THE BENEFIT OF FUZZY LOGIC TO PROTECTION OF CULTURAL AND HISTORICAL HERITAGE

Renata, ĎURAČIOVÁ [1], Tibor, LIESKOVSKÝ [2], Eva, STOPKOVÁ [3], Kristína, KROČKOVÁ [4]

[1]Department of Theoretical Geodesy, Faculty of Civil Engineering, Slovak University of Technology in Bratislava, Radlinského 11, 813 68, Bratislava, Slovakia

[1]renata.duraciova@stuba.sk, [2]tibor.lieskovsky@stuba.sk, [3]eva.stopkova@stuba.sk, [4]kristina.krockova@stuba.sk

Abstract

The development of urban agglomerations and human settlements brings increasing a pressure to effective protection and preservation of cultural and historical heritage. Geographic information systems could play here an important role, on the one hand to help register and manage approachable information about existing cultural and historical heritage, on the other hand they can help in the prediction of areas with high cultural and historical potential that are threatened by the development of human settlements and the city transformations. Early prediction of endangered areas allows their consequent protection, respectively inclusion into the urban planning and suchlike.

This contribution deals with different aspects of the application of fuzzy logic in archaeological predictive modelling, which represent an effective tool to identify areas with potential of archaeological sites presence. The base constitutes the issue of determining the course of fuzzy membership functions for each input layer to the prediction. Onwards we deal with the difference between aggregate functions and individual t-norms that have a direct influence to the creation of predictive models. Finally, we devote to the fuzzy clustering that allows you to more effectively identify and single out categories of input features – archaeological sites. As well an increasing the process quality of archaeological predictive modelling, from creating individual prediction models for various components of archaeological sites, up to the possibilities of enhancing a validation quality of predictive models.

Keywords: cultural and historical heritage, fuzzy logic, predictive modelling, archaeology, t-norms, clustering

SOURCES OF UNCERTAINTY IN ARCHAEOLOGY

One of the main issues in the application of GIS in archaeology is the geometric positioning of the archaeological site. In the current state of archaeological evidence in Slovakia can be accuracy of positioning moved within 100 m, especially for older and unverified records. This value is based on the rules of the Central Registry of Archaeological Sites in Slovakia (CEANS) where the basic unit of the registration system in the database is a point with diameter of 4 mm placed on the map with a scale of 1: 25 000, what means 100 m in the field (Bujna, 1993). This point form record of archaeological site is a significant source of uncertainty (Fig. 1). It means the generalization towards reality insofar as archaeological sites comprise large area polygons. The issue is determining the centroid of known archaeological site (core of site versus its geometric centre) as well as its scope. It is rare that the site is exposed and whole area of the site is known.

Archaeological findings itself fall to various transformation processes; doom process or spatial movements, such as erosion, soil movements, or suchlike. Similarly, in modelling certain phenomena and the application of theoretical knowledge is necessary to take into account human behaviour, not always based on rational principles and characterized by high variability and adaptability. This is reflected in the vagueness of the opinions in the investigation of human history and human behaviour ("proximity to water", "suitable slope", "suitable land") what is the cause of confusion in definition of the certain standards in spatial analyzes. All these facts bring to spatial analyzes significant degree of uncertainty and distortion. Fuzzy logic allows better reflecting the natural estimated properties and modelling the uncertainty of archaeological spatial data

(Lieskovský et al., 2011). Fuzzy sets were used for modelling of spatial data for example in (*Hwang* and *Thill, 200*5).

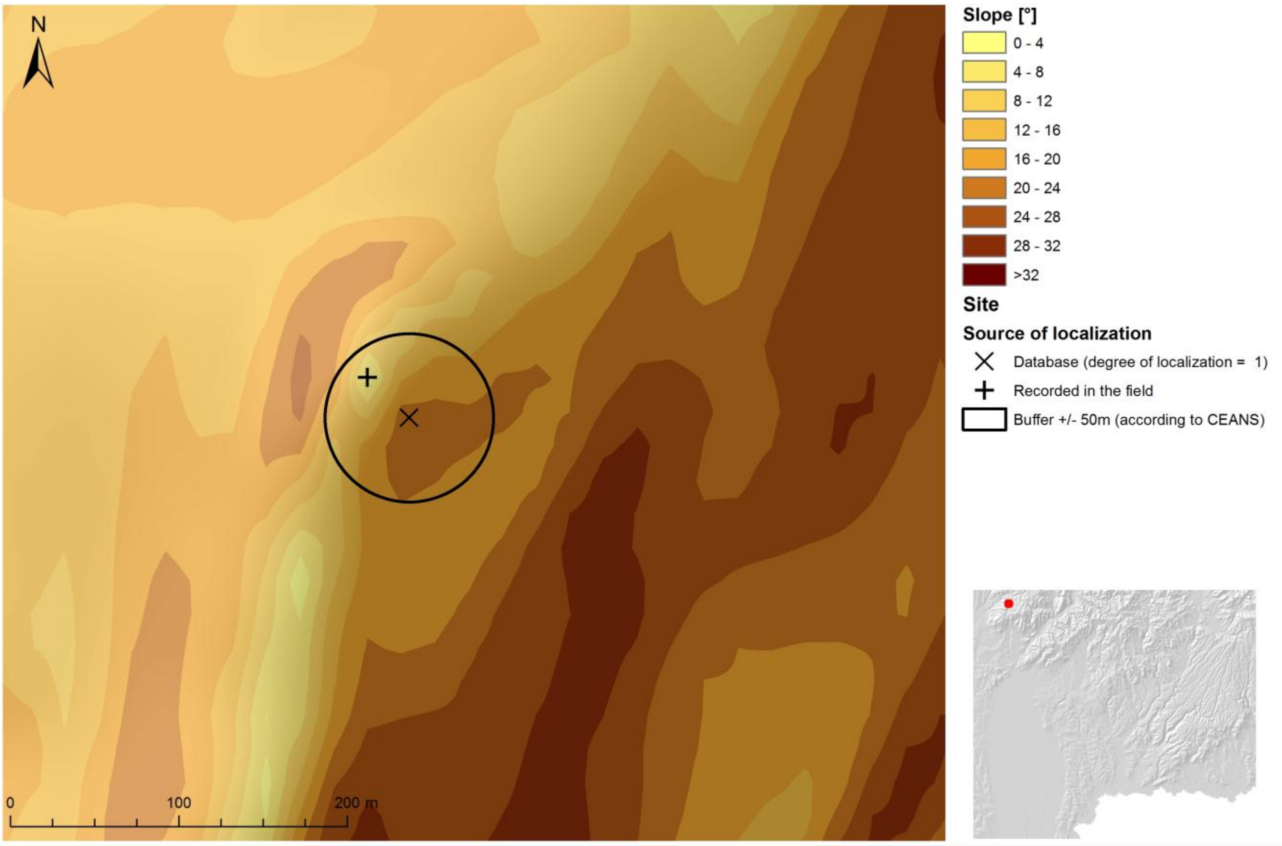


Fig. 1. The difference in positions of the site (real position and position according to database of archaeological sites evidence) (Lieskovský, 2011)

BASIC CONCEPTS OF FUZZY SETS

The characteristic function of the crisp set $\chi_A : X \rightarrow \{0,1\}$ can be formalised (Novák, 2000) as:

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases} \tag{1}$$

The fuzzy set *A* of a universe *X* is defined by a membership function $\mu_A(x)$ such that $X \rightarrow [0,1]$, where $\mu_A(x)$ is the membership value of *x* in *A* (Kainz, 2012). (A fuzzy set is a set of elements $x \in X$ (*X* is called universe), where each of them is assigned by a degree of membership $\mu_A(x)$, whose values must be between zero (no membership) and one (definite membership).)

Fuzzy sets are, therefore, means that provides the ability to mathematical description of concepts of vagueness and to a work with them. Degree of membership reflects the rate to which the element belongs to the set.

The shape and parameters of the membership functions can be determined based on practical experience or on the known properties of the analyzed phenomenon. Trapezoidal (piecewise linear) membership function is the most commonly use one (Fig. 2a):

$$\mu_A(x) = \begin{cases} 0 & \text{if } x < a, \\ \dfrac{x-a}{b-a} & \text{if } a \le x \le b, \\ 1 & \text{if } b < x < c, \\ \dfrac{d-x}{d-c} & \text{if } c \le x \le d, \\ 0 & \text{if } x > d. \end{cases} \tag{2}$$

Other frequently used membership function shapes are e.g. Gaussian or sinusoidal (Fig. 2b).

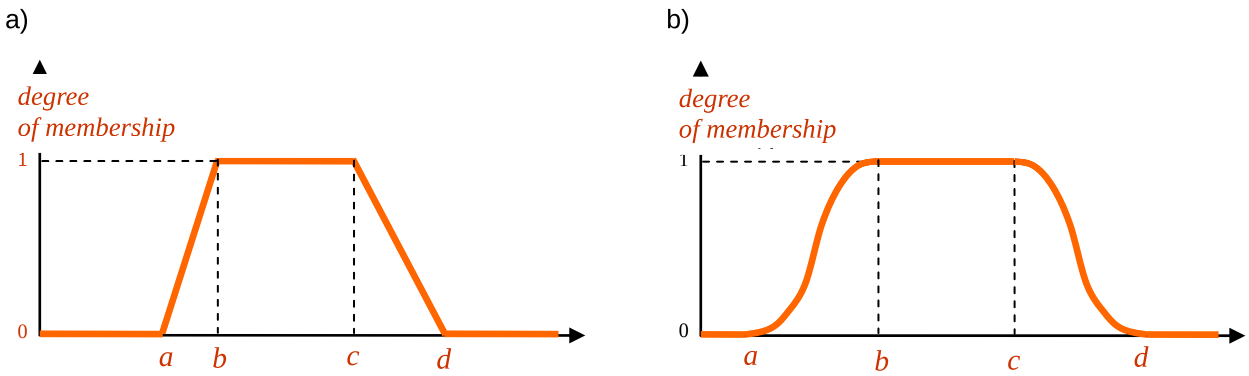


Fig. 2. a) Trapezoidal and b) sinusoidal membership function

PARAMETER DETERMINATION OF FUZZY MEMBERSHIP FUNCTIONS

In determining the parameters of fuzzy membership functions, we use similar approaches as for predictive modelling (Lieskovský et al., 2011):

Deductive approach is appropriate in cases when sufficient knowledge exists about the expected knowledge of phenomena based on the results of archaeological excavations, archaeological theory, and possibly sufficient experiences. Similarly, deductive approach is the only option of modelling in areas with absence of information about archaeological localities, or in case where can be assumed that the data are not statistically relevant (low amount, uneven examination of the interest area).

Application of inductive approach allows us partially to objectify some assertions and used assumptions; and at the same time, it is suitable for the determination of individual parameter weights. However the approach does not fully model some phenomenon and is subject to distortion caused by uneven examination of the interest area, the accuracy of localization and failure the differentiation of various types of archaeological sites (Šmejda, 2003).

In case of a deductive approach to modelling, the course of membership function of a fuzzy set representing the proper conditions for archaeological sites (e.g. Fig. 3) was determined based on expert judgment or based on available data in the archaeological literature. An alternative could be to determine a function course of the statistical processing of input spatial data (for instance extrapolated from the frequency histogram of individual characteristics of localities and landscape). Probably the most appropriate approach is the combination of two methods described above. For example, to create a model in deductive approach with using fuzzy logic is possible to adjust the shape of membership functions based on the results of statistical tests of known parameters and in a similar way to set the weight of each layer. On the other hand, in the case of applying mainly inductive approach we can eliminate remote and extreme values of the statistical tests based on the deductive assumptions (Lieskovský, 2011).

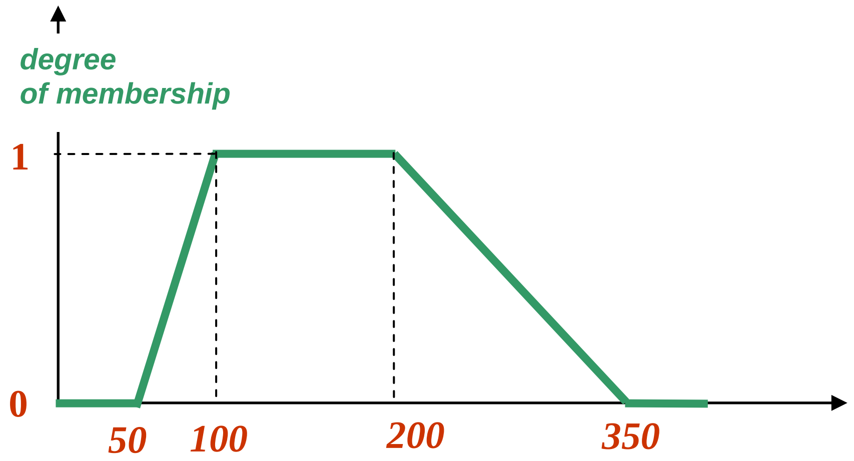


Fig. 3. Fuzzy membership function of distance between sites and watercourses

DETERMINATION OF INPUT LAYERS WEIGHTS

In most references about determination of weights for archaeological modelling is used deductive method of determining the weights. Inductive determination of weights designed by us is based on comparison of differences in two distribution of the phenomenon at sites f (x) and phenomenon in the landscape g (x). We come out from assumption that the difference of two distributions is bigger the given phenomenon have greater influence on the choice of site (Fig. 4 – Fig.6). Coefficient of differences between two distributions defines as follows (Lieskovský, 2011):

$$w(f,g)=\int_{-\infty}^{\infty}\left|f(x)-g(x)dx\right| \tag{3}$$

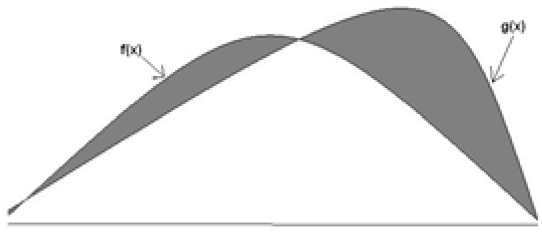


Fig. 4. Example of difference in distributions of phenomenon at sites and phenomenon in landscape

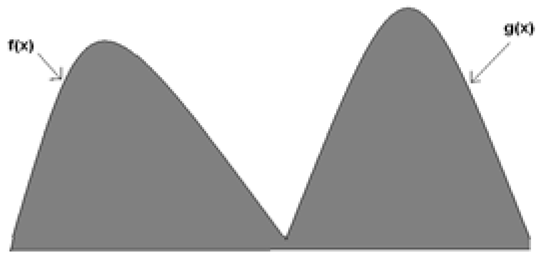


Fig. 5. The maximum difference in distributions of phenomenon at sites and phenomenon in landscape – difference gets value of 2

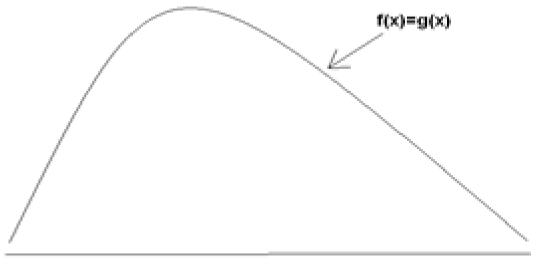


Fig. 6. The same distribution of phenomenon at sites and in landscape - difference gets value of zero

Since this rate differences is in range [0,2] it is proper to normalize:

$$w^N(f,g) = \frac{1}{2} w(f,g) \tag{4}$$

Examples of fuzzy layers are presented in Fig. 7 and their estimated weights are listed in Tab. 1.

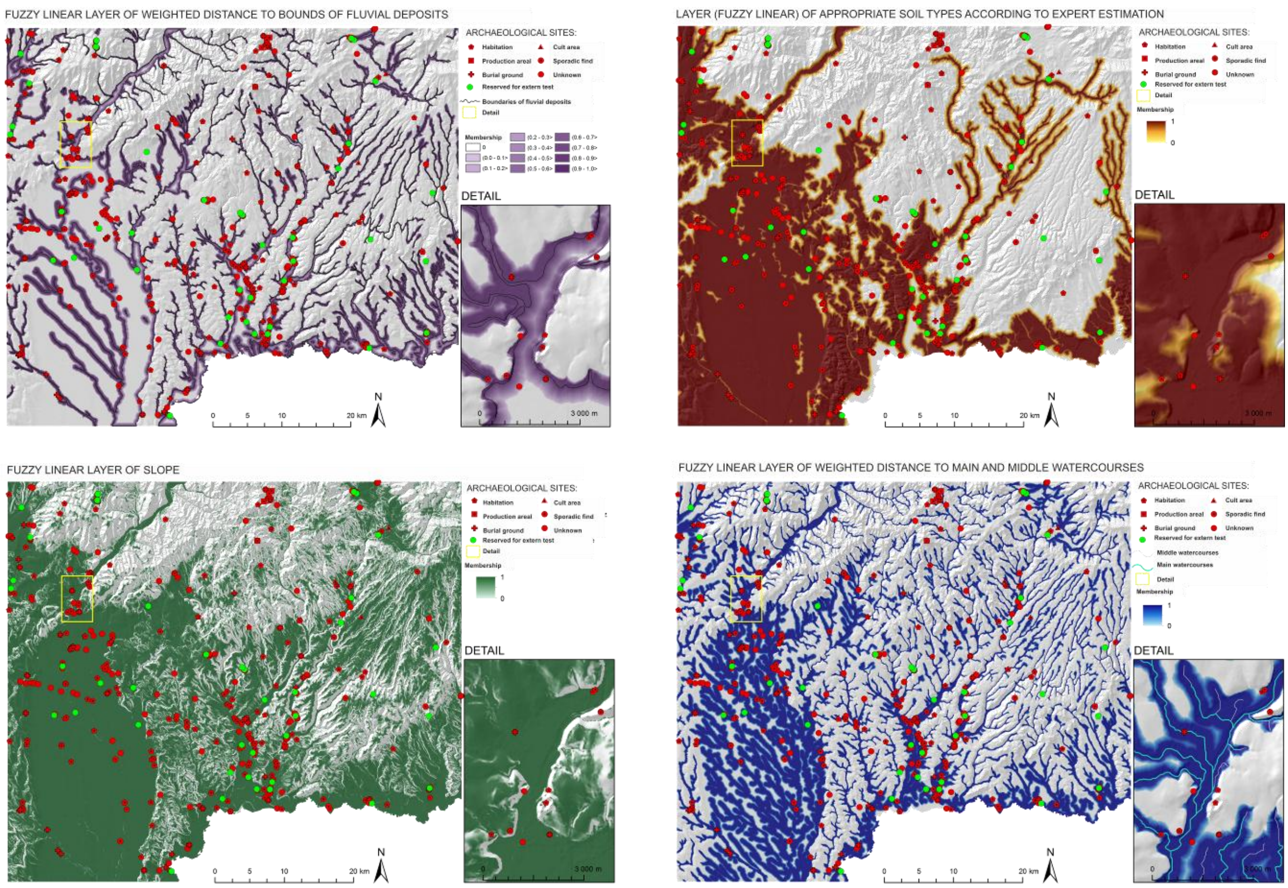


Fig. 7. Examples of fuzzy layers entering the prediction model with examples of calculated weights

Table 1. Examples of estimated weights for each layer

Layer name	Weight
Fuzzy linear layer of cost weighted distance to main and middle watercourses	0.588
Cost weighted distance to bounds of fluvial deposits	0.843
Slope	0.376
Main representation of soil type in perimeter of 500 m	0.5241

FUZZY LOGICAL OPERATIONS AND AGGREGATION FUNCTIONS

Logical operations like intersection, union and complement, which are corresponding to propositional operations conjunction, disjunction and negation in mathematical logic, are the basic functions with sets in GIS environment. Logical operators are also part of the query to spatial databases in SQL. The basic operations are defined in Boolean algebra for crisp sets. In spatial analysis for example, we need to select sites, which (Ďuračiová et al., 2011):

- must be near to fluvial deposits (less than 250 m),
- must have moderate slope or favourable aspect (e.g. south-west),
- must not be far away from a river (more than 350 m), but must not be too close (less than 50 m),
- must be in the lowland.

If:

- *F* is a set of sites located within 250 meters from boundaries of fluvial deposits,
- *S* is a set of sites with moderate slope (e.g. slope within 10°),
- *A* is the set of sites in south-west aspect,
- *L* is the set of sites with altitude less than 300 m,
- *R1* is the set of sites whose distance from the river is ≤ 50 m,
- *R2* is a set of sites whose distance from the river is ≥ 350 m,

then for a set of appropriate sites applies:

$$V = F \cap (S \cup D) \cap N \cap \neg(R1 \cap R2). \quad (5)$$

Fuzzy logical operators (with the truth-value of interval [0,1]) are the basis of fuzzy propositional calculus for operations with fuzzy sets (Navara & Olšák, 2004). In fuzzy set theory:

- fuzzy standard intersection of two sets *A, B* is fuzzy set *A*∩*B* with membership function (Fig. 8):

$$\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x)), \quad (6)$$

- fuzzy standard union of two sets *A, B* is fuzzy set *A*∪*B* with membership function (Fig. 8)::

$$\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x)), \quad (7)$$

- fuzzy standard complement of set *A* is fuzzy set $\bar{A}$ with membership function:

$$\mu_{\bar{A}} = 1 - \mu_A(x). \quad (8)$$

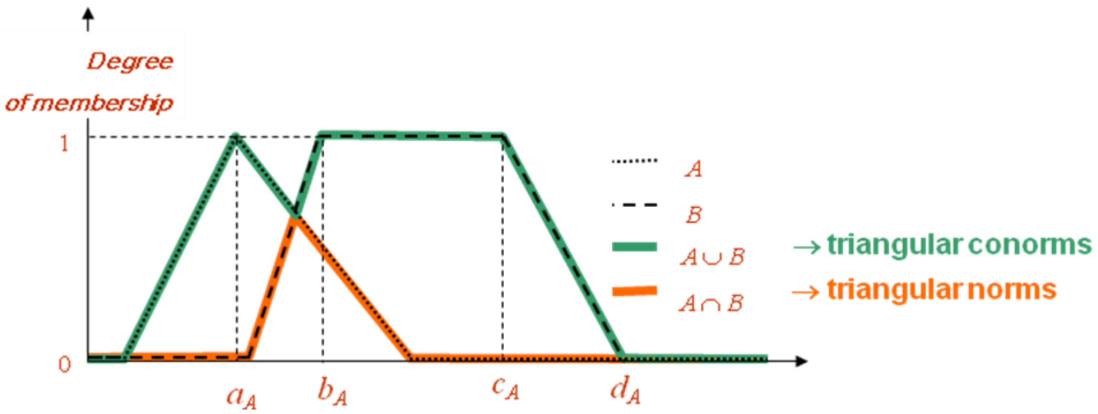


Fig. 8. Fuzzy intersection (conjunction) and fuzzy union (disjunction)

Generally, fuzzy intersection (conjunction) is in many cases defined by the so-called triangular norms (t-norms). The generalization of the classical union (disjunction) is interpreted by triangular conorms (t-conorms) (Kolesárová & Kováčová, 2004). The most commonly used t-norms are (Fig. 9):

$$T_M(x,y) = \min(x,y)$$ - minimum t-norm (9)

(it corresponds to the standard intersection)

$$T_P(x,y) = xy$$ - product t-norm (10)

$$T_Ł(x,y) = \max(0, x+y-1)$$ - Łukasiewicz t-norm (11)

$$T_D(x,y) = \begin{cases} \min(x,y), & \text{if } \max(x,y)=1 \\ 0, & \text{else} \end{cases}$$ - drastic t-norm (12)

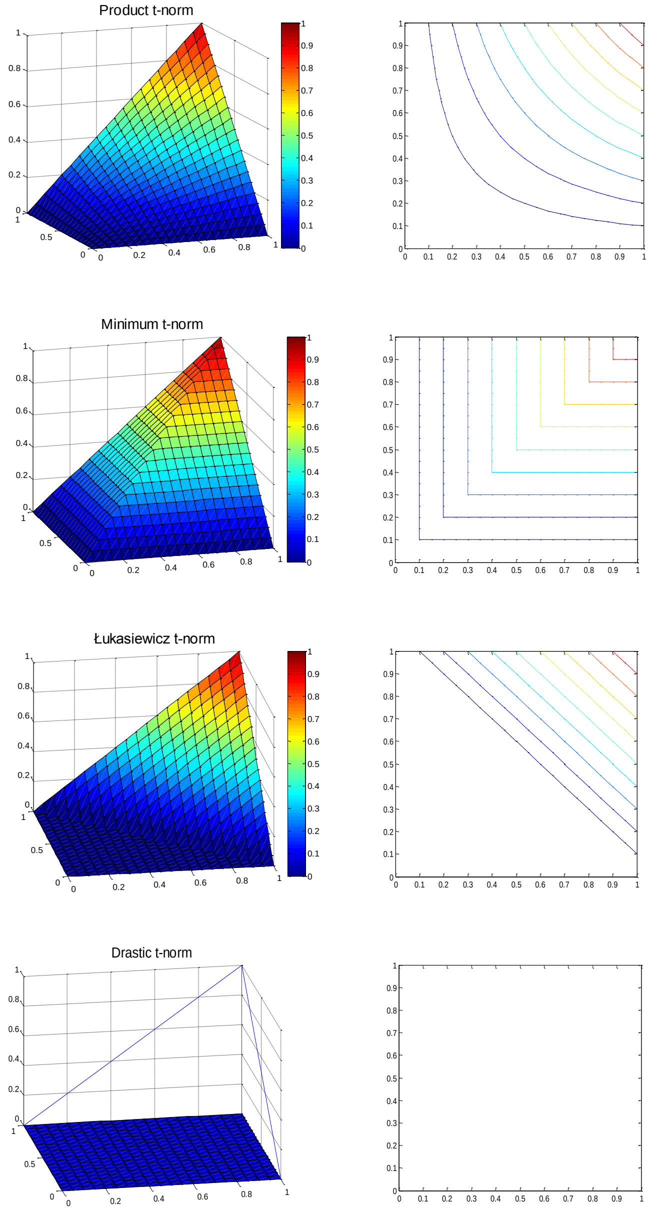


Fig. 9. The most commonly used t-norms

Weighted average is an appropriate aggregation function for the introduction of various weights of each criterion modelled by fuzzy sets. The impact of choice of overlapping function (t-norm, weighted average) to the resulting predictive model (detail) is shown in Fig. 10.

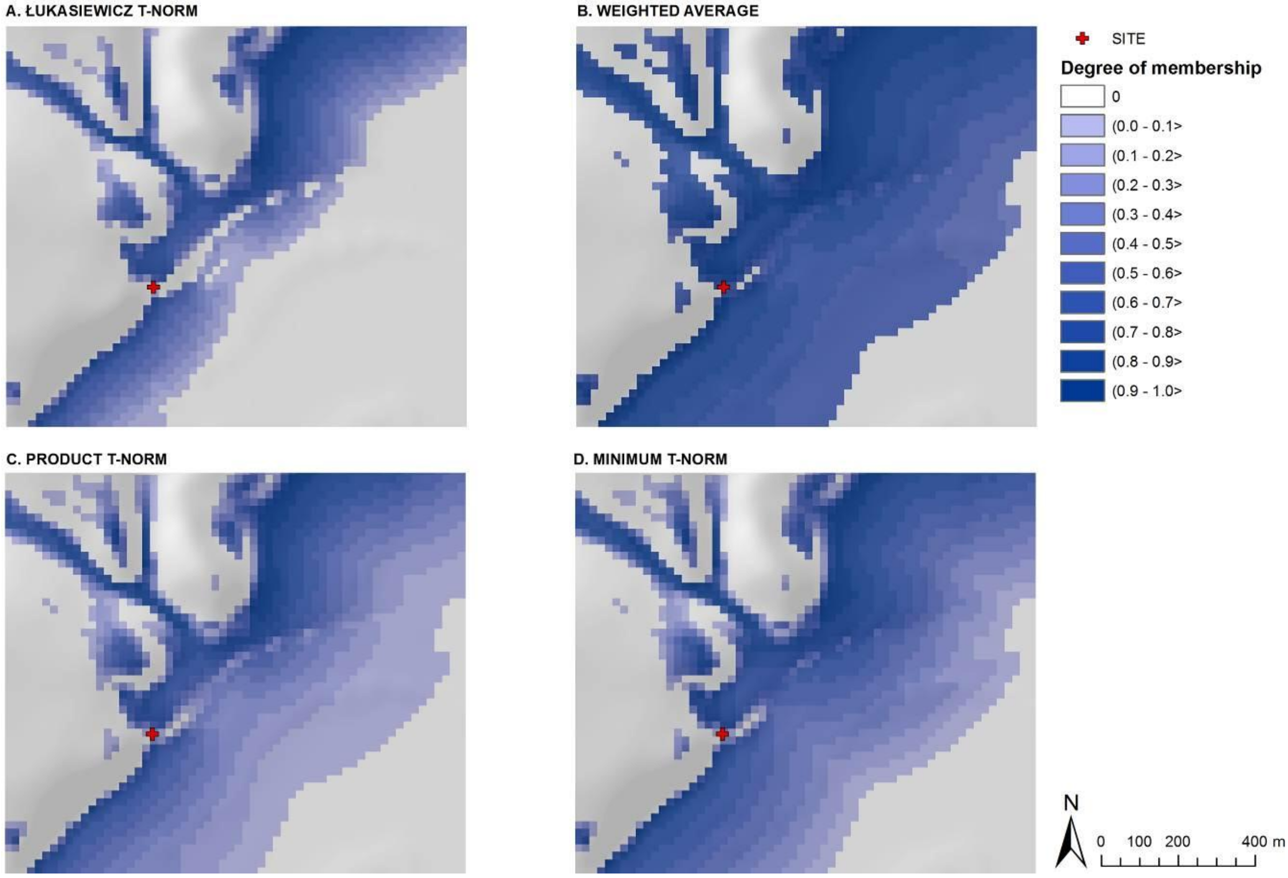


Fig. 10. The impact of choice of overlapping function (t-norm, weighted average) to the resulting predictive model (detail)

Spatial decision making using fuzzy sets is discussed for example in (*Morris* and *Jankowski, 2005*) or in (*Šmejda, 2003*).

POSSIBILITEIS OF FUZZY LOGIC USAGE IN CLUSTER ANALYSIS

Determination of parameters of fuzzy membership functions could be in some cases complicated by fact, that knowledge of the phenomenon may be inadequate to describe all its variants. For example, if we analyze optimal conditions for settling habitation (shown upper), these criteria may valid for casual settlement in ancient ages. However, hill forts were settled, according to (*Štefanovičová, 2001*), at steep slopes or close to rivers, wetlands or at lakes and they could not be described by the criteria. Prediction model based on these criteria could not detect potential site - hill fort. If type of habitation (casual or hill fort) is known, it is possible to develop more prediction models based on criteria that describe properties of both groups.

Parameters of fuzzy membership functions have been estimated using data of archaeological sites evidence. The items differ in accuracy of position and of attributes and they are different in complexity of information. Absence of note about habitation type may lead to some uncertainties, for example during prediction model testing there have been 13 known sites undetected - that means almost 1% of total number of sites.

One of the possibilities to eliminate the risk of potential site rejecting based on properties that not completely suit to the prediction model is application of fuzzy cluster analysis. This means data clustering - determining if element belongs to the membership of similar elements or not. Fuzzy theory enables to determine, if the element belongs to more than one group and also ratio of this relationship (*Mika and Lakhmi, 2006*). Grouping of similar elements enables to differentiate special cases and obtain information about variations of the data. Fuzzy C-Means (FCM) is one of the most often used methods based on the minimalism of sum of weighted differences between feature values and average value of the group (*Bezdek, 1981*).

Analytical tool *Grouping analysis* based on similar principle (*ESRI, 2012*) have been developed and implemented to ArcGIS 10.1 (ESRI) into toolbox *Mapping Clusters*. Topic is quite circumstantial; we will try to explain experiment focused on surveying of hill forts common signs, specifically of potential differences in their site terrain configurationWe expected hill fort to be located at:

- raised locality with smooth or little more steep slope (flatland),
- locality in medium or higher hills with steep slope,
- locality in medium or higher hills with very smooth slope (top of the hill).

Analysis results are summarized in Fig. 11, Graph 1 and Graph 2. Graph 1 describes memberships of elements with minimal sum of weighted differences between values of elevation or slope and the mean of the membership. Graph 2 combines memberships according to distribution of elevation and slope. According to experimental criteria, memberships have been differed as *hill forts* and *habitations* (Fig. 11). As optimal number of memberships was empirically confirmed membership of five.

The result of experiment have been tested by comparing with known castles, hill forts and fortifications, that have been researched in literature, maps and in some cases found as note in database. For check was used also layer of sites that never been predicted before during testing of prediction models. Fig. 11 shows coincidence of majority of known fortifications with unpredicted sites and with membership of *hill forts*. Exceptions might represent Levice Castle (located at flatland close to the town) and Čabraď Castle (protected more by river than by steep terrain). Other massive exception we can found in Banská Štiavnica surrounding, there is too many elements in membership of *hill forts* because of special terrain configuration.

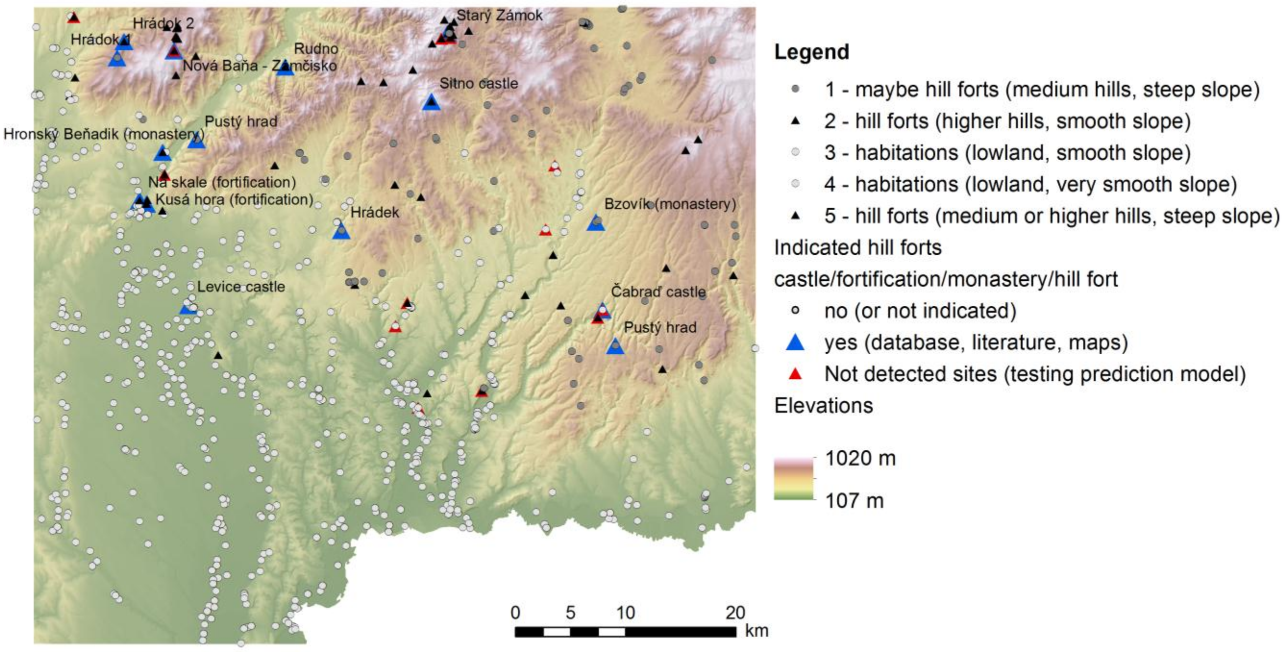


Fig. 11. Results of cluster analysis of sites - *hill forts* and *habitations* memberships

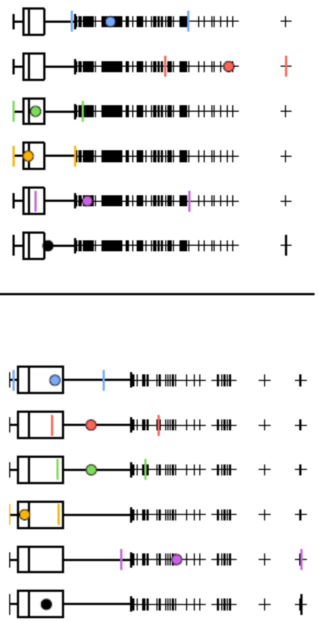
Variable-Wise Summary

H_ELEV_M: R2 = 0.86

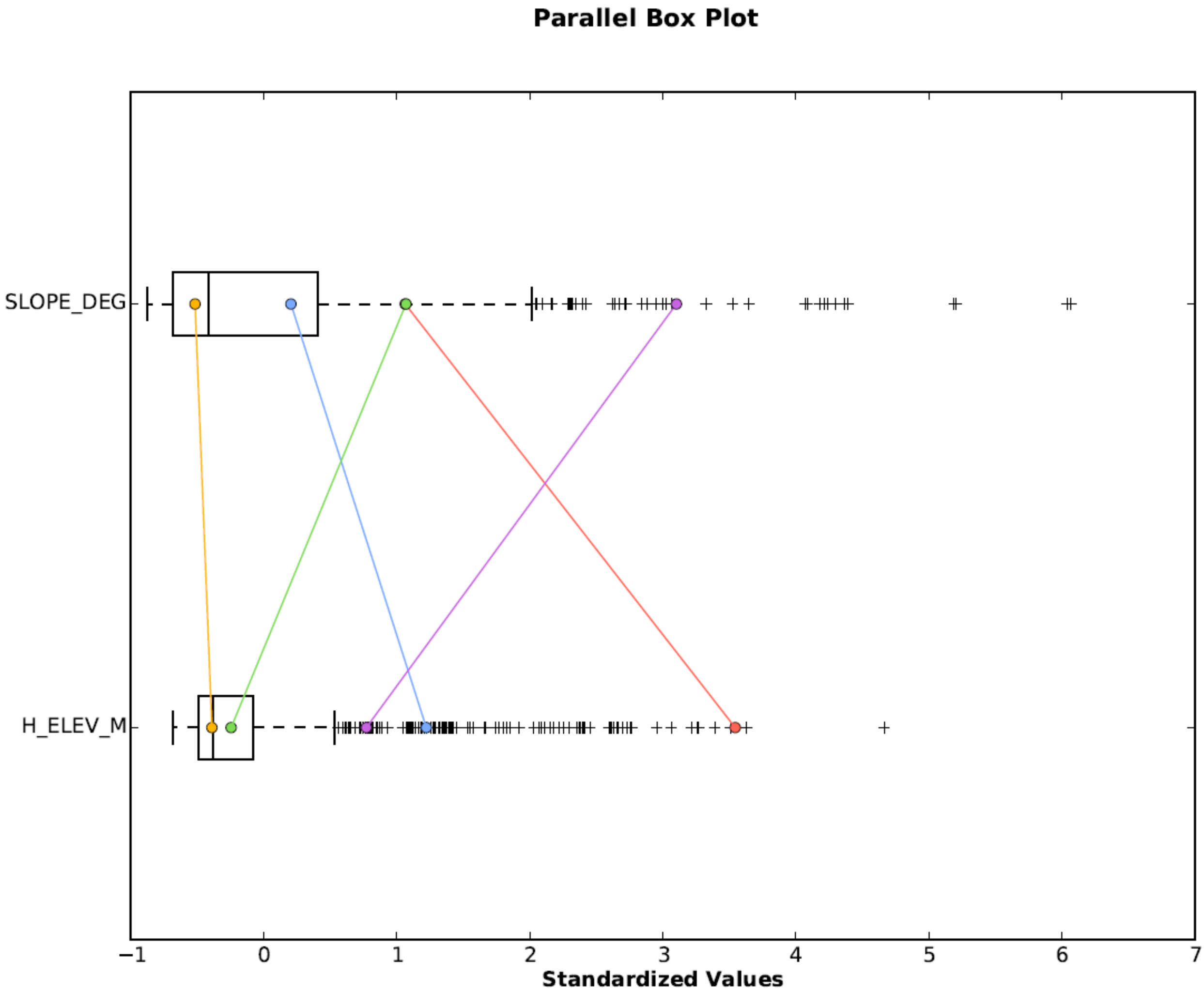
Group	Mean	Std. Dev.	Min	Max	Share
1	431.5814	80.9031	303.7827	688.0181	0.4293
2	820.3468	165.1267	611.9446	1007.7792	0.4422
3	185.9384	49.2733	113.9975	342.2664	0.2550
4	161.7254	32.4184	112.6567	314.5240	0.2255
5	356.1201	144.7384	187.0491	690.0831	0.5620
Total	226.7521	167.4192	112.6567	1007.7792	1.0000

SLOPE_DEG: R2 = 0.83

Group	Mean	Std. Dev.	Min	Max	Share
1	4.4887	2.0845	0.3580	9.3029	0.3112
2	8.0438	2.3141	4.1717	14.7037	0.3664
3	8.0673	2.4141	4.7821	13.3524	0.2982
4	1.5067	1.1311	0.0092	4.8367	0.1680
5	16.4802	3.8969	11.0452	28.7526	0.6161
Total	3.6376	4.1382	0.0092	28.7526	1.0000



Graph 1. Memberships according to characteristics of elevation and slope



Graph 2. Combinations of variables describing memberships

Topic will be research in the future because of potential in adding more parameters (for example soil types) of cluster analysis or in defining specific properties of fortifications settled in river or wetland environment. The results could be more precise after obtaining more information about known hill fort from literature and other sources. Membership 1 (potential hill forts) should be also analysed because it might contain habitations except hill forts - it depends on similar terrain configuration.

CONCLUSION

Application of fuzzy approach in GIS means the contribution to increasing of quality and plausibility of spatial analyses. Fuzzy sets enable to model, analyse and describe phenomena in more natural way: with regard on their constitution and (in)accuracy of available data describing. Fuzzy approach may be useful for example in archaeological prediction modelling. High degree of inaccuracy is caused by absence of common approach to data acquisition, by long term of recording archaeological data and by generalization of polygonal sites to point data. In addition, quality and scale of other spatial data used in spatial analyses belong to important aspects of inaccuracy (for example maps of soil types as results of interpretation and interpolation of point data). Fuzzy sets mean more accurate way how to model the settlement propriety of the land in the past. It is more relevant to model for example fact, that localities within a distance of 350 m could be the most appropriate to settlement and within a distance of 500 m the appropriateness decreases, than to claim, that only localities within a distance of 350 m could be settled.

Application of fuzzy logic itself does not guarantee automatically more accurate and correct results in case of inconsistent knowledge of the phenomenon, or in case of too many degrees of freedom. It helps to quantify known phenomena in more objective way like to eliminate inaccuracy of them or minimize loss of approximate information too. That is the reason why multi-criteria decision-making with usage of fuzzy logic principle may be adequate method to develop prediction models in GIS. It provides high flexibility that enables to adapt model to reality by setting appropriate fuzzy sets and by using correct rules of evaluation.

ACKNOWLEDGMENT

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0249-07.

REFERENCES

Bezdek, J. C. (1981) Pattern Recognition with Fuzzy Objective Function Algorithms. Norwell (Massachusetts, USA) : Kluwer Academic Publishers. ISBN 0306406713.

Bujna, J. and com. (1993) CEANS – Centrálna evidencia archeologických nálezísk na Slovensku – Projekt systému. In Slovenská archeológia XLI 2. Nitra, AÚ SAV, 1993, pp. 367-386.

Ďuračiová, R. - Lieskovský, T. - Kročková, K. - Sabo, M. (2011) Multikriteriálne rozhodovanie pomocou fuzzy množín v prostredí GIS a jeho využitie v archeologickej predikcii. In: Geodetický a kartografický obzor: odborný a vedecký časopis Roč. 57 (99), č. 8., pp. 205 - 215. ISSN 0016-7096.

ESRI (2012) ArcGIS Forum [online]: http://forums.arcgis.com/threads/20288-fuzzy-c-means-cluster-analysis [cit. 2012-10-19]

Hwang, S. and Thill, J. (2005) Modeling Localities with Fuzzy Sets and GIS. In: Petry, F. E., – Robinson, V. B. – Cobb, M. A. Fuzzy modeling with Spatial Information for Geographic Problems. Berlin: Springer, 344 p. ISBN 3-540-23713-5.

Kainz, W. Fuzzy Logic and GIS [online]: http://www.scribd.com/doc/44096805/Fuzzy-in-GIS-Basic [cit. 2012-10-19].

Kolesárová, A. – Kováčová, M. (2004) Fuzzy množiny a ich aplikácie. Bratislava, STU v Bratislave, 160 p. ISBN 80-227-2036-4.

Lieskovský, T., Faixová Chalachanová, J., Ďuračiová R., Blažová, E. (2011) Archeologické predikčné modelovanie z pohľadu geoinformatiky. Bratislava, REMPrint s.r.o., ISBN 978-80-970741-0-4.

Lieskovský, T. (2011) Využitie geografických informačných systémov v predikčnom modelovaní v archeológii. Dizertačná práca. KGZA SvF, Slovenská technická univerzita, Bratislava.

Mika, S. I. and Lakhmi, C. J. (2006) Innovation in Fuzzy Clustering. Berlin : Springer-Verlag, [online]: http://books.google.sk/books?id=myOoxdsjx4wC&printsec=frontcover&hl=sk&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false [cit. 2012-10-19]

Morris, A. and Jankowski, P. (2005) Spatial Decision Making Using Fuzzy GIS. In: Petry, F. E., – Robinson, V. B. – Cobb, M. A. Fuzzy modeling with Spatial Information for Geographic Problems. Berlin: Springer, 344 p. ISBN 3-540-23713-5.

Navara, M. and Olšák, P. (2002) Základy fuzzy množín. Praha, Vydavatelství ČVUT, ISBN 80-01-0258-3.

Novák, V. (2000) Základy fuzzy modelování. Praha, BEN, 2000, 80-7300-009-1.

Šmejda, L. (2003) Možnosti využití techniky „Multi-Criteria Evaluation“ v prostorové archeologii. Příspěvky k prostorové archeologii 1. Plzeň, Aleš Čeněk, pp. 231-246.

Štefanovičová, T. (2001) Tisícročné hradiská Slovanov. In Historická revue. Júl 2001. [online]: http://www.historickarevue.com/archiv_8.html [cit. 2012-10-19].