.e. the relations) is preeminent in determining the archaeological potential.

- *Yield (belonging to [0,1])* is the amount of potential each cell keeps for itself, with respect to the rest, which is distributed on the basis of the weight matrix *S*. So the closer *yield* is to *1*, the more each cell keeps potential for itself.

## 3. The Results

The map of archaeological potential was created for each one of the 7 layers, one for each archaeological period under consideration: Protohistory, Etruscan Period, Roman Period, Late Roman Period, Early Medieval Period, Late Medieval Period, Modern Age, Contemporary Age. The final result was tested by means of the data of 14 new cores. Those data could not be included in a specific archaeological period, because of the nature itself of the method of investigation. Hence the validation of the results provided by the algorithm was performed on the overall archaeological potential. The overall absolute potential was divided into 5 levels, in the same way as the estimated archaeological potential. The comparison was performed computing the difference between the level of the overall estimated archaeological potential and the level of the overall

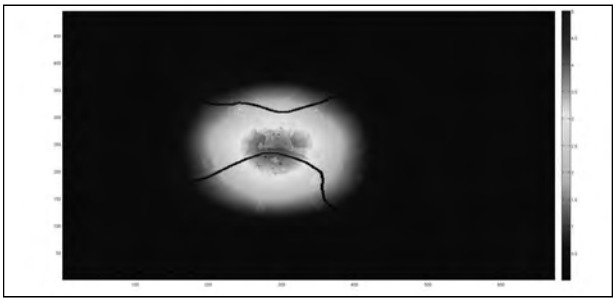

***Figure 4***. *The result of the application of the algorithm to estimate the archaeological potential, for the Late Medieval Period.*

absolute potential. The proportion of exactly estimated potential levels is *9/13 = 69,2%*, while the maximum error is *1*, with an average error of *0.3077* (Dubbini 2013).

The results presented, including the archaeological potential map, are to be considered the first steps towards an automatic, formally definable, and repeatable approach to the computation of archaeological potential. Of course no completely automated procedure would be possible in this and any task involving social and human behaviour, so also in the proposed algorithm the procedure is controlled by the users (archaeologists), who can manage the whole process assigning values to parameters. For these reasons, the map of archaeological potential should be always evaluated in conjunction with the interpreted archaeological data published in MappaGIS (MAPPA 2014b), and with the raw data released as open data in MOD (MAPPA 2014c). In this way, the predictive map of archaeological potential is a useful and powerful tool both for land management and for archaeological research.

## Acknowledgements

MAPPA project has been supported by *Regione Toscana* and is carried out by the Department of Archaeological Sciences, Department of Earth Sciences, and the Department of Mathematics of the University of Pisa, with external collaboration from the Regional Directorate for Cultural and Landscape Heritage of Tuscany, the Superintendency for Archaeological Heritage of Tuscany, the Superintendency for Architectural, Landscape and Ethno-anthropological Heritage for the Provinces of Pisa and Livorno, the Municipality of Pisa, the National Institute of Geophysics and Vulcanology, the National Aerial Photograph Archive, and the Digital Culture Laboratory – CISIAU Interdepartmental Centre of Information Services for the Humanities.

## Bibliography

Anichini, F., M. Bini, F. Fabiani, G. Gattiglia, S. Giacomelli, M. L. Gualandi, M. Pappalardo, and G. Sarti. 2011. "Definition of the parameters of the Archaeological Potential of an urban area." *MapPapers* 1:47–49. Accessed October 26, 2013. doi: 10.4456/MAPPA.2011.04.

Anichini, F., F. Fabiani, G. Gattiglia, and F. Ghizzani Marcìa. 2013. "Questione di pesi: valori, parametri, relazioni per il calcolo del potenziale archeologico." In *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archeologico. Volume II*, edited by F. Anichini, N. Dubbini, F. Fabiani, G. Gattiglia, and M. L. Gualandi, 89–100. Roma: Edizioni Nuova Cultura.

Bini, D., N. Dubbini, and S. Steffè. 2011. "*Mathematical models for the determination of archaeological potential.*" *MapPapers* 1:77–85. Accessed October 26, 2013. doi: 10.4456/MAPPA.2011.08.

Bini, D., N. Dubbini, and S. Steffè. 2012. "*On the two main issues about the application of page rank for the determination of archaeological potential.*" *MapPapers* 2:45–50. Accessed October 26, 2013. doi: 10.4456/MAPPA.2012.04.

Dubbini, N. 2013. "L'algoritmo MAPPA." In *MAPPA. Metodologie Applicate alla Predittività del Potenziale Archeologico. Volume II*, edited by F. Anichini, N. Dubbini, F. Fabiani, G. Gattiglia, and M. L. Gualandi, 101–114. Roma: Edizioni Nuova Cultura.

Dubbini, N., and G. Gattiglia. 2012. "Mathematical Models for the Determination of Archaeological Potential." In *Archaeology in the Digital Era. Volume II. e-Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*, edited by G. Earl, T. Sly, A. Chrysanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska, and D. Wheatley, 710-719. Amsterdam: Amsterdam University Press.

Fabiani, F., and G. Gattiglia. 2012. "The digital archiving structure." In *MAPPA. Methodology Applied to Archaeological Potential Predicitivity. Volume I*, edited by F. Anichini, F. Fabiani, G. Gattiglia, and M. L. Gualandi, 4 –73. Roma: Edizioni Nuova Cultura.

Langville, A. N., and C. D. Meyer. 2006. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton: Princeton University Press.

MAPPA 2014a. "Mappa project." Accessed January 15. http://www.mappaproject.org/.

MAPPA 2014b. "MappaGIS." Accessed January 15. http:// www.mappaproject.org/webgis/.

MAPPA 2014c. "MAPPAopenDATA." Accessed January 15. http:// www.mappaproject.org/mod/.

Shannon, C. E. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27:379–423, 623–656.