

# The Potenza Valley Survey. A GIS analysis of the first field campaign.

G. Verhoeven

Ghent University, Department of Archaeology and Ancient History of Europe, Blandijnberg 2, B-9000 Ghent

## 1. THE POTENZA VALLEY SURVEY

In January 2000, Ghent University (Belgium) initiated a geo-archaeological research project called the 'Potenza Valley Survey' (PVS) in the 'Regione delle Marche' (The Marches). This interdisciplinary project was set up by the Departments of Archaeology and Geography. Under the direction of Prof. Frank Vermeulen, the major aim of the PVS was -and still is, as the project will last until at least 2006- to measure the evolution of social complexity within the valley of the Potenza, one of the Marches' major rivers.

The PVS wants to examine how populous the region was, what the different types of settlements were and how the settlement patterns changed, in time as well as in space. Therefore, the project got the subtitle 'From Acculturation to Social Complexity in Antiquity: A Regional Geo-Archaeological and Historical Approach'. To fulfil the initially outlined goals<sup>1</sup>, three sample zones, systematically spaced around the river, are focused on (Fig. 1).

The Potenza -called *Flosis* in Roman times- flows over its ca. 80 km long course from the Apennines through a wide and fertile Apennine foothill landscape before it ultimately runs into the generally flat Adriatic coastline zone. Already important in Prehistory<sup>2</sup>, the Potenza river valley became one of the most commercial routes of the Central Italian Protohistory and even in later periods the valley remained an important corridor for political, economical and cultural contacts between the Thyrrenian and the Adriatic coast. In Roman times, several cities developed in or near the valley floor (*Potentia*, *Helvia Recina*, *Trea*, *Septempeda* & *Prolaqueum* cf. Fig. 1) and even a southern branch of the Roman Via Flaminia passed through it. In the subsequent period, the area remained of importance as it formed the contact zone between Longobards and Byzantines. These are only a few reasons why a geo-archaeological survey was launched in this particular area.

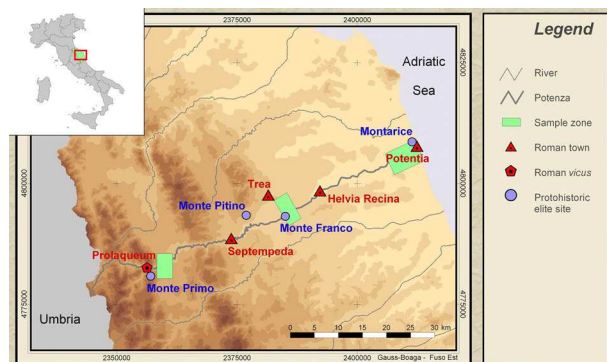


Figure 1. Location of the three sample zones.

The geo-archaeological research consists of four important cornerstones: aerial survey, field survey, geomorphological survey and a historical survey. The aerial survey comprises the collecting of vertical photographs<sup>3</sup> and satellite images. Besides this passive part, the survey also has an active part: taking oblique aerial photographs from a low flying aircraft. Till now, this has been done in a very thorough way: on the one hand, different flights were made in different seasons; on the other hand, pictures were taken of the whole valley with a specific interest in the three sample zones.

Systematic prospection of the ploughed fields, always carried out in the month of September, is the method used in the field survey. In search for archaeological evidence, intensive linewalking was chosen as prospection method. Therefore, an interval of 10 to 15 meter between different walkers is aimed at. Scholars of the Department of Geography conduct the geomorphological part of the project by making as many field observations as possible. Additionally, some augering is also performed to reconstruct the evolution of the coastline and study the problem of alluviation and colluviation.<sup>4</sup> Finally, an intensive historical survey to study known archaeological sites of all periods into detail is implemented as well. Moreover, some research of toponymical and historical written information is ongoing.

Since the PVS started more than three years ago, a substantial amount of information is gathered.<sup>5</sup> Managing such quantities of -almost all geographically linked- data has become much easier with the introduction of GIS. Linking spatial data to non-spatial (attribute) data, GIS has become a commonly used management and analysis tool in a lot of disciplines, also in archaeology. From the

beginning of the project, it was the aim of director Frank Vermeulen to incorporate all gathered survey data into a GIS-context.

## 2. DISSERTATION

At the time of writing this article, one must be aware that the project is already going on for almost four years. During this period, the GIS was set up, spatial data collected and the first analyses realised. A part of this work was performed by students. By writing a dissertation on a PVS-subject, they were offered the chance to work in an interdisciplinary project and apply their theoretical knowledge in a practical way. One of these dissertations was made by the author.<sup>6</sup> Besides a theoretical approach of the terms “field survey” and “Geographical Information System”, an analysis of the results of the first field campaign completed this thesis.<sup>7</sup> This paper will focus on the latter part.

## 3. METADATA AND QUALITY CONTROL

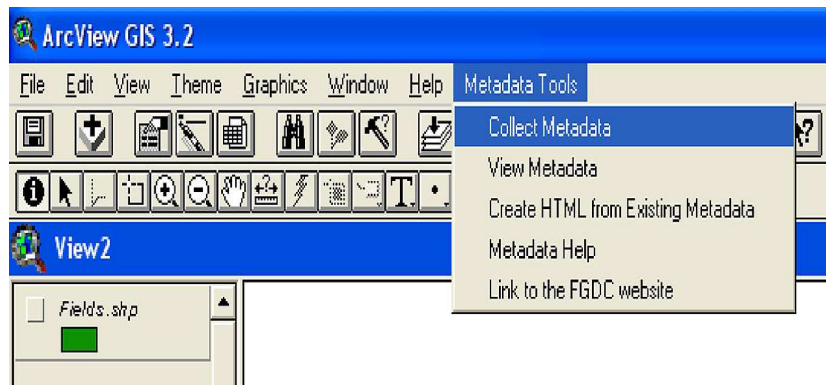
Before any visualization and analyzing could occur, it was important to be(come) aware of the data quality. After all, it is important to understand that it is possible to have ‘error free’ attribute data (in terms of the method employed), but impossible to have ‘error free’ spatial data, because the multi-dimensional reality will always differ from its digital representation. Therefore, the entering of spatial information can rightly be seen as the Achilles’ tendon of almost every GIS.<sup>8</sup>

Quality affecting factors that can be mentioned are the accuracy and precision of the source data, the interpolation methods employed, the scale of the data, the georeferencing system used, the data collection technique and the sampling strategy that are applied, the process of scanning and digitizing, etc. In this respect, a lot of attention was paid to metadata and -a limited amount of- quality control.

### 3.1. Metadata

Metadata can be seen as ‘data about the data’ as it describes the content, quality, condition and other characteristics of data. It provides us with information about the who, what, when, where, why and how of a data set.

In a GIS, metadata are often neglected. Unjustly so, because archaeological research is often undertaken by researchers using digital data created by others. The former have limited understanding about the limitations or purpose of that data. Therefore, documenting data can be helpful to others and yourself (even the person who digitized the data, may forget items like the precision and the accuracy of the data, the area it covers, the date of creation, etc.) In consequence, proper metadata are critical in preserving the usefulness of data over time and it is crucial to supply all kinds of GIS-data with this information. Different metadata standards exist. In the PVS, the geospatial information is documented using the Federal Geographic Data Committee’s (FGDC) Content Standard for Digital Geospatial Metadata, which the Federal Geographic Data Committee approved in June 1998.<sup>9</sup> To build FGDC-compliant metadata, an easy-to-use application called the ArcView Metadata Collector was achieved (Fig. 2).<sup>10</sup>



**Figure 2. The drop-down menu which appears in ArcView 3.x after installing the Metadata Collector.**

### 3.2. Quality control

Besides metadata, quality control has to be considered another important topic. Using spatial data

in a GIS-operation implies that error (defined as the discrepancy between a given -digital- value and its value in reality) in the input will propagate to the output of the operation. In this respect, the resulting output is a function of the input values and inaccurate input values will automatically affect the computed results.<sup>11</sup> Therefore, it is significant to test the accuracy of some digital maps. One example is the *Digital Elevation Model* or DEM. In the majority of GIS-data sets, the DEM is the most important layer. Except information about the height, it is also used to generate slope and aspect and it serves as well as a basis for *cost path analysis*, *distance calculations* etc. As a consequence, it is fundamentally to study the topographical features of archaeological sites. One always needs an interpolation algorithm to construct a usable DEM. Some digitized high points or contour lines form the basis of the DEM, but they do not form a continuous surface with height values yet. To retrieve this, an interpolation method is run and will attribute the places without data a figure that represents the height of that specific place. However, different interpolation methods produce different terrain values. We must surely accept that the final DEM will be an approximation of a continuous phenomenon, but how closely this approximation reflects reality can be calculated.<sup>12</sup> In the thesis, the accuracy of the DEM applied for the first sample zone was calculated. This control occurred on the basis of two tests. Initially, the digitized contour lines were checked visually by generating a TIN (*Triangulated Irregular Network*). A closer look at the characteristic points and lines as well as the hillshaded relief revealed some profound errors, owed to the manual labelling. Afterwards, a quantitative test was run. Eleven different interpolation methods were executed.<sup>13</sup> The predicted elevations for a given DEM were compared to some reference points (none of them was located on a contour line) and the discrepancy between both was calculated. With these figures, a global RMSE (*Root Mean Square Error*, which is mathematically the same as standard deviation) could be calculated for a given DEM. This procedure was repeated for each interpolation. Subsequently, different errors could be compared and the lowest RMSE revealed the preferable DEM.

#### 4. ANALYSES

The first part of the analyses encompassed the set up of a questionnaire. Brainstorming with different people led to a questionnaire, which in its initial phase -the questions are still adapted and updated during the project- could be divided into two parts: analysis of data linked with the cadastral maps and analysis of environmental characteristics. The former part covered items as the calculation of the precise prospected area (3,2 km<sup>2</sup>), determination of the proportional ratio of the visibility classes and the computation of the proportional ratio of the fields with erosion (29,9%). Furthermore, a *chi-squared test of goodness-of-fit* revealed for a *significance level*  $\alpha = 0,05$  that sites are equally distributed across fields with and without erosion. A determination of the off-site proportions (0 = 5,7%; 1 = 56,5%; 2 = 19,2%; 3 = 18,6%)<sup>14</sup> as well as a calculation and visualization of the total and periodical site density (total = 17,19 sites/km<sup>2</sup>; pre- and protohistoric = 3,13 sites/km<sup>2</sup>; Roman = 13,75 sites/km<sup>2</sup>; medieval and post-medieval = 1,12 sites/km<sup>2</sup>) completed this first part. One can notice the high density of Roman sites. However, a further subdivision per period is needed, before any conclusions can be drawn from these figures.

The second part dealt with the relation between the sites and the environmental characteristics as height, slope, distance to the nearest water source, etc. In an archaeological GIS, analysis of such environmental characteristics can occur in two different ways: the *contextual analysis*, which brings together the context of sites and finds and analyzes them without any proof (e.g. *site catchment analysis*) and the *locational analysis*, identical to the contextual except the attempt to proof the hypothesis.<sup>15</sup> It was the latter kind of research that was performed. All the analyses are based on a study by Hodder and Orton.<sup>16</sup> They examined the distribution of 173 Iron Age coin finds in relation to Roman road locations in central and southern England. Using manual methods of pre-GIS days, they performed a statistical analysis which revealed a significant association between the coin distribution and the Roman road network. With GIS, such analyses can be executed a lot faster and with results that are even more precise and accurate.

The distance to the nearest waterway was the first examined background feature. To illustrate, the first step is to create a distance grid, where every cell holds a figure that represents the distance of

that cell to the nearest waterway (Fig. 3). This grid serves as a background distribution. The distance of the sites to the nearest waterway can be extracted from this background distribution. Subsequently, the following question may be asked: is this a normal sample of our background or is the location of the sample cells (the sites) unusual in relation to our general background values (the study area)? Therefore, both the background and the sample values were plotted cumulatively and the *Kolmogorov-Smirnov one-sample test* computed. This test statistically compares the sample distribution against the background distribution and tells the researcher if the sample of site locations differs significantly from the total background. If it does, one can accept that the sites were placed in relation to the examined feature. If the difference between the two cumulative curves doesn't reach a critical value, sites were placed randomly considering the background feature and the significance level.

In this way, the relation of the sites (per main period) and isolated finds to the following environmental characteristics was determined: elevation, slope and aspect; distance to the waterways, to the Potenza, the springs and the total hydrographical network; distance to the nearest road, to the nearest flint; view on the Potenza (Table 1). Some relations did occur (e.g. the isolated stone age artefacts were obviously linked to the springs and the Roman sites to the contemporary rural roads), but generally spoken, no specific connection between the environmental characteristics and the sites could be deduced for the moment, partly due to -as underlined before- the fact that the periods were too broad delineated for detailed analyses.

## 5. COMMENTS AND CONCLUSION

The results presented here are a first, archaeological-analytical examination of the archaeological results gathered in the first field campaign of the Potenza Valley Survey. Notwithstanding the fact that the performed one-sample statistical methods were suitable, the results have to be seen as preliminary. When ongoing rectification and interpretation of the archaeological oblique airphotos will be finished, a digital soilmap is implemented in the GIS, the geomorphologists have clarified the changing courses of the rivers as well as the varying locations of the springs and, finally, a chronological refinement of the different sites is achieved, it will be worthwhile to redo the analyses to get more precise -and hopefully more accurate- results. Furthermore, it is possible that more and/or other statistical procedures will have to be followed. However, a final solid reconstruction of the (archaeological) landscape will be an interdisciplinary combination of opinions and interpretations, made by different scholars, thus avoiding (too) one-sided visions and conclusions.

## ACKNOWLEDGEMENTS

The PVS is an interdisciplinary geo-archaeological survey. The director of the project is Prof. Frank Vermeulen. Prof. Morgan De Dapper leads the geomorphological research.

The survey is carried out within the Regione delle Marche and the PVS-team is very grateful to the So-printendenza Archeologica delle Marche and especially to Soprintendente Dott. Giuliano de Marinis and Dottoressa Edwige Percossi -Ispettrice of the area- for their support and expertise.

## REFERENCES

Federal Geographic Data Committee 1998. FGDC-STD-001-1998. Content Standard for Digital Geospatial Metadata (revised June 1998). Washington, D.C.

Fisher, P. 1999. Geographical Information Systems: Today and Tomorrow?. In: Gillings, M., Mattingly, D. & van Dalen, J. (eds.). *Geographical Information Systems and Landscape Archaeology*. Oxford, 5-11. (= The Archaeology of Mediterranean Landscapes 3)

Gillings, M. & Wise, A. (eds.) 1999. GIS Guide to Good Practice. Oxford. (= Archaeology Data Service)

Hageman, J. & Bennett, D. 2000. Construction of Digital Elevation Models for Archaeological Applications. In: Wescott, K. & Brandon, R. (eds.). *Practical Applications of GIS for Archaeologists. A Predictive Modeling Toolkit*. London, 113-127.

Heuvelink, G. 1993. Error Propagation in Quantitative Spatial Modelling. Applications in Geographical Information Systems. Utrecht. (= Nederlandse Geografische studies 163)

Hodder, I. & Orton, C. 1976. *Spatial Analysis in Archaeology*. Cambridge. (= New Studies in Archaeology)

Miller, P. 1999. The Importance of Metadata to Archaeology: One View from within the Archaeology Data Service. In: Dingwall, L., Exon, S., Gaffney, V., Laffin, S. & van Leusen, M. (eds.). *Archaeology in the Age of the Internet. Computer Applications and Quantitative Methods in Archaeology. CAA 97. Proceedings of the 25th Anniversary CAA Conference Held at the University of Birmingham, UK 10-13 April 1997*. Oxford, 133-136. (= BAR International Series 750)

Percossi Serenelli, E. 1981. Le vie di penetrazione commerciale nel Piceno in età protostorica. Nota preliminare. PICUS. Studi e ricerche sulle Marche nell'antichità 1: 135-144.

Verhoeven, G. 2002. The Potenza Valley Survey. Een GIS-analyse van de eerste veldcampagne en een duiding van de begrippen "Geografisch InformatieSysteem" en "field survey". Gent. (Unpublished Dissertation UGent)

Vermeulen, F. 2002a. The Potenza Valley Survey (Marche). In: Attema, P. *et al.* *New Developments in Italian Landscape Archaeology*. Oxford, 104-106. (= BAR International Series 1091)

Vermeulen, F. 2002b. Uit de lucht gegrepen: de bijdrage van archeologische luchtfotografie in het "Potenza Survey" project. *Tijdschrift voor Mediterrane Archeologie* 26: 33-42.

Vermeulen, F. 2003. The Potenza Valley Survey: Preliminary Report on Field Campaign 2002. *Babesch* 78. (forthcoming)

Vermeulen, F. & Boullart, C. 2001. The Potenza Valley Survey: Preliminary Report on Field Campaign 2000. *Babesch* 76: 1-18.

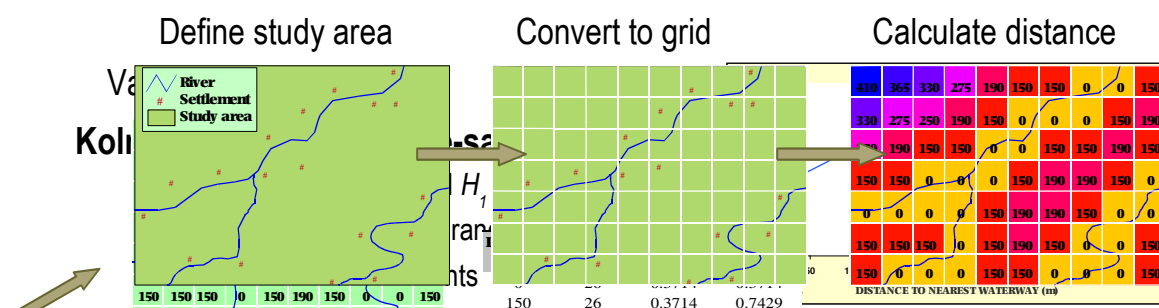
Vermeulen, F., De Dapper, M., Boullart, C., De Vliegheer, B-M. & Goethals, T. 2002. Geoarchaeological Approaches in the Valley of the Potenza (Marches, Central-Italy). *Proceedings of the congress of Paris 2002*. (forthcoming)

Vermeulen, F., Monsieur, P. & Boullart, C. 2002. The Potenza Valley Survey: Preliminary Report on Field Campaign 2001. *Babesch* 77: 49-71.

Voorrips, A. 1998. Electronic Information Systems in Archaeology. Some Notes and Comments. *Archeologia e Calcolatori* 9: 251-267.

<http://www.csc.noaa.gov/metadata/text/download.html>

S.n. NOAA Coastal Services Center. *ArcView Metadata Collector v2.0 Extension*. 07/01/2003



Select significance level:  $\alpha = 0,05$   
 Calculate critical value  $d = 1,645$

275	2	0.0286	0.9429
330	2	0.0286	0.9714
365	1	0.0143	0.9857
410	1	0.0143	1.0000
Total		70	1

Create cumulative curves

Values of sample

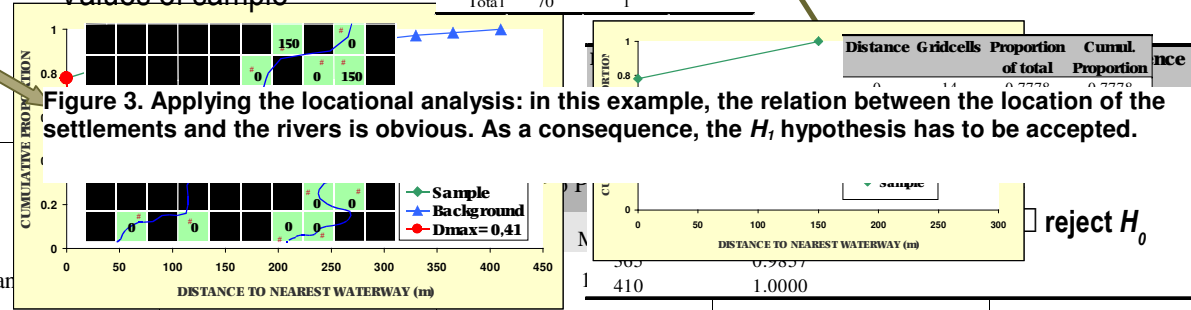


Figure 3. Applying the locational analysis: in this example, the relation between the location of the settlements and the rivers is obvious. As a consequence, the  $H_0$  hypothesis has to be accepted.

Pre- and protohistory	10	529	194	No	130	2797	1029	No	32	959	425	No	0	385	159	No
Roman	156	250	202	No	710	1342	895	No	461	581	531	No	149	250	198	No
(Post-)medieval	22	297	180	No	211	1168	490	Yes	82	875	257	Yes	28	278	116	No
Isolated finds																

	Elevation (m)				Slope (°)				Aspect (°N)					
	Min	Max	Mean	Relation	Min	Max	Mean	Relation	Flat	0 - 90	90 - 180	180 - 270	270 - 360	Relation
Pre- and protohistory	344	470	394	No	4	12	8,5	No	0%	10%	30%	0%	60%	No
Roman	330	474	392	No	0	17	7,2	No	13,6%	31,8%	29,6%	6,8%	18,2%	No
(Post-)medieval	397	410	403	No	0	16	8,4	No	25%	25%	25%	25%	0%	No
Isolated finds	340	400	378	No	0	14,5	6,4	No	28,6%	14,3%	21,4%	21,4%	14,3%	No

	Distance to nearest road (m)				Distance to nearest flint (m)				View on the Potenza			No view on the Potenza		
	Min	Max	Mean	Relation	Min	Max	Mean	Relation	Site count	%	Relation	Site count	%	Relation
Pre- and protohistory	10	240	90	No	32	233	141	No	6	60	No	4	40	No
Roman	0	331	102	Yes					21	47,7	No	23	52,3	No
(Post-)medieval	10	322	137	No					2	50	No	2	50	No
Isolated finds					30	305	221	Yes						

Table 1



- <sup>1</sup> For a detailed overview of the main goals, the reader is kindly asked to consult 'Vermeulen 2002a'.
- <sup>2</sup> Percossi Serenelli 1981
- <sup>3</sup> Mostly from the R(oyal) A(ir) F(orce).
- <sup>4</sup> An overview of this work as well as the first results can be read in 'Vermeulen, De Dapper, Boullart, De Vliegheer & Goethals 2002'.
- <sup>5</sup> For more detailed information about the project and the work performed till now in the three sample zones, the following articles are recommended: 'Vermeulen 2002b', 'Vermeulen 2003', 'Vermeulen & Boullart 2001' & 'Vermeulen, Monsieur & Boullart 2002'.
- <sup>6</sup> Verhoeven 2002
- <sup>7</sup> The configuration of the PC used in these analyses: *hardware*: 2.0 Ghz Intel processor; 256 MB RAM; 75 GB hard disk; 1,44 MB floppy drive, Sony CD-RW 24/10/40; *software*: ArcView 3.2 as GIS-software, Matlab 6.0.088 to perform statistical analyses and Excel to create graphs. All programs were working on a Microsoft Windows XP (home edition) operating system.
- <sup>8</sup> Voorrips 1998, 255
- <sup>9</sup> Federal Geographic Data Committee 1998
- <sup>10</sup> <http://www.csc.noaa.gov/metadata/text/download.html>
- <sup>11</sup> Heuvelink 1993, 23-25
- <sup>12</sup> Hageman & Bennett 2000, 114-115
- <sup>13</sup> Exponential, Circular, Spherical, Gaussian and Linear Ordinary Kriging, Universal Kriging –with linear and quadratic drift-; Spline, Inverse Distance Weighting; Trend and TIN. All of them were run with different values for the variable parameters. As a result, 57 grids were calculated.
- <sup>14</sup> Off-site 0: 0 artefacts /50m; 1: <=5 artefacts/50m; 2: <15 artefacts/50m and off-site 3: >=15 artefacts/50m
- <sup>15</sup> Fisher 1999, 8
- <sup>16</sup> Hodder & Orton 1976, 226-229