Reports on Geodesy and Geoinformatics vol. 104/2017; pp. 115-130 DOI: 10.1515/rgg-2017-0020

Original article

Received: 19 July 2017 / Accepted: 20 December 2017

FUZZY SIMILARITY AND FUZZY INCLUSION MEASURES IN POLYLINE MATCHING: A CASE STUDY OF POTENTIAL STREAMS IDENTIFICATION FOR ARCHAEOLOGICAL MODELLING IN GIS

Renata Ďuračiová, Alexandra Rášová, Tibor Lieskovský

Department of Theoretical Geodesy, Faculty of Civil Engineering,
Slovak University of Technology in Bratislava

Abstract

When combining spatial data from various sources, it is often important to determine similarity or identity of spatial objects. Besides the differences in geometry, representations of spatial objects are inevitably more or less uncertain. Fuzzy set theory can be used to address both modelling of the spatial objects uncertainty and determining the identity, similarity, and inclusion of two sets as fuzzy identity, fuzzy similarity, and fuzzy inclusion. In this paper, we propose to use fuzzy measures to determine the similarity or identity of two uncertain spatial object representations in geographic information systems. Labelling the spatial objects by the degree of their similarity or inclusion measure makes the process of their identification more efficient. It reduces the need for a manual control. This leads to a more simple process of spatial datasets update from external data sources. We use this approach to get an accurate and correct representation of historical streams, which is derived from contemporary digital elevation model, i.e. we identify the segments that are similar to the streams depicted on historical maps.

Keywords: *spatial data uncertainty, similarity measure, fuzzy inclusion, spatial object matching, identity determination*

1. Introduction

The analysis, usage, and integration of spatial data from various data sources in geographic information systems (GIS) often require determining whether the spatial objects acquired from these sources are identical or at least similar. Spatial objects can be represented in different ways depending on the data sources. For instance, it is practically impossible that objects like rivers or roads will be represented as spatial objects with identical geometry by two independently created maps. Therefore, the determination of identity or similarity of two spatial objects can be problematic and is furthermore complicated by the inherent uncertainty of the object representation itself. The process of merging information from multiple data sources into one version is described by the terms “spatial data matching” (Walter and Fritsch, 1999) or “conflation” (Longley et al., 2001); a number of various methods used for the conflation of geospatial data is discussed, for example, by Cobb et al. (1998), Samal et al. (2004), Seth and Samal (2008), Chen and Knoblock (2008), Koukoletsos et al. (2012) or Toomanian et al. (2013). Some of the methods to detect or determine the similarity of objects are implemented in commonly used GIS software. Examples of such tools are *Similarity search* to identify which candidate spatial objects are most similar to input objects based on attributes, *Feature matching* to find corresponding features from two similar datasets based on the search distance, *Spatial adjustment rubbersheeting* to align one layer with another that is in close proximity. However, none of these is suitable for determining the identity or the similarity of uncertain representation of linear objects. The question of accuracy and uncertainty in GIS is recognised from the early years of GIS and it is still present in current research, e.g. (Heuvelink and Brown, 2008), either as an issue of positional accuracy (Goodchild and Hunter, 1997), uncertainty in GIS (Zhang and Goodchild, 2002), or fuzzy modeling in GIS (Petry et al., 2005).

Fuzzy set theory, introduced by Zadeh (1965), provides means to represent uncertain data and this concept can be extended to spatial data. In this paper, we use generalized concepts of fuzzy inclusion, fuzzy similarity, and fuzzy identity of two fuzzy sets (Bandemer 2006, Wang 1997, Wygralak 1983, Zadeh 1965). We propose a new method to determine the fuzzy similarity and the fuzzy inclusion of two uncertain polyline objects. The fuzzy approach used in this paper can improve, for instance, hydrological modelling in archaeological analysis or other applications, where a comparison of two uncertain polyline objects is necessary. In our case study, we use this method to identify matching streams derived from different sources, a modern digital terrain model (DTM) and historical maps. The main issue is that both stream datasets are uncertain and have a different geometry.

2. Methodology

2.1. Polyline similarity measures in GIS

Measures of similarity are different measures of the distance, the correlation, or the association; all of these are based on the proximity of objects or, on the contrary, on the distance of two objects, or their attributes. Similarity is usually recorded as a value in the intervals [−1, 1] or [0, 1], where 1 indicates the maximum of similarity, and -1 (0 respectively) the minimal similarity.

The similarity of two polyline objects can easily be determined using the distance measures (e.g. Euclidean, Manhattan, Minkowski, Mahalanobis), as long as the number of vertices on one polyline exactly matches the number of vertices on the other one. The polylines that represent, for instance, streams or roads in various data sources don’t generally fulfil this condition. Therefore, a different concept for the determination of the similarity measure of two polyline objects is necessary.

The similarity of two polylines with different number of vertices, or with vertices that cannot be assigned to each other, can be determined by the Hausdorff distance and the Fréchete distance.

The Hausdorff distance $\delta_H(A, B)$ between two sets A and B is defined as follows (Hausdor, 1914):

$$\delta_H(A,B) = \max(\sup_{a \in A} \inf_{b \in B} d(a,b), \sup_{b \in B} \inf_{a \in A} d(a,b)), \tag{1}$$

where sup represents the supremum, inf the infimum, and d is the underlying metric in the plane, e.g. the Euclidean distance. The Fréchet distance δ_F between the curves $P;Q : [0; 1] \rightarrow R^n$ is defined as follows (Ewing, 1985):

$$\delta_F(P,Q) = \inf_{\alpha,\beta} \max_{t \in [0,1]} \left\| P(\alpha(t)) - Q(\beta(t)) \right\|, \tag{2}$$

where α,β: $[0,1] \rightarrow [0,1]$ range over the class of all continuous and monotone increasing functions.

Alt and Godau (1995) described at p. 76 the Fréchet distance intuitively: 'Suppose a man is walking his dog, he is walking on the curve, the dog on the other. Both are allowed to control their speed but are not allowed to go backwards. Then the Fréchet distance of the curves is the minimal length of a leash that is necessary.'

However, neither of these methods considers the uncertainty of both polylines and they are both relatively computationally intensive. The uncertainty of objects can be expressed using fuzzy sets $\overline{A}$ and $\overline{B}$ (Fig. 1). To consider the uncertainty and similarity of arbitrary polyline objects a and b simultaneously, we employ the fuzzy set theory and compute the similarity of two fuzzy sets.

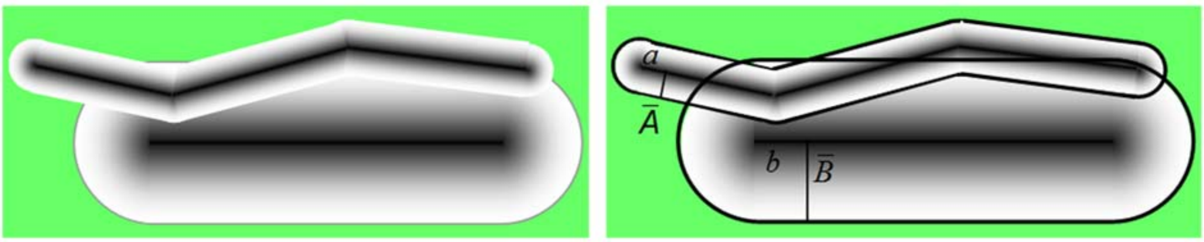

Fig. 1. The polyline objects with their areas of uncertainty (left), modelling of areas of uncertainty by fuzzy sets (right)

2.2. Uncertainty modelling with fuzzy sets

The concept of a fuzzy set was introduced by Zadeh (1965), p. 339: 'Let X be a space of points (objects), with a generic element of X denoted by x. Thus, $X = \{x\}$. A fuzzy set A in X is characterized by a membership (characteristic) function $\mu_A(x)$ which associates with each point in X a real number in the interval [0,1], with the value of $\mu_A(x)$ at x representing the 'grade of membership' of x in A.'

The shape and the parameters of the membership functions can be, in general, determined empirically or based on known properties of the analysed phenomenon. The special cases of piecewise linear membership functions, which could be used to represent the uncertainty of linear spatial objects, are shown in Figure 2.

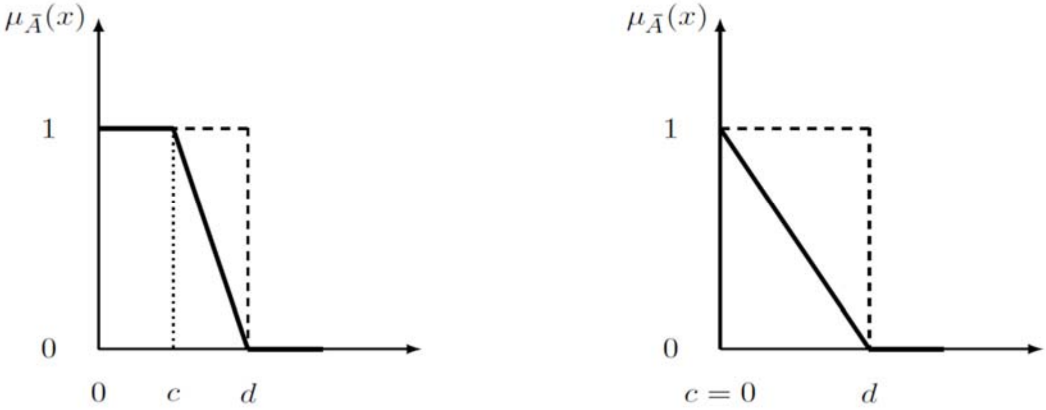

Fig. 2. Linear membership functions of fuzzy sets: L-function (left), triangular function (right)

The L-membership function (Fig. 2 left) is defined by:

$$\mu_A(x) = \begin{cases} 1 & \text{if} \quad x < c, \\ \dfrac{d-x}{d-c} & \text{if} \quad c \le x \le d, \\ 0 & \text{if} \quad x > d \end{cases} \tag{3}$$

and the triangular membership function (Fig. 2 right), for c = 0, by:

$$\mu_A(x) = \begin{cases} 1 & \text{if} \quad x < 0, \\ \dfrac{d-x}{d} & \text{if} \quad 0 \le x \le d \\ 0 & \text{if} \quad x > d \end{cases} . \tag{4}$$

The triangular or the L-membership function can also be seen as a simple approximation of the Gaussian membership function (Fig. 3), which is a more reliable description of the uncertainty of a spatial object:

$$\mu_A(x) = e^{\frac{-(x-c)^2}{2\sigma^2}}, \tag{5}$$

where σ^2 is the variance and c the expected value.

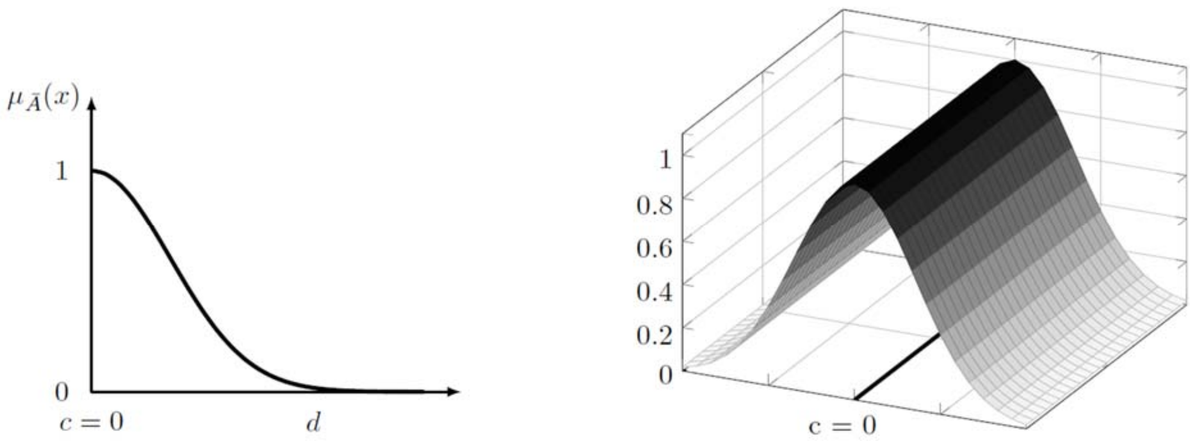

Fig. 3. 3D representation of the graph of the Gaussian membership function (left) and the uncertainty of a line represented by the Gaussian function (right)

2.3. Identity and similarity of two fuzzy sets

Fuzzy sets are considered to be a generalization of the classical set theory. In a similar way as we can generalize the characteristic function $\chi_A : X \to \{0,1\}$ to the membership function $\mu_{\overline{A}} : X \to [0,1]$, it is possible to generalize the identity, similarity, and inclusion of two sets to fuzzy identity, fuzzy similarity, and fuzzy inclusion of two fuzzy sets.

According to Bandemer (2006), two classical (crisp) sets A and B are identical if $A = A \cap B = B$ or, in equivalent form, if $A \subseteq B \wedge B \subseteq A$. The idea of similarity can then be based on the conception of two sets being similar if they are approximately equal. The indefinite term *approximately equal* is based on the conception that a set of objects outside their intersection $A \cap B$ is small when compared to their union $A \cup B$. To define the term small set, we need to define an appropriate measure for the size of the set. If the set is finite, the size of the set is the number of elements, while an integral of the set is used for infinite sets. In the case of an unknown character of the set, the appropriate measure is a generalization of the number of the elements for infinite sets called cardinality (card). The similarity $sim(A,B)$ of two sets A and B (also known as the Jaccard similarity coefficient (Jaccard, 1901)) can then be defined by (Bandemer, 2006):

$$sim(A,B) = \frac{card(A \cap B)}{card(A \cup B)}. \tag{6}$$

Let $\overline{A}$ and $\overline{B}$ be fuzzy sets in X. Then

$$sim(\overline{A},\overline{B}) = \frac{card(\overline{A} \cap \overline{B})}{card(\overline{A} \cup \overline{B})}. \tag{7}$$

For a finite fuzzy set $\overline{A}$, the cardinality $card(\overline{A})$ is defined as the sum of the membership degrees of a fuzzy set (Dhar, 2013):

$$card(\overline{A}) = \left|\overline{A}\right| = \sum_{x \in X} \mu_A(x). \tag{8}$$

Using Zadeh's definitions of fuzzy intersection and union (Zadeh, 1965):

$$\mu_{\overline{A} \cap \overline{B}} = \min(\mu_A(x), \mu_B(x)) \tag{9}$$

$$\mu_{\overline{A} \cup \overline{B}} = \max(\mu_A(x), \mu_B(x)) \tag{10}$$

we define the similarity measure of two fuzzy sets as:

$$sim(\overline{A},\overline{B}) = \frac{\sum_{x \in X} \min(\mu_{\overline{A}}(x), \mu_{\overline{B}}(x))}{\sum_{x \in X} \max(\mu_{\overline{A}}(x), \mu_{\overline{B}}(x))}. \tag{11}$$

This concept corresponds to the Jaccard similarity coefficient, but it can be applied to other similarity coefficients, e.g., Sørensen index (Sørensen, 1948), or Dice's coefficient (Dice, 1945) in a similar way.

2.4. Fuzzy subsethood and inclusion measure

Subsethood, as described by Young (1996), is an important concept in the area of fuzzy sets. Fuzzy subsethood allows one fuzzy set to contain other one to some degree between 0 and 1; it is a fuzzification of Zadeh's fuzzy set containment which is a crisp characteristic. If A and B are sets and every element of A is also an element of B, then A is a subset of (or *is included in*) B, which is denoted as $A \subseteq B$. An inclusion measure for the crisp sets is defined as:

$$\begin{aligned} & A \subseteq B \Rightarrow I(A,B) = 1 \\ & \neg(A \subseteq B) \Rightarrow I(A,B) = 0 \end{aligned} \tag{12}$$

with

$$A \subset B \Leftrightarrow |A \cap B| = |A| . \tag{13}$$

In fuzzy set theory, inclusion measure indicates the degree to which a given fuzzy set $\overline{A}$ is contained in another fuzzy set $\overline{B}$. For fuzzy sets $\overline{A}$ and $\overline{B}$ we consider:

$$\overline{A} \subset \overline{B} \Leftrightarrow |\overline{A} \cap \overline{B}| = |\overline{A}| . \tag{14}$$

Then, the degree of inclusion (the inclusion measure) for fuzzy sets $\overline{A}$ and $\overline{B}$ can be defined by:

$$I(\overline{A},\overline{B}) = \begin{cases} \dfrac{|\overline{A} \cap \overline{B}|}{|\overline{A}|}, & |\overline{A}| \neq 0 \\ 1, & |\overline{A}| = 0 \end{cases} . \tag{15}$$

Using Zadeh's definition of fuzzy intersection (8) and fuzzy union (9), an inclusion (or subsethood) measure is given by (Zeng and Li, 2006):

$$I(\overline{A},\overline{B}) = \begin{cases} \dfrac{\sum\limits_{x \in X} \min(\mu_{\overline{A}}(x), \mu_{\overline{B}}(x))}{\sum\limits_{x \in X} \mu_{\overline{A}}(x)}, & \overline{A} \neq \varnothing \\ 1, & \overline{A} = \varnothing \end{cases} . \tag{16}$$

Fuzzy inclusion can also be defined as a type of topological spatial relationship of two fuzzy spatial objects (Schneider, 2008; Tang et al., 2006).

2.5. New approach to the identification of uncertain spatial objects in GIS

Based on the above findings, we propose a new method for the spatial identification of objects from two datasets. This method takes into account the uncertainty of input data and also simplifies the conventional approach.

The area of uncertainty of vector data is modelled using raster data. Therefore, the raster of uncertainty is a finite fuzzy set. The proposed algorithm for computing the similarity and the inclusion measure of polyline objects is described by Activity

Diagram in Unified Modeling Language (UML) (Fig. 4). The advantage is that it can be used to compute similarity and inclusion measures using standard tools implemented in the GIS environment, such as *Buffer*, *Intersection*, *Union*, and so on.

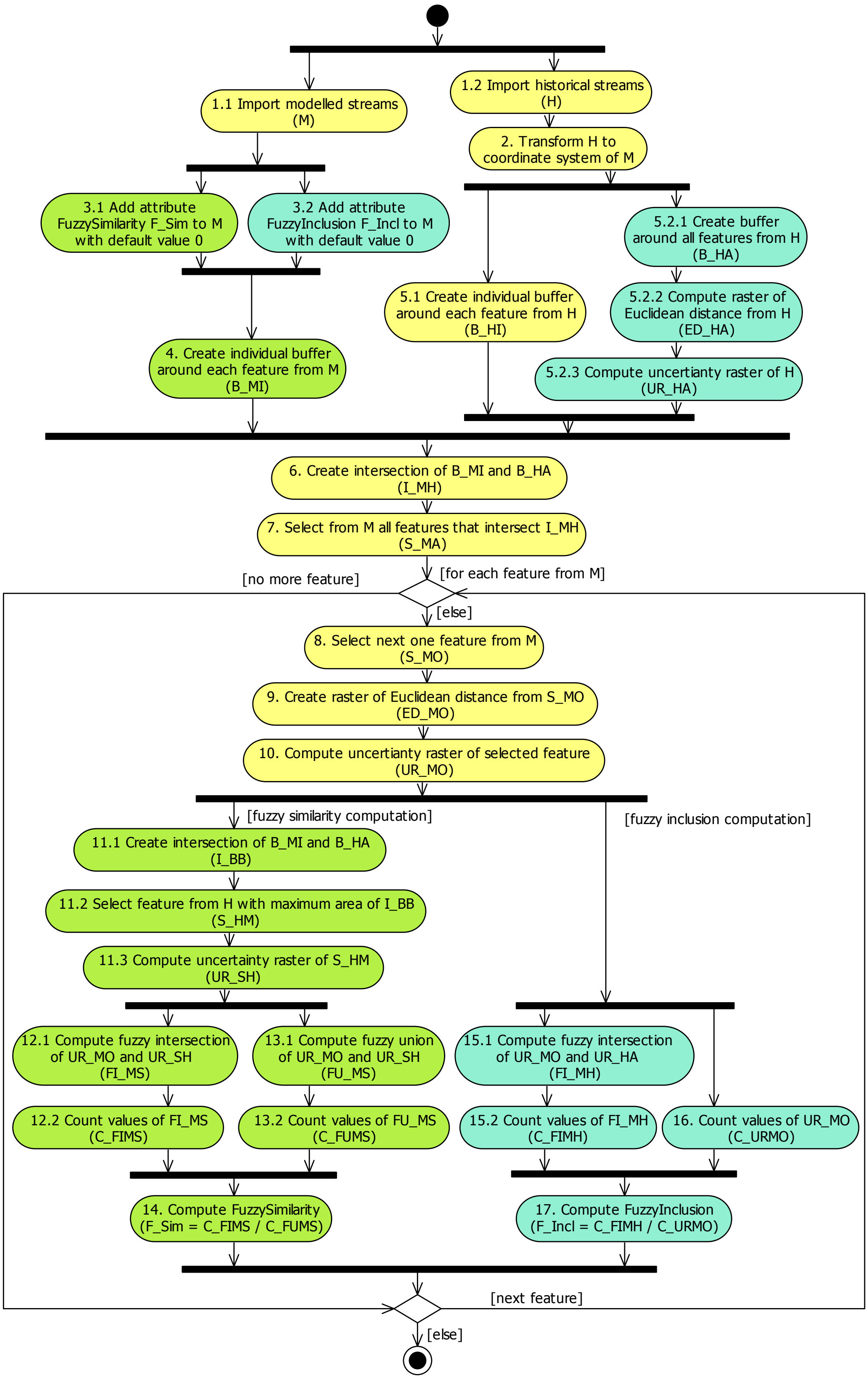

Fig. 4. Activity Diagram of Fuzzy Inclusion and Fuzzy Similarity computation

In general, the proposed approach can be used not only to identify uncertain polylines but also points and polygons. The similarity measure is appropriate for points, simple polygons and two polyline objects with matching segments (the same number of vertices is not necessary). The inclusion measure should be used if it is not possible to determine the matching segments in the polyline datasets. This can be applied to smooth line objects such as roads or streams, which is the most complicated situation for this task. That is why we are dealing with it in the case study.

3. Case study

The accessibility to the water sources, especially to the surface streams, is one of the most important factors in prediction of ancient settlements locations (Bátora and Tóth, 2014; Ford et al., 2009; Van Leusen et al., 2005; Lieskovský, 2011); the location of the streams themselves is important for archaeohydrological modelling of floods (Arnaud-Fassetta et al., 2010; Gillings, 1995), irrigation (Harrower, 2009; Harrower, 2010), or land-use potential (Bolten et al., 2006). Since current maps and datasets represent water sources in their present form (the vast majority of the streams in Europe were regulated and modified in 20th century), their usage in spatial analysis in archaeology is limited. Archaeological predictive modelling often requires information about the location of the streams and the water sources in the past. In some cases, historical maps can be used, but they are often not detailed enough and the hand-drawn vectorization combined with complicated georeferencing leads to limited precision. The positions of the streams could also be derived from the landscape morphology using a digital terrain model (DTM) and applying M8, MD, or MD8 algorithm, see e.g. (Wilson et al., 2008), which are commonly implemented in GIS software packages. Although the (natural) stream network can be derived from contemporary DTM, these results may not represent the streams in the original position in the past due to the landscape changes. Another approach to streams modelling based on the DTM is described e.g. in (Pilesjö and Hasan, 2014). These methods create possible streams; to obtain the real streams, the result needs to be compared to a map, an ortophoto, or other dataset. In our case study, we compare the modelled streams with the historical maps from the Second Military Survey of the Habsburg Empire. Therefore, using the approach described in previous chapter, we attempt to preserve the precision of the streams derived from the DTM with the correctness provided by historical maps. We use the inclusion measure to determine the segments of modelled streams that: i) can be considered as correct when compared to the historical maps, and ii) should be excluded from the subsequent analysis.

Our study area is located in the southern part of the central Slovakia (Fig. 5). The historical stream dataset was created by vectorization of the maps from the Second Military Survey of Habsburg Empire, which were georeferenced using identical points. This survey was performed in the 19th century (1806 - 1869), and it is the best available information about the streams before the regulation and modification of streams started in the 20th century. The First Military Survey maps (1764 - 1783) are only suitable for the verification of the stream presence and its topological characteristics, but cannot be used—considering the accuracy of the methods of creation—as a source of information about geometry, or the localization of streams (Fig. 6).

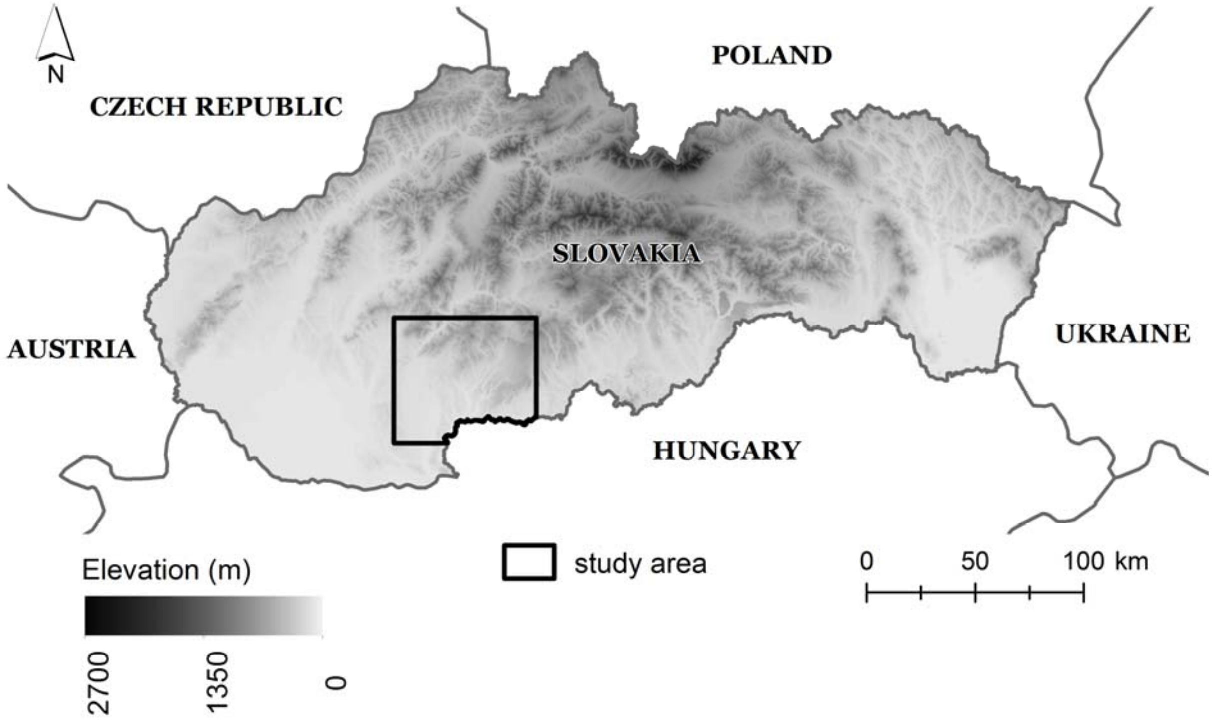

Fig. 5. The study area.

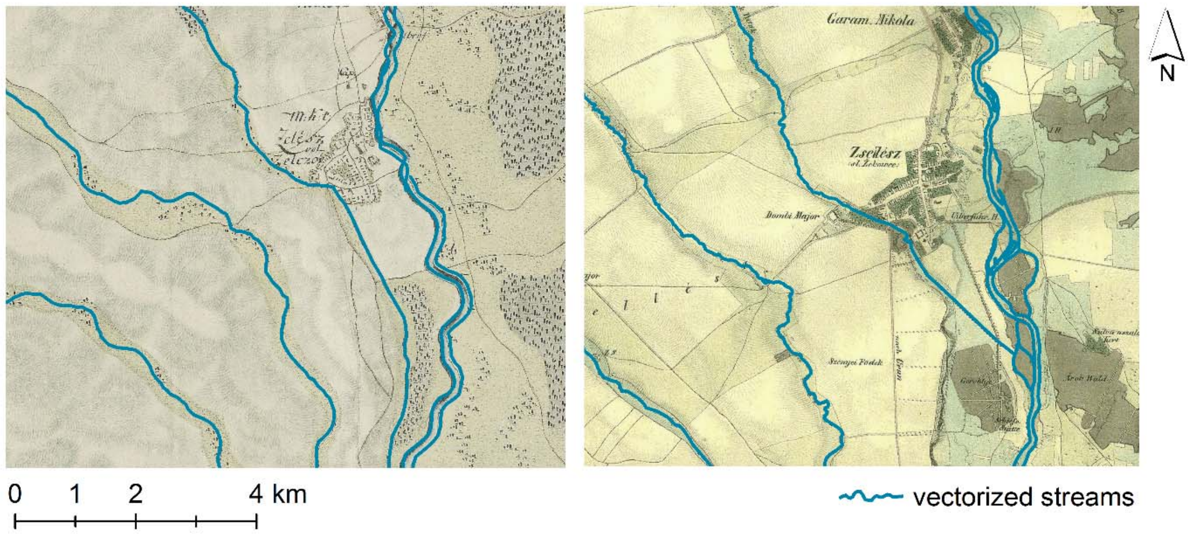

Fig. 6. Streams depicted on two historical maps: First Military Survey of the Habsburg Empire (1764 -1783) (left), b) Second Military Survey of the Habsburg Empire (1806 -1869) (right).

The historical stream dataset consists of 216 polyline segments with total length of 273.8 km. The zone of uncertainty for historical streams was set using the Gaussian membership function (Equation 5) with the expected value of uncertainty c_h = 0 m and the standard deviation σ_h = 140 m. These values we estimated from the residuals on 406 identical points (243 churches, 81 road crossings, 52 bridges, 26 stream confluences, 4 corners of buildings) depicted on the historical maps from the Second Military Survey of the Habsburg Empire. For identification, the current ZB*GIS*® database (the fundamental spatial database for GIS in Slovakia) we used.

We derived the modelled stream network dataset from the digital terrain model DMR-3.5 (a national DTM) with 10 m resolution using ArcGIS 10.2 Spatial Analyst's Hydrology toolset. The channel initiation threshold value was set to 2000; this value provided appropriate detail of the stream network with respect to the historical dataset. For the modelled (potential) streams, we also used the Gaussian membership function with the expected value c_m = 0 m and the standard deviation σ_m = 25 m. We excluded the modelled stream segments located outside the uncertainty zone of historical streams ($\mu_{\tilde{h}}(x)$ = 0), since their fuzzy inclusion measure would be equal to 0 (Equation 16). Then, there were 445 modelled stream segments left for the computation, with the total length of 320.9 km.

The historical and the modelled streams and their areas of uncertainty are shown in Figure 7.

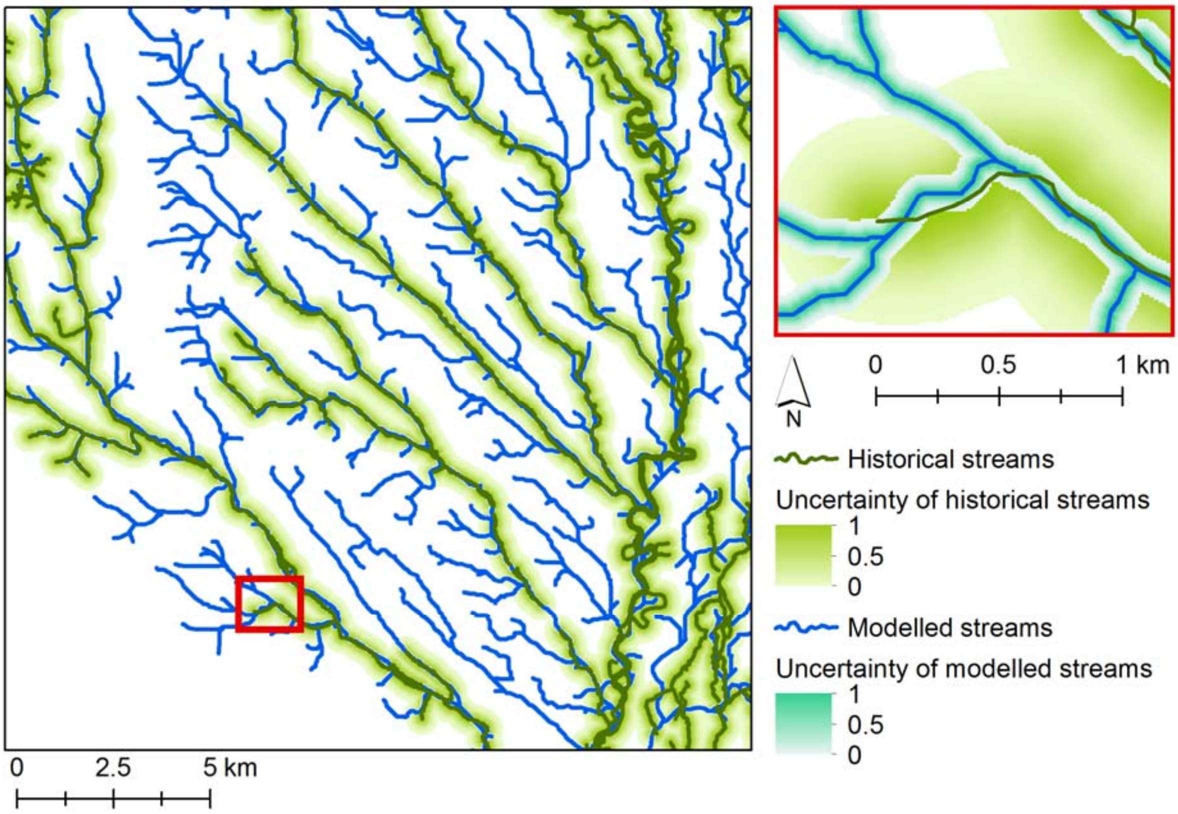

Fig. 7. Modelled and historical streams with their respective areas of uncertainty

4. Results

The computed inclusion values can help to determine, which of the modelled streams match the streams depicted on the historical maps (inclusion measure values close to 1), and which of them were probably too small or seasonal to be depicted on the maps or maybe not real at all (inclusion measure values close to 0). In Figure 8, we show the modelled streams with the values of inclusion measure.

With the computed inclusion measure, stored as numeric attribute of the modelled streams, we can rank the stream segments according to their correctness and set the criteria for the decision making process. These criteria may vary with different types of the landscapes and analytical purposes. In our case, we decided to consider objects with fuzzy inclusion measure higher than 0.8 to be accepted, and those with the value below 0.2 to be rejected; the rest of the objects need to be considered individually. In our case, more than a half of the objects can automatically be taken as correct (fuzzy inclusion measure larger than 0.80), and 11.5 % can be excluded from the dataset (Table 1).

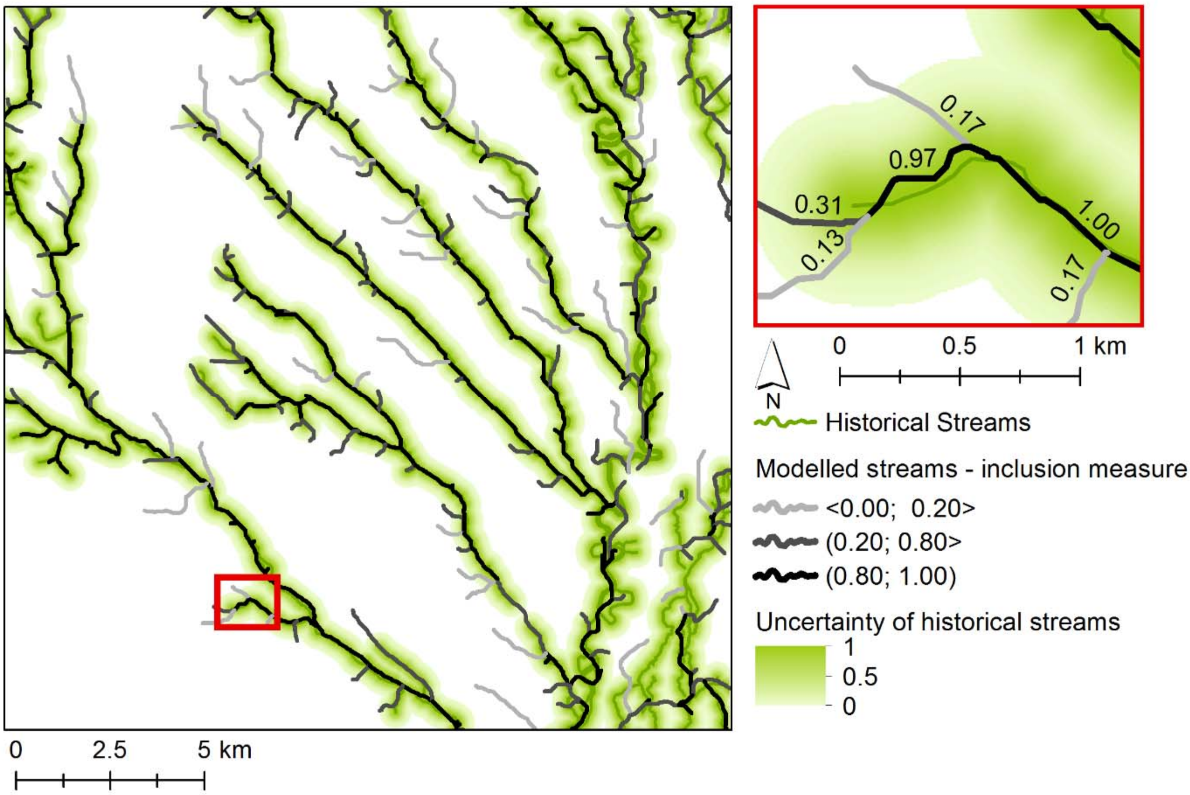

Fig. 8. Modelled streams with the value of inclusion measure

Table 1. Statistics of the fuzzy inclusion measure (FI) computed for the modelled stream segments

FI interval	[0.00;0,20]	[0.20;0,40]	[0.40;0,60]	[0.60;0,80]	[0.80;1,00]	Sum
# of feat.	51	46	47	41	260	445
% from total	11.5 %	10.3 %	10.6 %	9.2 %	58.4 %	100 %

The implications drawn from the inclusion measure can differ for various applications: the goal can be to exclude the incorrect segments or to extract only those with a higher degree of inclusion. Depending on the topography, the correct modelled stream segments can enter the archaeological analysis directly, or they can be used to update the matching segments of the historical streams, leaving parts that are not similar unchanged. In Figure 9, we show examples of dealing with the undetermined values $I \in (0.20; 0.80)$. There are three cases of segments with the inclusion value about 0.40, which are completely different. On the right, the segment does not correspond to the historical stream. In the middle, the segment looks partially correct, so we can preserve its matching part. While these two cases can represent smaller or occasional streams (not depicted on the maps), the modelled stream on the left is obviously wrong and cannot be used to improve the information about historical streams.

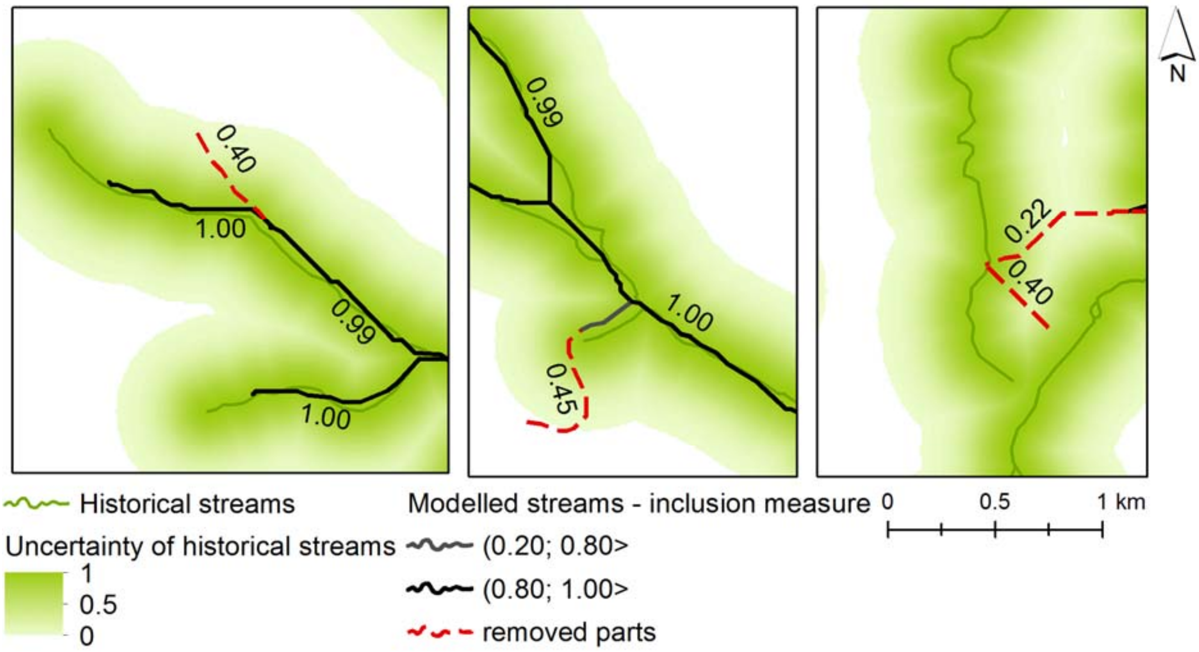

Fig. 9. Examples of updating possibilities. Three matching segments can be updated, the whole segment with I = 0.40 is excluded (right); only a part of the segment with I = 0.45 is preserved (middle); the update of is not possible (left)

In our example, the total number of segments of the modelled streams was about two times larger than the number of segments of the historical streams, with the total length of the modelled stream segments being about 1.2 larger. This is due to the fact that we usually depict smooth rivers and streams on maps, but the result of hydrological modelling usually contains also small segments of inflowing streams. These parts were not depicted on maps not because they were not present, but probably because of the cartographic generalization. From an archaeological point of view, it should be noted that the historical streams depicted on the Second Military survey maps represent the Middle ages and later time periods at the best. For more ancient time periods, the data related to the historical landscape morphology are unavailable; the difference between historical and modern morphology can be significant and in most cases it is not possible to use a modern DTM without considering the changes of the terrain. For this purpose, it is possible to use geomorphological, stratigraphic, geoarchaeological, and geobotanical data, which provide us with information about the changes in the landscape during long time periods (Arnaud-Fassetta et al., 2010; Lieskovský, 2011; Periman, 2005).

5. Conclusions and discussion

For a human, it is easy to determine the similarity of two objects. But, when working with large datasets, e.g. comparing data from different time periods or updating an existing dataset, it might be very time consuming to look at each object and to decide whether it is similar or not. For a computer, this task is more complicated and requires a mathematical definition, but once the algorithm is defined, it can be performed automatically with reduced need for manual control. In this paper, we have proposed an approach that uses the fuzzy inclusion or fuzzy similarity measure of two sets, use of which depends on whether we are comparing points, polylines, or polygons. This process considers the uncertainty of the geometric representation of spatial objects, which is their inherent property.

The advantage of the proposed algorithm is the possibility to effectively identify objects. We can select the objects that are similar or identical and use them in the analysis for updating, or select the dissimilar objects and exclude them from the dataset. This algorithm considers both the similarity of the shape and the similarity of the location of spatial objects. Moreover, the whole process of the determination of similarity or inclusion measure is enriched by modelling spatial objects' uncertainty using fuzzy sets. If needed, it is also possible to consider both the geometric representation and the values of chosen attributes of objects at the same time in the process. In this case, it is important to choose the appropriate measure of similarity, or aggregation of multiple similarity measures, whose determination depends on particular application.

Our algorithm consists of steps that can be performed using standard tools in GIS environment. We have applied this approach to compare two stream datasets: streams vectorised from the historical maps and streams derived from a modern digital terrain model. This task is useful, but is not limited to, for hydrological and predictive modelling in archaeology. We have computed the inclusion measure and used it to select the segments that i) can be considered as correct automatically ($I > 0.80$, 58.4 % of segments), ii) require individual decision ($0.20 < I \leq 0.80$, 30.1 %) and iii) are automatically rejected ($I \leq 0.20$, 11.5 %).

An open question is the setting of similarity or inclusion values' thresholds to determine particular segments or objects that should be (automatically) considered as similar or dissimilar. Strict thresholds lead to more reliable results, but greater need for manual processing (visual control and manual editing of some segments), which can be time consuming for large datasets. To make this process more effective, other means of automatic control can be employed depending on the purpose of the analysis, e.g. topological rules valid for stream datasets. Also, the parameters used to model uncertainty (i.e. the parameters of Gauss membership function) can be an issue. If the quality or source of input datasets is unknown, it might be difficult to estimate them correctly. Overall, the approach proposed in this paper enables to consider the uncertainty of both input datasets. Therefore, the output from this method, which then enters the decision-making process, is enriched by additional information on the quality of the data set. This is the main benefit of this method when compared to using only Euclidean distance or visual determination of the similarity of two uncertain spatial objects. This leads to a more reliable procedure that can be, at least to some extent, automatized.

Acknowledgments
The digital elevation model DMR-3.5 and the ZB*GIS*® database were provided by the Geodetic and Cartographic Institute Bratislava. This work was supported by the Grant No. 1/0682/16 of the Grant Agency of Slovak Republic VEGA.

References

Alt, H. & Godau, M. (1995). Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 5(1-2), 75-91. doi: 10.1142/S0218195995000064.

Arnaud-Fassetta, G., Carcaud, N., Castanet, C. & Salvador, P.-G. (2010). Fluviatile palaeoenvironments in archaeological context: Geographical position, methodological approach and global change – Hydrological risk issues. *Quaternary International*, 216(1-2), 93-117. doi: 10.1016/j.quaint.2009.03.009.

Bandemer, H.-W. (2006). *Mathematics of Uncertainty - Ideas, Methods, Application Problems.* Springer. doi: 10.1007/3-540-31228-5.

Bátora, J. & Tóth, P. (2014). Settlement Strategies in the Early Bronze Age in South-Western Slovakia. In: Kienlin, T., Valde-Nowak, P., Korczynska, M., Cappenberg, K. & Ociepka, J. (eds.) *Settlement, Communication and Exchange around the Western Carpathians*, Archaeopress, 325-340.

Bolten, A., Bubenzer, O. & Darius, F. (2006). A digital elevation model as a base for the reconstruction of Holocene land-use potential in arid regions. *Geoarchaeology*, 21(7), 751-762. doi: 10.1002/gea.20137.

Chen, C.-C. & Knoblock, C. A. (2008). Conflation of Geospatial Data. In: *Encyclopedia of GIS*, S. Shekar and H. Xiong, Eds. Springer US, Boston, ch. Conflation, 133-140. doi: 10.1007/978-0-387-35973-1_182.

Cobb, M. A., Chung, M. J., Foley III, H., Petry, F. E., Shaw, K. B. & Miller, H. V. (1998). No Title. *Geoinformatica*, 2(1), 7-35. doi: 10.1023/A:1009788905049.

Dhar M. (2013). Cardinality of Fuzzy Sets: An Overview. *International Journal of Energy, Information and Communications*, 1, 15-22.

Dice, L. R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3), 297-302. doi: 10.2307/1932409.

Ewing, G. M. (1985). *Calculus of Variations with Applications*. Dover Publications, New York.

Ford, A., Clarke, K. C. & Raines, G. (2009). Modeling Settlement Patterns of the Late Classic Maya Civilization with Bayesian Methods and Geographic Information Systems. *Annals of the Association of American Geographers*, 99(3), 1-25. doi: 10.1080/00045600902931785.

Gillings, M. (1995). GIS and the Tisza Flood-Plain: Landscape and Settlement Evolution in North-Eastern Hungary. In: *The Impact of Geographic Information Systems on Archaeology: a European Perspective*, G. Lock and Z. Stancic, Eds. Taylor & Francis, New York, 67-84.

Goodchild, M. F. & Hunter, G. J. (1997) A simple positional accuracy measure for linear features. *International Journal of Geographical Information Science* 11(3), 299-306. doi: 10.1080/136588197242419.

Harrower, M. J. (2009). Is the hydraulic hypothesis dead yet? Irrigation and social change in ancient Yemen. *World Archaeology*, 41(1), 58-72. doi: 10.1080/00438240802668354.

Harrower, M. J. (2010). Geographic Information Systems (GIS) hydrological modeling in archaeology: an example from the origins of irrigation in Southwest Arabia (Yemen). *Journal of Archaeological Science*, 37(7), 1447-1452. doi: 10.1016/j.jas.2010.01.004.

Hausdorff, F. (1914). *Grundzüge der Mengenlehre*. Veit, Leipzig.

Heuvelink, G. B. M. & Brown, J. D. (2016). Uncertain Environmental Variables in GIS. In: *Encyclopedia of GIS*, S. Shekar and H. Xiong, Eds. Springer International Publishing, ch. Uncertain, 1-9. doi: 10.1007/978-3-319-23519-6_1422-2.

Jaccard, P. (1901). Étude comparative de la distribution orale dans une portion des Alpes et des Jura. *Bulletin de la Societe Vaudoise des Sciences Naturelles*, 37, 547-579.

Koukoletsos, T., Haklay, M. & Ellul, C. (2012). Assessing data completeness of VGI through an automated matching procedure for linear data. *Transactions in GIS*, 16(4), 477-498. Doi: 10.1111/j.1467-9671.2012.01304.x

Lieskovský, T. (2011). Využitie geografických informačných systémov v predikčnom modelovaní v archeológii (Unpublished doctoral dissertation), Slovak University of Technology.

Longley, P. A., Goodchild, M. F., Maguire, D. J. & Rhind, D. W. (1999). *Geographical Information Systems and Science*. John Willey &Sons.

Periman, R. D. (2005). Modeling landscapes and past vegetation patterns of New Mexico's Rio del Oso Valley. *Geoarchaeology*, 20(2), 193-210. doi: 10.1002/gea.20043.

Petry, F. E., Robinson, V. B. & Cobb, M. A.(2005). *Fuzzy Modeling with Spatial Information for Geographic Problems*. Springer Berlin Heidelberg, Berlin, Heidelberg. doi: 10.1007/b138243.

Pilesjö, P. & Hasan, A. (2014) A Triangular Form-based Multiple Flow Algorithm to Estimate Overland Flow Distribution and Accumulation on a Digital Elevation Model. Transactions in GIS 18, 108-124. doi:10.1111/tgis.12015

Samal, A., Seth, S. & Cueto, K. (2004). A feature-based approach to conflation of geospatial sources. *International Journal of Geographical Information Science*, 18(5), 459-489. doi: 10.1080/13658810410001658076.

Schneider, M. (2008). Fuzzy Spatial Data Types for Spatial Uncertainty Management in Databases. In: *Handbook of Research on Fuzzy Information Processing in Databases*. IGI Global, 490-515. doi: 10.4018/978-1-59904-853-6.ch019.

Seth, S. & Samal, A. (2016). Conflation of Features. In: *Encyclopedia of GIS*, S. Shekar and H. Xiong, Eds. Springer International Publishing, 1-7. doi: 10.1007/978-3-319-23519-6_181-2.

Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab*, 5(4), 1-34.

Tang, X., Fang, Y. & Kainz, W. (2006). Fuzzy Topological Relations between Fuzzy Spatial Objects. In: *Fuzzy Systems and Knowledge Discovery, Proceedings, vol. 4223*, 324-333. doi: 10.1007/11881599_37.

Toomanian, A., Harrie, L., Mansourian, A. & Pilesjö, P. (2013). Automatic integration of spatial data in viewing services. J*ournal of Spatial Information Science*, 6, 43-58. doi: 10.5311/JOSIS.2013.6.87.

Van Leusen, M., Van Leusen, M., Deeben, J., Deeben, J., Hallewas, D., Hallewas, D., Kamermans, H., Kamermans, H., Verhagen, P., Verhagen, P., Zoetbrood, P. & Zoetbrood, P. (2005). A Baseline for Predictive Modelling in the Netherlands. Predictive Modelling for Archaeological Heritage Managment: A research agenda, Amersfoort, 25-92.

Walter, V. & Fritsch, D. (1999). Matching spatial data sets: a statistical approach. *International Journal of Geographical Information Science*, 13(5), 445-473. doi: 10.1080/136588199241157

Wen-June, W. (1997). New similarity measures on fuzzy sets and on elements. *Fuzzy Sets and Systems*, 85(3), 305-309. doi: 10.1016/0165-0114(95)00365-7.

Wilson, J. P., Aggett, G., Yongxin, D. & Lam, C. S. (2008). Water in the Landscape: A Review of Contemporary Flow Routing Algorithms. In: *Advances in Digital Terrain Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg, 213-236. doi: 10.1007/978-3-540-77800-4_12.

Wygralak, M. (1983). Fuzzy inclusion and fuzzy equality of two fuzzy subsets, fuzzy operations for fuzzy subsets. *Fuzzy Sets and Systems*, 10(1-3), 157-168. doi: 10.1016/S0165-0114(83)80112-2.

Young, V. R. (1996). Fuzzy subsethood. *Fuzzy Sets and Systems*, 77(3), 371-384. doi: 10.1016/0165-0114(95)00045-3.
Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. doi: 10.1016/S0019-9958(65)90241-X.
Zeng, W. & Li, H. (2006). Inclusion measures, similarity measures, and the fuzziness of fuzzy sets and their relations. *International Journal of Intelligent Systems*, 21, 639-653. doi: 10.1002/int.20152.
Zhang, J. & Goodchild, M. F. (2002). *Uncertainty in Geographical Information*. Taylor & Francis. doi: 10.1111/j.1467-8306.2003.09304014_8.x.

Authors:

Ing. Renata Ďuračiová, PhD. [1)], renata.duraciova@stuba.sk
Ing. Alexandra Rášová [1)], alexandra.rasova@stuba.sk
Ing. Tibor Lieskovský, PhD. [1)], tibor.lieskovsky@stuba.sk
[1)] Department of Theoretical Geodesy,
Faculty of Civil Engineering,
Slovak University of Technology in Bratislava
Radlinského 11, 810 05, Bratislava, Slovak Republic