Managing Uncertainty in Archaeological GIS Applications

Franco Niccolucci

University of Florence, Italy

Introduction

GIS have been used in archaeological applications since more than a decade now, and there has been a deep discussion confronting the GIS paradigm and way of use with the fundaments of archaeological methodology. Our goal in the present paper is neither to contribute to such debate nor to summarize it, but just to present some features having a potential impact on archaeological investigation.

Archaeological GIS applications are, sometimes, just digital cartography, that is a handy way of storing and neatly reproducing archaeological maps. There is nothing incorrect in that, but such applications do not exploit the potential of GIS. In other, and more interesting, cases, GIS is a key factor to reach novel archaeological conclusions that could not be achieved without using the GIS. One such example is viewshed analysis, where a 3D model of the relevant area is used to analyze the line-of-sight from particular points, as sites, hillforts or other locations, and hence draw conclusions on settlement patterns. Other applications involve prediction of yet undiscovered sites basing on spatial analysis of settlement patterns.

In general, GIS may be used to synthesize spatial factors and produce thematic representations and to filter archaeological data according to specific parameters, representing the results on maps in order to us them as an aid for archaeological interpretation. The reliability of such topographic synthesis is a crucial factor for the correctness of the interpretation, but it is often acritically given for granted, perhaps because it is produced by a computer, which by definition makes no mistakes.

There are, on the contrary, two main sources of error: one derives from the processing and the second from the nature of archaeological data. The consequences of both are often underestimated: sometimes ignoring them is harmless, in other cases it is potentially dangerous for the correctness of the archaeological results. In no case, however, they should stay uncontrolled.

Processing errors

Errors deriving from GIS processing are, in general, independent of the specific archaeological problem but may have archaeological consequences. Geographers have studied since long the different kinds of error that may affect a GIS, distinguishing among obvious sources of errors (as lack of precision or detail in data, human mistakes, etc.), errors deriving from natural or measurement variations (as lack in positional accuracy, faulty observations and measurement, instrument errors, e.g. deriving from a GPS, or seasonal changes, e.g. in water courses), and processing errors, i.e. introduced by rounding off and by propagation and cascading effect caused by the computational algorithms used to process data.

Accuracy in geographical data for archaeological applications has been analysed in very few cases, for instance in (D'Andrea et al 2002) and (Topouzi et al 2002). Both these cases are related to GPS use and report the precision obtained by the instrument used to locate archaeological features in space and to geo-reference them in a GIS. Most papers concerning GIS applications quote the base layer precision, but in many others, unfortunately, the base system, describing in digital terms the morphology of the terrain, is not reported and the reader cannot figure out the accuracy of the underlying geographical layer. For instance, it may make the difference if the digital topographic base was obtained by digitization of a 1:25 000 map or from a 1:100 000 map, or directly from a digital map with known resolution provided by some national agency as Ordnance Survey in UK or IGM in Italy. It is hoped that in such cases authors understand that mixing different sources reduces the reliability of the result to the poorest of them, even if they do not deal with the issue in the paper. This is not always the case in archaeological GIS literature. In a recent paper concerning "geo-archaeological investigations" sources are described as follows: "remote sense imagery, Russian cartography, geo-archaeological topographic surveys, unsupervised classification, palaeo-environmental interpretations etc." ("etc." is in the original). Elsewhere in the same paper, survey data are described as "collected with a PDA, a small (Trimble) antenna and the software Arcpad". According to this description, the reliability of such a mixed data salad is unpredictable, and the archaeological conclusions are absolutely unreliable. So it is unclear why this project was reported in such an undetermined way - hoping at least that it has not been carried on with the same indetermination.

One critical point in GIS applications is the use of a DEM (Digital Elevation Model), that is a virtual continuous representation of the terrain based on the interpolation of known elevations. Thus the quality of the final result is determined by the quality of the base points and by the algorithm used for the interpolation. Awareness of this dependence is not widespread. In the CAA2003 proceedings volume, some authors correctly point out that this issue is critical and describe in detail (as in Farinetti and Sbonias 2004) the base layers and the algorithm they use; a paper is dedicated to discuss general features and problems of the different DEM algorithms (Beex 2004). In another paper, before discussing alternate solutions for their DEM, authors state that "it is well known that even with a slightly different quality of DEMs we can obtain completely different results of path simulation" (Podobnikar, Tecco Hvalak and Dular 2004). Their statement is correct, but not so well known, because several of the following papers use freely DEMs, often in a substantial way for the paper goals, without letting the reader know anything about the base system precision, the interpolation method used and the accuracy required to draw the conclusions presented in the papers. Applications as viewshed analysis or cost-surface analysis should, on the contrary, always require the description of the model features, as the base layers and the interpolation algorithm used, and some verification that the accuracy of the resulting model is sufficient for the goal.

Unfortunately, DEM verification is not straightforward. The above quoted paper (Podobnikar, Tecco Hvalak and Dular 2004) proposes a sort of supervised procedure that mixing a-priori knowledge with GIS calculation, aims at progressively and iteratively arrive to "eliminate gross errors of DEM" in determining ancient paths (cost-surface analysis). More common is the use of a Monte Carlo simulation to test the quality of the DEM (Nackaerts, Govers and Loots 1999). In loose terms, this consists of a series of tests on random sets of points, reconstructing the sample statistics from the test results. Other authors have more pragmatically suggested to use the GIS as a guide to interpretation and verify the results by direct inspection in situ.

In conclusion, every GIS model has a given data accuracy that is better to report in all cases and a confidence interval that may be difficult to determine. It is plausible that most GIS applications fall within the confidence interval, but when the model accuracy is critical for the archaeological interpretation it is necessary to double-check the model with known mathematical and statistical methods as the above quoted ones, or, at least, to verify the model accuracy in a neighbourhood of the relevant points, to assure that the conclusion is not affected by model imprecision.

Inherent uncertainty in archaeological concepts

Archaeological (spatial) concepts are often defined in an imprecise way. This consideration applies even to the most basic spatial archaeological concept, the "site". Even if in most cases it is possible to state that a point is internal to a particular site, or that it is not, the site border cannot be represented by a sharp line. We often do so as a simplification, but in some cases it may be an over-simplification. Having imprecisely defined borders is a rather common feature for geographical objects (see Burrough and Frank 1996 and, for a application to history, Benvenuti and Niccolucci 1996), unless they are determined by some a-priori definition, e.g. an administrative or political one. Such concepts are therefore better represented as fuzzy concepts. Let us remind that a *fuzzy set* is defined as a generalization of ordinary sets. Membership in a fuzzy set is not a yes-no condition, rather it has a continuum of possibilities varying from 0 (not belonging) to 1 (belonging). The function associating to every element of a fuzzy set the corresponding membership index is called the *degree of membership* or the *membership function* or, shortly, the membership.

As opposite to fuzzy sets, ordinary sets are called "crisp". A crisp set is defined as a set whose elements have a membership assuming the values 0 or 1 only.

So fuzzy 2-dimensional sets, also called *fuzzy regions*, are in fact 3-dimensional, because in every point they add to the two point coordinates a third number, the membership in the set, expressing the possibility that the point belongs to the set. Such sets can be visualized in two dimensions with a varying intensity of a colour, full colour being associated to 1 and blank to 0.

Also the attributes of archaeological objects may be fuzzy, because their value is not known with precision and is assigned with some uncertainty expressed with a number. A typical example of this are age and gender of the deceased buried in a cemetery. Finally, conditions may be fuzzy because they refer to linguistic concepts as spatial concepts ("near to") or object attributes ("young person", "prestige goods").

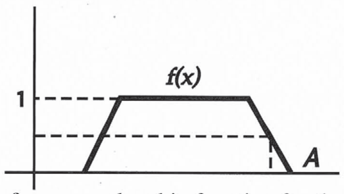
Thus, many of the usual concepts have a fuzzy counterpart that is usually more effective in describing archaeological concepts. The fuzzy extension of an archaeological GIS develops therefore in two directions: fuzzy extension of the object database, allowing fuzzy attributes for the entities in the GIS alphanumeric database, and fuzzy extension of the geographical objects, allowing fuzzy definitions and fuzzy relations.

Let us summarise briefly the definitions of these fuzzy extensions for the underlying database.

A *fuzzy label* is an attribute that can assume a set of values, the attribute domain, each associated to a number between 0 and 1. In the crisp label case, the attribute can assume only one of the possible values in its domain, while in the fuzzy extension each of the values is possible, and has a number attached to it, expressing the possibility of the corresponding value. A common example is the gender of a person buried in a tomb. In the crisp perspective, the deceased has to be assigned to the male or female gender, or perhaps remain indeterminate, so two values are possible: Male and Female (and unassigned, or NULL), but one and only one may be assigned. In the fuzzy extension, all the values are considered as possible and a numeric index, its membership as previously defined, is attached to each of them to express its possibility, i.e. its feasibility or likeliness. Such a fuzzy gender is represented as follows:

$\{(M; x_M), (F; x_F)\}$

that is by a set of couples where the first element is the attribute value and the second



A fuzzy membership function for the set *A*

is its membership. For example:

{(M; 0.7), (F; 0.5)}.

A *fuzzy value* in the numeric domain N is a function assigning to every value of N the corresponding possibility index a number in [0, 1]. If N is finite this is a particular case of fuzzy label, in which the attribute values are numeric. A typical example is the age of a deceased: in the crisp case, only one numeric value can be assigned to the age, while in the fuzzy model several values are possible, those considered impossible having 0 as possibility index. Another useful concept is the *fuzzy constant*. This is used to express with precision, in fuzzy terms, common linguistic concepts that lack a precise definition or have a context-dependent definition. A good example is the term "young" as opposed to adult and/or elderly. Not only it includes a range of ages, with fuzzy limits (unless one uses a legal definition, e.g. the full legal age, as a limit), but it is also related to a specific context, the age range of maturity being different nowadays from the past. So a fuzzy constant is a pre-determined value for a fuzzy label or a fuzzy value. For instance, one could define "young" in a modern context as having membership 1 for ages 14 to 17, 0.5 for 12 and 13 and 0 for other ages.

The presence of fuzzy concepts requires also extensions of relations called *fuzzy operators*. It indeed makes little sense asking if two fuzzy labels are equal, for instance if the persons buried in two tombs are of the same gender, because the gender is fuzzy for each of them. In fact also equality, the simplest relation, needs an extension called *fuzzy equality*, sometimes denoted with ~. To be eligible for equality, of course, the two attributes need to be of the same kind, that is have the same domain, otherwise they cannot be compared. This is true also for crisp attributes.

The fuzzy equality of two fuzzy labels takes into account every possible label value and compares the corresponding membership of each of the two operands, i.e. the two objects being compared on that attribute. It retains the minimum of these two values, this being the "worst" possible case. Then it considers all the possible label values and determines the highest joint possibility index, which is the value of the fuzzy equality. For instance, to compare

{(M; 0.7), (F; 0.5)} ~ {(M; 0.4), (F; 0.9)}

one obtains that the minimum for M is 0.4 and for F is 0.5, so the value of fuzzy equality is the largest of the two, i.e. 0.5. This definition works well and give the result one might expect in special cases. Firstly, it gives 1 for equal crisp objects, and 0 for different ones:

{(M; 1), (F; 0)} ~ {(M; 1), (F; 0)} = max (1 , 0) = 1

{(M; 1), (F; 0)} ~ {(M; 0), (F; 1)} = max (0 , 0) = 0.

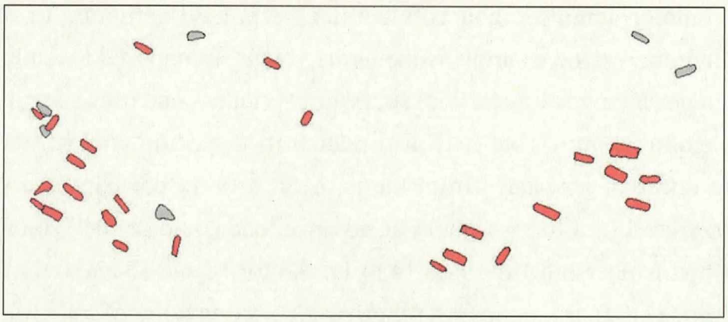
Second, it works well when comparing a crisp object with a fuzzy one

{(M; 0.7), (F; 0.5)} ~ {(M; 1), (F; 0)} = max(0.7 , 0) = 0.7

Further fuzzy operations include counting, for which we do not add "1" for every item, but its membership. Similarly, averaging is a weighted average with membership used as weights, e.g. to calculate the deceased average age of a set of tombs.

With such a definition, it is possible to select objects fulfilling specific fuzzy conditions and map them. For instance, the following figure represents tombs having pre-determined levels of equality to a constant and has been obtained using the fuzzy GIS of a cemetery dating to the 9th - 3rd Century B.C.E. in Southern Italy (Niccolucci, D'Andrea and Crescioli 2001).

Fuzzy spatial operations generalize the classical GIS model in which objects have a



Map representing tombs resulting from a fuzzy selection with the criterion "fuzzy equal to 40 years old". Red (darker) ones have a fuzzy equality greater then 0.5, grey (lighter) ones less than 0.5 and greater than 0. GIS model of the Etruscan cemetery at Pontecagnano (Italy), described in Niccolucci, D'Andrea and Crescioli 2001.

location and various attributes. In the fuzzy GIS model, objects have a fuzzy location, defined as a fuzzy 2-dimensional or 3-dimensional set, and fuzzy attributes as we have just described. Operations may be classified as *local*, i.e. determining new fuzzy values for objects according to their location (e.g. overlay operations), *focal*, i.e. determining new fuzzy values for objects basing on their neighbourhood (e.g. distance), and *zonal*, i.e. operations computing fuzzy values according to a zone that contains them (e.g. select operations).

As an example of local operations, let us consider locations having "good" insulation and "moderate" slope. The two concepts and the associated regions are both fuzzy so it is necessary to define a fuzzy intersection or overlay, which is itself a new fuzzy region whose membership function is defined as the minimum of the two membership functions of the originating sets: if x is a location and m_A denotes the membership function

$m_{A \cap B} = \min (m_A, m_B)$

The above concept may be useful, for instance, to test locations susceptible of being archaeological site when it is assumed that the above conditions are good indicators for the presence of a site. A more complicate case arises when one wants to determine if an actual archaeological site is placed in such an overlay region: this can be dealt with by considering the fuzzy intersection of the fuzzy location where the archaeological site is placed with the above fuzzy overlay region and reasoning on it.

Distance is an example of focal operation. We will not go into the details of the definition of the distance of a point from a fuzzy region, but just sketch out the procedure. Let us assume, for the sake of simplicity, that a fuzzy region A is based on three crisp regions A_1, A_2 and A_3, with a very simple membership function, equalling m_1 on A_1, m_2 on A_2, and m_3 on A_3, m_1, m_2 and m_3 being three numbers in the range [0, 1].

$m_A = \{(A_1; m_1), (A_2; m_2), (A_3; m_3)\}$.

This is in fact no restriction, because if the membership function assumes discrete values, sets as A_1, A_2, $A_3 \cdots$ may be always defined as "level" sets. In the general case, a "slicing" process allows to generalize the procedure to continuous membership functions. Anyway, it is possible in the example to evaluate the distances d_1, d_2 and d_3 of a fixed (crisp) point from each of the regions A_1, A_2 and A_3 in the traditional, non-fuzzy way. Then the *fuzzy distance* of the given point from the fuzzy set A is defined as the fuzzy value d assuming the values d_1, d_2 and d_3 with membership m_1, m_2 and m_3:

$$d = \{(d_1; m_1), (d_2; m_2), (d_3; m_3)\}.$$

In other words, the fuzzy distance has a membership function very similar to the fuzzy region, with values corresponding to the non-fuzzy distance from the level sets of the given region. Things become a bit more complicate when the point representing the location is a fuzzy location, but we will not go into further details.

A possible application of the above can be the evaluation of areas "near" to an archaeological site. This case requires the definition of "near" as a fuzzy constant and then the comparison of the distances with this "near" concept. If the focus is not fuzzy, as for instance "locations near to a river" or "near to a spring" the computation is simpler and requires only comparing a traditional distance with the fuzzy constant "near". To consider the case of "archaeological sites near a spring" it is sufficient to use the spring location as the crisp location and the site as the fuzzy region and apply the previous procedure.

Finally, zonal operations include the creation of thematic maps and examples have already been described above when dealing with select operations.

Apart from the already quoted model of the Etruscan cemetery in Southern Italy (Niccolucci D'Andrea and Crescioli 2001), a predictive model based on fuzzy logic has been presented at CAA2002 (Hatzinikolau et al 2003). In this model the authors suggest fuzzy definitions of some linguistic criteria and verify a set of rules to predict settlement. Due to the scale used and the smallness of the relevant sites, archaeological sites are assimilated to (crisp) points and tested against a set of rules based on the above criteria: e.g. "if the distance of a location from a spring is short and distance from agricultural land is short and slope is smooth then the membership of the location in the set of candidate sites is 0.6", what they express by saying "then the location is a site with 60% certainty". They present in the paper some 2-D maps of these fuzzy regions - or better the union of all the fuzzy regions corresponding to different criteria - and place actual archaeological sites on them, verifying the goodness of fit of the predictive model.

Introducing fuzziness into GIS has been the object of computer science research since several years (see for instance Stefanakis, Vazirgiannis and Sellis 1996, Galindo, Medina, Pons and Cubero 1998 and Schneider 1999) but, unfortunately, in practice it is not easy to develop real applications because commercial GIS software has little provision, if any, for fuzzy concepts and fuzzy processing tools are limited or absent. This may be no problem, because in most cases the crisp approximation to fuzzy concepts works well, in the average uncertainty balances and the archaeological GIS may avoid the complication of fuzziness.

When fuzzy logic is required, Open Source (OS) systems have a definite advantage because they allow insertion of user-defined types (for the database) and functions (for the GIS). One such example is the GRASS GIS that can use the Postgres RDBMS as underlying database. Users of these OS systems have created some fuzzy functions for both, and these may be further extended programming the necessary modules and recompiling the sources of the packages. It is not an end-user task, but it may be accomplished rather easily with some programming skills, and archaeologists will surely find a colleague at the computing lab willing to do the job for them. Fuzzy function extensions for GRASS have been made available by François Delclaux at ftp://ftp.mpl.ird.fr/pub/delclaux/rfuzzy/ as raster functions, while fuzzy operators and labels have been implemented for Postgres by the author of the present paper and are available upon request.

Another issue that usually discourages newcomers to fuzzy logic is the question: "how is the membership function determined?" The concept here is expressing with a numeric value the degree of confidence evaluated by experts, first of all the researcher carrying on the investigation and making explicit by quantification vague concepts that are nonetheless commonplace, as the ones previously quoted like "near", "young", and so on. Therefore the evaluation is subjective, but is made explicit and other researchers may disagree and argue. The evaluation may however be based on other factors to give a sound, less questionable basis. For instance, gender and age assignments in the previously quoted cemetery example were based on the results of physical anthropological investigations undertaken on the skeletons, producing an anthropological index according to the Acsadi and Nemeskeri method (Acsádi and Nemeskéri 1970), which can be the starting point for assigning the membership functions to each fuzzy attribute.

In the same case study, dating based on grave goods was assigned a trapezoidal membership function defined by experts with "tails" of 25% of the time interval. The trapezoidal function is the most common for fuzzy values, because it is easy to compute and corresponds to the concept of certainty of "yes" for inside values, certainty of "no" for outside values and uncertainty for borders, smoothly linking "yes" and "no". It is also easy to store in a database because it requires only four parameters when the fuzzy value is clearly unimodal, as it is the most frequent case. In the above quoted paper (Niccolucci D'Andrea and Crescioli 2001) explanation and justification are given in detail.

Conclusions

The above considerations show that beyond the debate on quantitative methods there is still some work to do to make quantitative analysis using GIS reliable. Computational difficulty will not be a problem when procedures will be created for routine processing.While the accuracy issues on the underlying geographic layer may benefit by widespread adoption of geographical methods and the archaeological approach may limit to apply results developed elsewhere and analyse results from an archaeological perspective, fuzzy GIS need a stronger involvement because they heavily rely on expertise. Only the archaeologist expert in the problem under consideration is able to configure the membership functions in a meaningful way, and the technology limits to make correct deductions from these initial assumptions. There is a continuous feedback into the assumptions resulting from the progress in the investigation, which modifies the model adapting it to new perspectives emerging from the current state of the research. Fuzzy models may be better communicated to other researchers and agreed or discussed, because subjective parameters are expressed with numbers, the way mankind has used since a remote past to communicate concisely and with no misunderstandings. Paradoxically, in this case extending the realm of quantification to include even the simplest concepts may hopefully reconcile those who foster the use of computers and quantitative methods and those who claim that human subjectivity - luckily enough - cannot be removed from archaeological investigation.

References

Acsádi G.Y. and Nemeskéri J., 1970. *History of human life span and mortality.* Budapest: Akademiai Kiado.

Beex, W. 2004. Use and Abuse of Digital Terrain/Elevation Models. In: AA. VV. (eds.) *Enter the Past - The E-way into the Four Dimensions of Cultural Heritage.* Bar International Series 1227. Oxford, Archaeopress, 240-242.

Benvenuti A. and Niccolucci F., 1996. Historical objects with indeterminate boundaries. Paper presented at the IX International AHC Conference, Moscow.

Burrough P. A. and Frank A. U., 1996. *Geographic Objects with Indeterminate Boundaries.* London, Taylor & Francis.

Crescioli M., D'Andrea A. and Niccolucci F. , 2000. A GIS-based analysis of the Etruscan cemetery of Pontecagnano using fuzzy logic. In: G. R. Lock (ed.) *Beyond the Map: Archaeology and Spatial Technologies.* Amsterdam, IOS Press, 157-179.

D'Andrea, A., Gabrielli, R., Reali, E., Tucci, S. and Verde, G. , 2002. Data acquistion and three-dimensional models of medieval castles: a comparison between two case-studies Rocca Ricciarda (Tuscany) and Wu'ayra (Petra, Jordan). In: F. Niccolucci (ed.) *Virtual Archaeology.* BAR International Series 1075. Oxford, Archaeopress, 129-138.

Farinetti E. and Sbonias K, 2004. Fields of Wheat Back to the Land. A GIS Environment for the Study of Medieval Village History in Central Greece. In: AA. VV. (eds.) *Enter the Past - The E-way into the Four Dimensions of Cultural Heritage.* Bar International Series 1227. Oxford, Archaeopress, 271-275.

Galindo J., Medina J. M., Pons O. and Cubero J. C., 1998. A Server for Fuzzy SQL Queries. In: T. Andreasen, H. Christiansen, and H. L. Larsen (eds.) *Flexible Query Answering Systems. Proceedings of the Third International Conference, FQAS '98, Roskilde.* Lecture Notes in Artificial Intelligence 1495. Berlin, Springer, 164-174.

Hatzinikolaou E, Hatzichristos T, Siolas A and Matzourani E, 2003. Predicting Archaeological Site Locations Using G.I.S. and Fuzzy Logic. In: M. Doerr and A. Sarris (eds.) *The Digital Heritage of Archaeology.* Athens, Archive of Monuments and Publications, Hellenic Ministry of Culture.

Loots L., Nackaerts K. and Waelkens M., 1999. Fuzzy viewshed analysis of the hellenistic City. Defence system at Sagalassos. In: L. Dingwall, S.C. Exon, V. L. Gaffney, S. Laflin and P.M. Van Leusen (eds.) *Archaeology in the Age of the Internet.* BAR International Series 750. Oxford, Archaeopress, 63-65.

Nackaerts K., Govers G. and Loots L., 1999. The use of Monte Carlo techniques for the estimation of visibility. In: L. Dingwall, S. C. Exon, V. L. Gaffney, S. Laflin and P. M. Van Leusen (eds.) *Archaeology in the Age ot the Internet.* BAR International Series 750. Oxford, Archaeopress, 82.

Niccolucci, F., D'Andrea A. and Crescioli, M., 2001. Archaeological applications of fuzzy databases. In: Z. Stancic and T. Veljanovski (eds.) *Computing Archaeology for Understanding the Past.* BAR International Seris 931. Oxford, Archaeopress, 107-115.

Podobnikar T., Tecco Hvala S. and Dular J., 2004. Iterative Approach to Ancient Paths Modelling in the Iron

Age Study of the Doleniska Region (Slovenia). In: AA. VV. (eds.) *Enter the Past - The E-way into the Four Dimensions of Cultural Heritage*. Bar International Series 1227. Oxford, Archaeopress, 254-257.

Schneider M., 1999. Uncertainty Management for Spatial Data in Databases: Fuzzy Spatial Data Types. In: R. H. Güting, D. Papadias and F. Lochovsky (eds.), *Advances in Spatial Databases, Proceedings of the 6th International Symposium, SSD '99, Hong Kong*. Lecture Notes in Computer Sciences 1651. Berlin, Springer.

Stefanakis E., Vazirgiannis M. and Sellis T., 1996. Incorporating Fuzzy Logic Methodologies into GIS Operations. In: T. Sellis. and D. Georgoulis (eds.), *Proceedings of the Athens International Conference on Urban Regional Environmental Planning and Informatics to Planning in an Era of Transition*, Athens, Greece, National Technical University of Athens.

Topouzi, S., Tripolitsiotis, A., Sarris, A., Mertikas, S. and Soetens, S. 2001, Errors & Inaccuracies in Repositioning of Archaeological Sites. In G. Burenhult and J Arvidsson (eds.) *Archaeolgical Informatics: Pushing the Envelope*. BAR International Series 1016. Oxford, Archaeopress, 417-426.

考古学GISアプリケーションの不確実性とその対応

フランコ・ニコルッチ

フローレンス大学

我々は、考古学における広範なGISの応用の際、「過去のイベントとその状態に関する我々の知識は、間接的な知識の結果であり、研究者の解釈を基礎としている」ということを常に意識しているわけではない。したがって、それは資料批判のエラーと不確定性の影響を大きく受けることになる。通常、それらは平均化され、データは確実であるかのように扱われる。しかし、不確定性の正しい管理が処理結果の信頼性を醸成し、そしてエラーを積み重ねることで引き起こされるミスを避けることによって、境界線上のデータの処理を可能とする。

例えばGPSのような、自動的にデータを収集するような自動化されたツールの利用は、そうしたツールによって誘発されるエラーと、システムにおけるそうしたツール普及の評価とを常に伴っていなくてはならない。残念ながら、こうした適用評価の分析は、考古学的な応用においては非常に希である。この論文では、こうした方法論的研究について、まず既存の方法について簡潔にまとめ、推奨すべき方法について提案し議論する。

これとは別に、"不完全な"概念定義に関連する問題もある。それがデータの処理において非常に頻繁に使用されるにもかかわらず，この問題は上記の問題と同様に議論されていない。たとえば、空間概念で言えば「丘」などの定義であり、人類学的概念で言えば「青年」などの定義である。これらは、－どこが平野の終わりでどこが丘の始まりか？－、－いつが青年の終わりでいつから大人になるのか？－、というような、いわゆる"Soriteのパラドクス"の影響を受けている。

このような場合、"実在"と"不在"の境界は鮮明なラインではなく、こうした不明瞭さはパラドクスを決定的にする。そしてそれは現在のコンピュータツールのほとんどが明確に定義されたデータを処理するよう設計され、ファジーな実体を扱えない、ということだけでそうなっているのである。

この論文では、特に空間的なアプローチが必要とされる場合に、考古学的な文脈の中で不確実に、あるいは不完全に定義された概念を扱う方法を提起する。考古学的な事例研究として、墓域における墳墓の空間分布研究、考古学調査の結果として生じる遺跡の定義、環境や地形に基礎をおいたセトルメントパターン分析などを扱う。