Research Article
Fuzzy Reliability in Spatial Databases

Ferdinando Di Martino[1] and Salvatore Sessa[1,2]

[1] *Università degli Studi di Napoli Federico II, Dipartimento di Architettura, Via Monteoliveto 3, 80134 Napoli, Italy*
[2] *Università degli Studi di Napoli Federico II, Centro Interdipartimentale per l'Analisi e la Progettazione Urbana Luigi Pisciotti, Via Toledo 402, 80134 Napoli, Italy*

Correspondence should be addressed to Salvatore Sessa; sessa@unina.it

Today it is very difficult to evaluate the quality of spatial databases, mainly for the heterogeneity of input data. We define a fuzzy process for evaluating the reliability of a spatial database: the area of study is partitioned in isoreliable zones, defined as homogeneous zones in terms of data quality and environmental characteristics. We model a spatial database in thematic datasets; each thematic dataset concerns a specific spatial domain and includes a set of layers. We estimate the reliability of each thematic dataset and therefore the overall reliability of the spatial database. We have tested this method on the spatial dataset of the town of Cava de' Tirreni (Italy).

1. Introduction

Fuzzy rule-based models are applied in geographical information systems (GIS) [1–3] and we use our previous approach [4, 5] for estimating the reliability of spatial databases. There the concept of geodata "reliability" was introduced as a fuzzy measure of the quality of geodata, based on the analysis of uncertainty and quality of the data. Strictly speaking, in [5] the authors implement a tool called (Fuzzy Spatial Reliability Analysis) Fuzzy-SRA [6] for studying the reliability of the intrinsic vulnerability of aquifers by utilizing the DRASTIC model, encapsulated in a GIS; in [4] Fuzzy-SRA is used for estimating the reliability of the aerophotogrammetric set of geographic layers of the island of Procida (near Naples, Italy) and in [7] Fuzzy-SRA is applied in a GIS tool for implementing a fuzzy rule-based system for analyzing the eruption risk of the famous vulcan Vesuvius.

As the first step, we need to divide the geographic area of study in isoreliable zones, that is, in zones having (quasi) homogeneous data quality and geographical characteristics. An expert sets the characteristics related to the quality of each layer (e.g., the percent of uncoded spot elevation features). Each characteristic, called "parameter," is a measurable entity

that could affect the quality of the dataset. After calculating the value of a parameter, a fuzzification process is applied for estimating the quality of the set of layers, where each fuzzy set is given by a triangular fuzzy number (TFN), which in turn is identified from a linguistic label. In other words, an isoreliable zone is a subarea of the area of study in which the quality of the geodata is homogeneous; that is, the values of the parameters are similar and with the same geographical characteristics (e.g., a flat country).

The expert creates a fuzzy partition, labelling the TFNs with linguistic labels (say) $\alpha_1, \alpha_2, \ldots, \alpha_n$ (see, e.g., Table 3) for each parameter. An isoreliable zone is associated with the linguistic label of the corresponding TFN, for which the membership degree of the parameter is the highest one. This process is iterated for each parameter of each layer. The elements of Fuzzy-SRA, associated with each parameter, are fuzzy attributes represented by a string. To clarify how these strings are composed, now we suppose, as an example, that we have partitioned the area of study in 5 isoreliable zones, $O1, \ldots, O5$; hence we create a fuzzy partition of the domain discourse of the parameter A in 6 fuzzy sets with linguistic labels, respectively, $\alpha_1, \alpha_2, \ldots, \alpha_6$. After the

TABLE 1: Example of association of isoreliable zones with linguistic labels.

Isoreliable zone	Linguistic label
O1	α_2
O2	α_1
O3	α_2
O4	α_6
O5	α_4

TABLE 2: Spatial datasets and related attributes used in [4].

Layer	Parameter
Spot elevations (SE)	Density of uncoded SE
	Density of coded SE per ha
Contour lines	Mean density of contour lines for ha/mean slope
	Standard deviation of contour lines for ha/mean slope
Buildings	Mean of the absolute value of perimeter-shape length
	Mean of the absolute value of shape area
Network streets	Mean density of buffer area covered by buildings
	Standard dev. of buffer area covered by buildings

TABLE 3: The TFNs of the linguistic labels.

Label	Description	a	M	b
$\alpha_1 = Cv$	Optimum reliability	0.80	0.90	1.00
$\alpha_2 = V$	Good reliability	0.65	0.75	0.80
$\alpha_3 = Mv$	Sufficient reliability	0.55	0.60	0.65
$\alpha_4 = F$	Mediocre reliability	0.45	0.50	0.55
$\alpha_5 = Sc$	Scanty reliability	0.35	0.40	0.45
$\alpha_6 = Bd$	Bad reliability	0.20	0.30	0.35
$\alpha_7 = Nl$	Null reliability	0.00	0.10	0.20

fuzzification process, we associate each isoreliable zone to a TFN as showed in Table 1.

A string is created for the parameter P_1 in the following form:

$$A = [O4]^{\alpha_6}[-]^{\alpha_5}[O5]^{\alpha_4}[-]^{\alpha_3}[O1, O3]^{\alpha_2}[O2]^{\alpha_1}, \quad (1)$$

where the symbol "$-$" indicates the absence of isoreliable zones to be associated with the corresponding TFN.

Now we suppose to create a partition in four TFNs of the domain of a second attribute P_2, obtaining the corresponding string B. The combination of the two strings A and B made by means of the new operation "Δ" is a string $C = (A\Delta B)$ (details are given in [6] and in Section 2.1) defined as

$$C = (A\Delta B) = [-]^{\gamma_8}[o_4]^{\gamma_7}[-]^{\gamma_6}[o_5]^{\gamma_5}[-]^{\gamma_4}[o1]^{\gamma_3}[o_3]^{\gamma_2}[o_2]^{\gamma_1}. \quad (2)$$

In this string new TFNs, labeled as $\gamma_1, \ldots, \gamma_8$, are obtained as well. Generally, if n (resp., m) is the number of TFNs for the
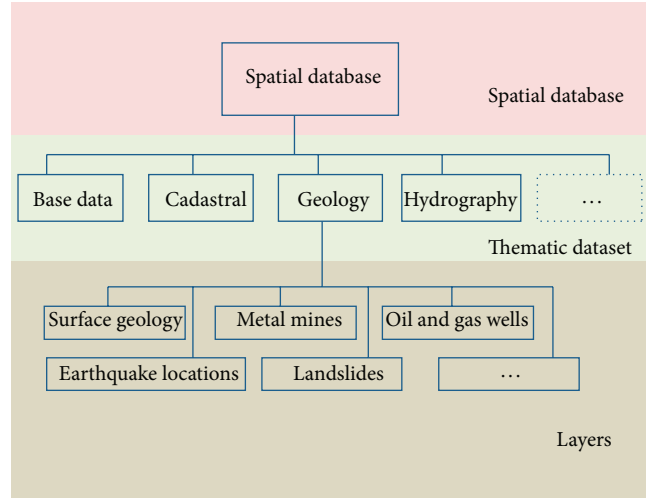

FIGURE 1: Spatial hierarchic database: three levels.

string A (resp., B), then the string C contains $(n+m-1)$ TFNs. The strings obtained for each layer are recombined as in (2) for obtaining a final string used for evaluating the reliability of the whole set of layers. For instance, we can consider (see [4] for details) four aerophotogrammetric layers as given in Table 2.

After calculating the strings for each layer, they are combined again by using operator (2) in order to obtain a final string. In this calculus a weight is associated with each layer, related to the role of that layer in the spatial database. For instance, a layer "buildings" can be more relevant with respect to the layer "infrastructures"; then the quality of the first layer affects the spatial database reliability more than the second one.

There is another question to be considered which consists in the fact that the weight associated with an attribute can change for different isoreliability zones. For instance, the quality of the dataset "spot elevation" affects the reliability of the spatial database in zones with strong slopes more than in smoothed zones. After assigning the weights, the final string is to be recalculated as showed in Section 2.3. The reliability index, to be assigned to each isoreliable zone, is given by the central value of the TFN in which that zone appears in the final string.

In our method we model a spatial database in three hierarchic levels: spatial database, thematic dataset, and layer. A spatial database is composed of several thematic datasets (e.g., geology, hydrology, etc.). Each thematic dataset contains more thematic layers (Figure 1).

Our method starts considering the layer level, assigning the single parameters to each layer. After determining the strings for each layer of a thematic dataset, they are combined with formula (2) for obtaining the final string of the thematic dataset; successively this string is recalculated by considering the weights assigned to each layer. After calculating the final strings of each thematic dataset, we combine these strings by using again formula (2) and we obtain the final string for the

whole spatial database. This string is to be recalculated by considering the weights assigned to each thematic dataset.

For each isoreliable zone, then we obtain the corresponding isoreliability index by taking the central value of the TFN related to that zone in the final string. We test our method considering the spatial database of the town of Cava de' Tirreni, near Salerno (Italy).

In Section 2 we present the algebraic structure given in [6]. Section 3 contains our method, Section 4 gives the results of our tests, and Section 5 is conclusions.

2. Definition of the Algebraic Structure

2.1. The Operations. We recall the main properties of the algebraic structure given in [6]. Let U be the universe of discourse and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ an ordered n-tuple of linguistic labels, each composed from one or more linguistic modifiers and a variable, as, for example, "$\alpha_1 = $ False," "$\alpha_2 = $ More or Less Good," ..., "$\alpha_i = $ Good," "$\alpha_{i+1} = $ Very Good," ..., "$\alpha_n = $ Completely Good," and each represented by suitable TFNs denoted also by α_i, $i = 1, 2, \ldots, n$ (see, e.g., Table 3). Let A be a fuzzy attribute, that is, a map $A : U \rightarrow \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, represented by a string of the following type:

$$A = [a_n]^{\alpha_n} [a_{n-1}]^{\alpha_{n-1}} \cdots [a_1]^{\alpha_1}, \tag{3}$$

where $a_i = A^{-1}(\alpha_i)$ is a subset of U, also called "class" in the sequel. If $A^{-1}(\alpha_i) = \emptyset$, then we write $a_i = [-]$. Let B be another fuzzy attribute represented by following string:

$$B = [b_m]^{\beta_m} [b_{m-1}]^{\beta_{m-1}} \cdots [b_1]^{\beta_1}, \tag{4}$$

where the used symbols have a similar meaning to the above ones. In accordance with [6], we define the operation Δ between A and B by setting

$$C = (A \Delta B) = [c_{m+n-1}]^{\gamma_{m+n-1}} [c_{m+n-2}]^{\gamma_{m+n-2}} \cdots [c_1]^{\gamma_1}, \tag{5}$$

where, by assuming $n \geq m$ without loss of generality, the subsets $\{c_i\}$ are given from the following formulas for $i = 1, \ldots, m+n-1$:

$$c_i = \begin{cases} \bigcup_{j=1,\ldots,i} \left(a_{i-j+1} \cap b_j \right) & \text{if } 1 \leq i \leq m-1, \\ \bigcup_{j=1,\ldots,m} \left(a_{i-j+1} \cap b_j \right) & \text{if } m \leq i \leq n-1, \\ \bigcup_{j=i-n+1,\ldots,m} \left(a_{i-j+1} \cap b_j \right) & \text{if } n \leq i \leq m+n-1. \end{cases} \tag{6}$$

As suggested in [6], the subsets c_i can be calculated by using a simple rule based on the usual arithmetical

multiplication. The TFNs γ_i, for $i = 1, \ldots, m+n-1$, are indeed given by

$$\gamma_i = \begin{cases} \dfrac{1}{k1 + k2} \cdot \displaystyle\sum_{j=1}^{i} d2_j \cdot d1_{i-j+1} \cdot \left(k1 \cdot \alpha_{i-j+1} + k2 \cdot \beta_j \right) \\ \qquad\qquad\qquad \text{if } 1 \leq i \leq m-1, \\ \dfrac{1}{k1 + k2} \cdot \displaystyle\sum_{j=1}^{m} d2_j \cdot d1_{i-j+1} \cdot \left(k1 \cdot \alpha_{i-j+1} + k2 \cdot \beta_j \right) \\ \qquad\qquad\qquad \text{if } m \leq i \leq n-1, \\ \dfrac{1}{k1 + k2} \cdot \displaystyle\sum_{j=i-n+1}^{m} d2_j \cdot d1_{i-j+1} \cdot \left(k1 \cdot \alpha_{i-j+1} + k2 \cdot \beta_j \right) \\ \qquad\qquad\qquad \text{if } n \leq i \leq m+n-1 \end{cases} \tag{7}$$

with the above coefficients d_i, for $i = 1, \ldots, m+n-1$, defined by

$$d_i = \begin{cases} \displaystyle\sum_{j=1}^{i} d2_j \cdot d1_{i-j+1} & \text{if } 1 \leq i \leq m-1, \\ \displaystyle\sum_{j=1}^{m} d2_j \cdot d1_{i-j+1} & \text{if } m \leq i \leq n-1, \\ \displaystyle\sum_{j=i-n+1}^{m} d2_j \cdot d1_{i-j+1} & \text{if } n \leq i \leq m+n-1. \end{cases} \tag{8}$$

The index $d1_i$ (resp., $d2_i$) represents the number of subsets $\{a_i\}$ (resp., $\{b_i\}$) of the string A (resp., B) involved in the operation of union performed to obtain the subsets $\{c_i\}$ of the resulting fuzzy attribute C, whereas the index $k1$ (resp., $k2$) stands for the total number of subsets $\{a_i\}$ of A (resp., $\{b_i\}$ of B) involved in the operation of intersection which gives the subsets $\{c_i\}$ of C.

2.2. The Weights of the Attributes. The first step, which precedes the above mentioned operations over the strings, consists in the determination of the weights of each attribute connected to a fixed zone because they can vary by changing zone. Strictly speaking, the above model implies the necessity to build a mean of the weights of the zones which have the same linguistic label in an attribute. This mean shall be the weight of that linguistic label, which in turn is multiplied for the middle point of the TFN, representing the same label, giving a number q, of which we consider the smallest integer contained in it, that is, INT (q). At the right of the same linguistic label, thus we create INT (q)-linguistic labels "approximated" with the procedure of Section 2.3. For example, we consider six zones $O1, \ldots, O6$ in which the fuzzy attribute A has received six values with the related weights $W1$ in accordance with Table 3. Then if $U = \{O1, O2, O3, O4, O5, O6\}$, then the fuzzy attribute A is represented by the following string:

$$\mathbf{A} = [O1, O3, O4]^{Cv} [O2]^{v} [O5]^{Mv} [O6]^{F} \tag{9}$$

and consider the linguistic label Cv. For simplicity, let us denote by W_{1i} the weight of the attribute A for the zones

Oi with $i = 1, 3, 4$. Then the mean value $W_{1,Cv}$ for Cv is equal to 2, to be multiplied for 1.0 (cf. Table 2) giving $N_{1,Cv} = \text{INT}(W_{1,Cv} \cdot 1.0) = 2$ which represents the number of new linguistic labels, inserted at the right of Cv. Other new linguistic labels shall not be inserted at the right of the three remaining labels since we have, with evident meaning of the symbology, $W_{1,V} = W_{12} = 1$, $W_{1,Mv} = W_{15} = 1$, and $W_{1,F} = W_{16} = 1$ obtaining $N_{1,V} = \text{INT}(W_{1,V} \cdot 0.7) = 0$, $N_{1,Mv} = \text{INT}(W_{1,Mv} \cdot 0.6) = 0$, and $N_{1,F} = \text{INT}(W_{1,F} \cdot 0.5) = 0$. Then we obtain the following finer string for the attribute **A**:

$$\mathbf{A} = [O1, O3, 04]^{Cv}[-]^{Cv,2}[-]^{Cv,1}[O2]^{V}[O5]^{Mv}[O6]^{F}. \quad (10)$$

This methodology gives the advantage to improve the position of the objects (in our case study, the isoreliable zones) in the set of the attributes, just bearing in mind the new linguistic labels with which the objects can be associated. The calculation of the membership functions for the TFNs, representing the new linguistic labels, is made in the following way.

Let β be the considered linguistic label present in the attribute **Ai** and let $N_{i,\beta}$ be the number of the new linguistic labels obtained with the above procedure. Let α be the linguistic label immediately following β in the linguistic labels of **Ai**. For every $t = 1, \ldots, N_{i,\beta}$, we put $a_{\beta,t} = a_{\alpha} + t^{*}(a_{\beta} - a_{\alpha})/(N_{i,\beta} + 1)$ and similarly for $M_{\beta,t}$ and $b_{\beta,t}$. Then $[a_{\beta,t}, M_{\beta,t}, b_{\beta,t}]$ is the TFN representative of the linguistic label β, t.

2.3. Approximation of the Linguistic Labels. Some TFNs obtained in the final fuzzy attribute, after the successive composition of several strings, must be reconverted in linguistic labels, which can be approximated to known TFNs using the following procedure.

Let β be the TFN to be approximated and α, γ TFNs known (i.e., the meaning of their linguistic labels is known) such that $M_{\alpha} \leq M_{\beta} \leq M_{\gamma}$. By setting $d = M_{\gamma} - M_{\alpha}$ and if $M_{\alpha} \leq M_{\beta} \leq M_{\alpha} + d/10$, then we put $\beta = \alpha$; if $M_{\alpha} + d/10 < M_{\beta} \leq M_{\alpha} + 3d/10$, then we say β is "Next To α" and we write $\beta = \text{NT}[\alpha]$; if $M_{\alpha} + 3d/10 < M_{\beta} \leq M_{\alpha} + 7d/10$, then we say β is "Included Between α and γ" and we write $\beta = \text{IB}[\alpha, \gamma]$; if $M_{\alpha} + 7d/10 < M_{\beta} \leq M_{\alpha} + 9d/10$, then we say β is "Before To γ" and we write $\beta = \text{BT}[\gamma]$; if $M_{\alpha} + 9d/10 < M_{\beta} \leq M_{\gamma}$, then we put $\beta = \gamma$. For instance, taking in account the TFNs of Table 3, let $\beta = \gamma_6$ as in Section 2.1. Since $M_V \leq M_{\beta} \leq M_{Cv}$ and $d = 0.30$, it is easily seen that $\beta = \text{IB}[V, Cv]$.

We note that no matter of comparison between a_{α}, a_{β}, and a_{γ} and similarly for b_{α}, b_{β}, and b_{γ} is requested in this procedure.

3. Fuzzy Reliability for Spatial Databases

We model a *spatial database* in a three-level hierarchic structure as showed in Figure 1. The spatial database is composed of *thematic datasets* referred to as specific spatial domains. Each thematic dataset is composed of *layer*, that is, of geo-referenced vectors or raster themes.

As in [4], after applying the algebraic structure on the strings corresponding to each parameter and recalculating the final string considering the weights assigned to the parameters, we reuse the algebraic structure operator applying it to the final strings associated with the layers of a thematic dataset and recalculating the final string obtained considering the weights associated with these layers. For obtaining the reliability index of the spatial database we apply the operator of the algebraic structure to the final strings associated with each thematic dataset and recalculate the obtained final string considering the weights assigned to each thematic dataset. Therefore we estimate the reliability of the spatial database in each isoreliable zone and apply the calculus on the algebraic structure as described in Section 2 in each level of our spatial database model. Below we describe the single steps that compose our method.

(1) The domain expert creates a partition of the area of study in isoreliable zones; each isoreliable zone is a geographical area, homogeneous in terms of data quality and environmental characteristics.

(2) For each layer the parameters are identified, that is, the observables that affect the quality of the layer, and assigned; for each isoreliable zone, the labels of the corresponding TFNs and the weights of each parameter are assigned as well. The expert creates a fuzzy partition in TFNs of the domain of each parameter.

(3) For each layer the operator of the algebraic structure [6] is applied on the parameters, obtaining a final string to be recalculated (as described in Section 2.2) by considering the weights assigned to the same parameters. We obtain the *index of reliability* of each layer; we call the map of this index *the reliability map* of the layer.

(4) In each thematic dataset its layers are identified as parameters; the string associated with a layer is given by the final string calculated for this layer. For each isoreliable zone the expert assigns the weights of each layer considering the impact of the layer on the quality of its thematic dataset.

(5) For each thematic dataset the operator of the algebraic structure [6] is applied on the related strings by obtaining a final string to be recalculated (as described in Section 2.2) by considering the weights assigned to the same layers. We obtain the *index of reliability* and the corresponding reliability map of the thematic dataset.

(6) Now we identify as parameters the thematic datasets of the spatial database; the string associated with a thematic dataset is given by the final string calculated for this thematic dataset. The expert assigns, for each isoreliable zone, the weights of each thematic dataset considering the impact of the layer on the quality of the spatial database.

(7) The operator of the algebraic structure [6] is applied on the string assigned to each thematic dataset, obtaining a final string to be recalculated (as described in Section 2.2) by considering weights assigned to

TABLE 4: Values for W_1.

ID	A	W_1
O1	Cv	3
O2	V	1
O3	Cv	2
O4	Cv	1
O5	Mv	1
O6	F	1

FIGURE 2: The five isoreliable zones of the area of study.

the same thematic datasets. We obtain the *index of reliability* and the related reliability map of the spatial database.

In Section 4 we present the results by applying our method on a spatial database based on the tool Fuzzy-SRA.

4. Test Results

In our tests the area of study is given by the town of Cava de' Tirreni (Italy). Considering the data quality and the environmental and climatic characteristics, the area of study is partitioned in five isoreliable zones as showed in Figure 2.

We consider the most significant thematic datasets and layers of the spatial database. In our test we consider 5 thematic datasets. In Table 5 we show the thematic datasets and the parameters chosen for each layer.

In the choice of the parameters, particular attention was focused on the absence of primary information connected to geographical entities (e.g., the height of a building or of a spot elevation's point).

Other characteristics that affect the quality of a layer consist of geometric and topological types of errors (isolated street lines or particles intersecting between them).

To simplify the calculus we create for each parameter a fuzzy partition in five TFNs, labeled as showed in Table 3. For brevity we show only the TFNs set for the parameter

TABLE 5: Layers and parameters of the spatial database.

Thematic dataset	Layer	Parameter
(1) Aerial photogrammetric data	(1.1) Streets	Density of isolated lines
		Density of unlabeled lines
	(1.2) Buildings	Density of intersecting polygons
		Density of polygons with uncoded height value
		Density of polygons with wrong height value
	(1.3) Spot elevations	Density of uncoded points
		Density of points with wrong elevation value
	(1.4) Contours	Density of uncoded lines
		Density of lines with wrong elevation value
(2) Cadastral data	(2.1) Terrain parcels	Density of intersecting polygons
		Density of uncoded polygons
	(2.2) Buildings parcels	Density of intersecting polygons
		Density of uncoded polygons
		Density of polygons not overlapping buildings
	(2.3) Land use	Density of intersecting polygons
		Density of unlabeled polygons
(3) Hydrology	(3.1) Rivers	Density of unlabeled lines
		Density of overlapping lines
	(3.2) Lakes	Density of intersecting polygons
		Density of unlabeled polygons
	(3.3) Aquifers	Density of polygons with wrong piezometric elevation value
(4) Geology	(4.1) Surface geology	Density of intersecting polygons
		Density of unlabeled polygons
	(4.2) Land slides	Density of intersecting polygons
		Density of unlabeled polygons
	(4.3) Alluvial zones	Density of intersecting polygons
		Density of unlabeled polygons

TABLE 5: Continued.

Thematic dataset	Layer	Parameter
(5) Infrastructural networks	(5.1) Road network	Density of overlapping lines
		Density of unlabeled arcs
	(5.2) Drainage system	Density of overlapping lines
		Density of unlabeled arcs
	(5.3) Water supply network	Density of overlapping lines
		Density of unlabeled arcs
	(5.4) Electricity grid	Density of overlapping lines
		Density of unlabeled arcs

TABLE 6: TFNs for the layer (1.1)—parameter "density of isolated lines."

Label	Description	a	M	b
Cv	Optimum reliability	0.00	0.005	0.01
V	Good reliability	0.01	0.015	0.02
Mv	Sufficient reliability	0.02	0.03	0.04
F	Mediocre reliability	0.04	0.05	0.06
Sc	Scanty reliability	0.06	0.08	0.10
Bd	Bad reliability	0.10	0.20	0.30
Nl	Null reliability	0.30	0.60	1.00

TABLE 7: TFNs for the Layer (2.1)—parameter "density of intersecting polygons."

Label	Description	a	M	b
Cv	Optimum reliability	0.00	0.005	0.01
V	Good reliability	0.01	0.02	0.03
Mv	Sufficient reliability	0.03	0.04	0.05
F	Mediocre reliability	0.05	0.07	0.09
Sc	Scanty reliability	0.09	0.12	0.15
Bd	Bad reliability	0.15	0.18	0.21
Nl	Null reliability	0.21	0.70	1.00

TABLE 8: Weights assigned for the layer (1.1)—parameter "density of isolated lines."

Zone	Weight
O1	4
O2	4
O3	4
O4	3
O5	3

"density of the isolated lines" of the layer 1.1 street and for the parameter "density of intersecting polygons of the layer 2.1—Terrain parcels."

For each isoreliable zone the weights of the parameters are assigned by an expert. For brevity, we show the weight

TABLE 9: Weights assigned for the layer (2.1)—parameter "density of intersecting polygons."

Zone	Weight
O1	5
O2	5
O3	4
O4	4
O5	4

TABLE 10: Weights assigned to the layers of each thematic dataset for the isoreliable zone $O1$.

Thematic dataset	Layer	Weight
(1) Aerial photogrammetric data	(1.1) Streets	4
	(1.2) Buildings	4
	(1.3) Spot elevations	3
	(1.4) Contours	3
(2) Cadastral data	(2.1) Terrain parcels	4
	(2.2) Buildings parcels	4
	(2.3) Land use	3
(3) Hydrology	(3.1) Rivers	3
	(3.2) Lakes	3
	(3.3) Aquifers	2
(4) Geology	(4.1) Surface geology	4
	(4.2) Land slides	3
	(4.3) Alluvial zones	3
(5) Infrastructural networks	(5.1) Road network	4
	(5.2) Drainage system	3
	(5.3) Water supply network	3
	(5.4) Electricity grid	2

assigned to the two previous parameters for each isoreliable zone.

After combining the strings related to each parameter of a layer, we obtain a final string that is recalculated considering the weights assigned to the same parameters. By using this string we obtain the reliability map of the layer. For brevity, considering Tables 4, 6, 7, 8, 9, we show the isoreliability maps for the layers 1.1 (Figure 3) and 2.1 (Figure 4) with final string, respectively, given by

$$\mathbf{A} = [O1]^{Cv}[O2]^{\mathrm{IB}[V,Cv]}[O3]^{Mv}[O5]^{\mathrm{NT}[Mv]}[O4]^{F},$$

$$\mathbf{B} = [O1, O2]^{Cv}[O3]^{\mathrm{NT}[V]}[O4, O5]^{\mathrm{IB}[Mv,F]}. \tag{11}$$

In these maps we note that the two less reliable zones are $O4$ and $O5$; in fact, in these zones the data are imprecise. After obtaining the final string for each layer of a thematic dataset, we combine them for obtaining the final string for the thematic dataset. Then we recalculate this final string by considering the weights assigned to the layers for each isoreliable zone. In Table 10 we show the weights assigned to each layer for the isoreliable zone $O1$.

TABLE 11: Reliability values obtained for the five thematic datasets.

Isoreliable zone	Reliab. value 1	Reliab. value 2	Reliab. value 3	Reliab. value 4	Reliab. value 5
O1	0.90	0.87	0.85	0.85	0.87
O2	0.86	0.85	0.82	0.82	0.85
O3	0.67	0.74	0.74	0.71	0.71
O4	0.44	0.53	0.53	0.50	0.44
O5	0.47	0.57	0.57	0.53	0.47

TABLE 12: Weights assigned to the thematic datasets for the isoreliable zone O1.

Isoreliable zone	Weights of dataset 1	Weights of dataset 2	Weights of dataset 3	Weights of dataset 4	Weights of dataset 5
O1	5	5	3	4	3
O2	5	5	3	4	3
O3	5	5	3	4	3
O4	5	5	4	5	3
O5	5	5	4	5	3
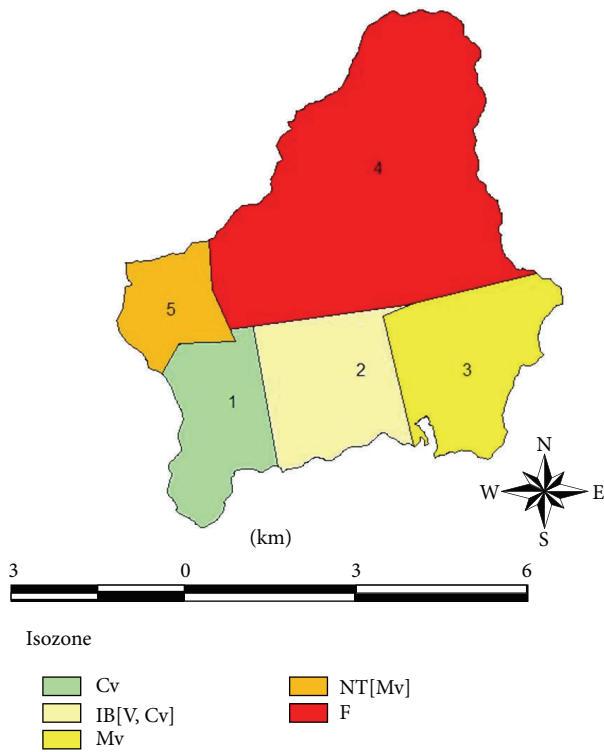
FIGURE 3: The reliability map for the layer 1.1—streets.

FIGURE 4: The reliability map for the layer 2.1—terrain parcels.

Figure 5 shows the thematic map of the thematic dataset "Aerial photogrammetric data." We consider as isoreliability values the central values of the TFN formed in the final string.

In Figure 6 we show the thematic map of the thematic dataset "Cadastral data."

The two reliability maps show that in the isoreliable zones O4 and O5 the quality of the data is poor. This result is confirmed for all the thematic datasets. In Table 11 we show the reliability values obtained for the five thematic datasets in each isoreliable zone.
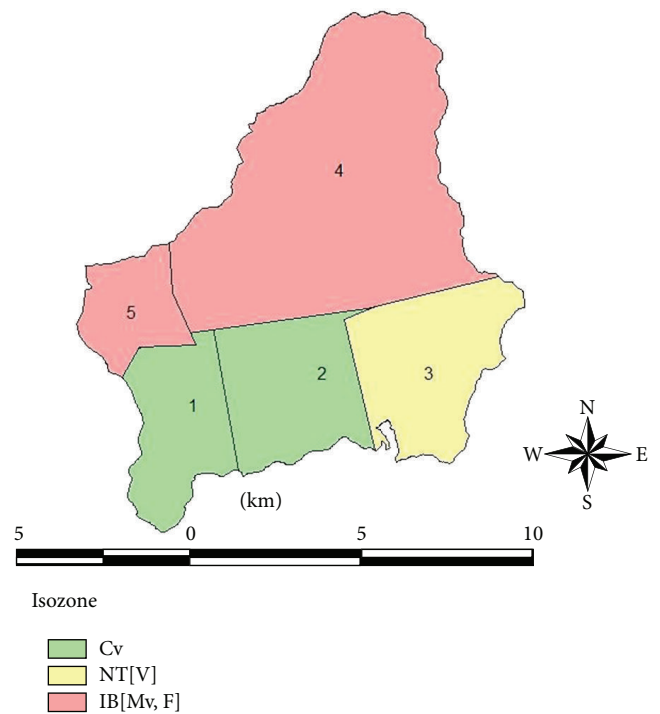
After the calculation of the final strings for each thematic dataset, we combine them for obtaining the final string of the spatial dataset of Cava de' Tirreni (Italy). This final string is recalculated by considering the weights assigned to the thematic datasets for each isoreliable zone. The weights assigned for the five isoreliable zones are showed in Table 12.

The weights of the spatial datasets 3 and 4 for the isoreliable zones O4 and O5 are different with respect to the ones assigned for the isoreliable zones O1, O2, and O3. Indeed in the isoreliable zones O4 and O5 the surface terrains slopes are significant and there are many hydrographic characteristics.
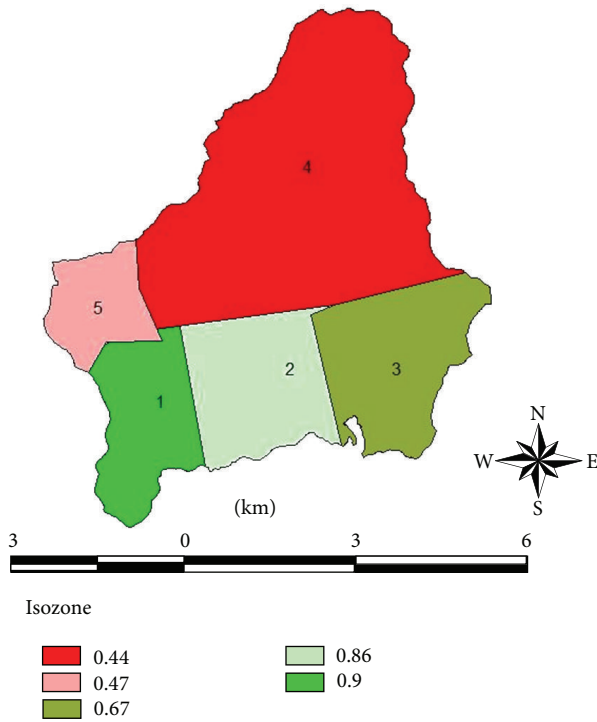
FIGURE 5: The reliability map for the thematic dataset "Aerial photogrammetric data."
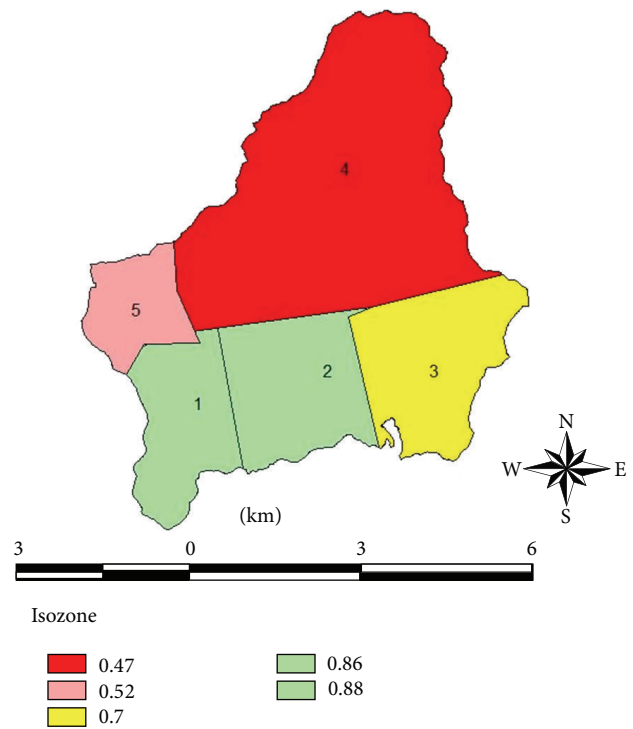

FIGURE 7: The reliability map for the spatial database of Cava de' Tirreni.
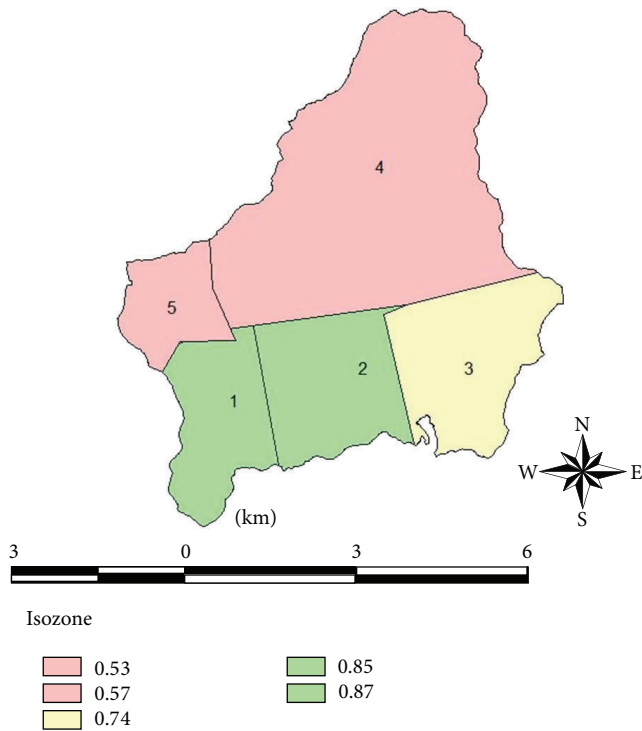

FIGURE 6: The reliability map for the thematic dataset "Cadastral data."

Finally in Figure 7 we show the reliability map for the spatial database.

The results in Figure 7 confirm the previous ones corresponding to the single thematic datasets. The reliability index is good in the isoreliable zones $O1$ and $O2$, fairly good in the isoreliable zone $O3$, and poor in the isoreliable zones $O4$ and $O5$.

5. Conclusions

To give an evaluation of the reliability of spatial database is a complex problem due to the lack of homogeneity of the spatial datasets and to the variation of the data quality on the area of study. Then the usage of a fuzzy logic approach is adequate for measuring the quality of spatial information. In this research we adopt the fuzzy algebraic structure [6] and the fuzzy reliability method applied in [4, 5] for evaluating the reliability of spatial datasets, in order to estimate the reliability of a whole spatial database.

We structure the spatial database in a three hierarchical levels, evaluating the reliability of the single layers, of the thematic datasets, and finally of the spatial database. We test our method on the spatial database of Cava de' Tirreni (Italy). An expert identifies the isoreliable zones and assigns the weights to each parameter, to the layers, and to the thematic datasets. We present the results obtained and the final reliability map of the spatial database on the area of study.

Acknowledgment

References

[1] A. Bardossy and L. Duckstein, *Fuzzy Rule-Based Modelling with Applications to Geophysical, Biological and Engineering Systems*, CRC Press, 1995.

[2] F. Di Martino and S. Sessa, "Spatial analysis with a tool GIS via systems of fuzzy relation equations," in *Computational Science and Its Applications: ICCSA 2011*, vol. 6783 of *Lecture Notes in Computer Science*, pp. 15–30, 2011.

[3] F. Di Martino and S. Sessa, "Spatial analysis and fuzzy relation equations," *Advances in Fuzzy Systems*, vol. 2011, Article ID 429498, 14 pages, 2011.

[4] F. Di Martino, V. Loia, S. Sessa, and M. Giordano, "An evaluation of the reliability of a GIS based on the fuzzy logic in a concrete case study," in *Fuzzy Modeling with Spatial Information for Geographic Problems*, F. E. Petry, V. B. Robinson, and M. A. Cobb, Eds., pp. 185–208, Springer, 2005.

[5] F. Di Martino, S. Sessa, and V. Loia, "A fuzzy-based tool for modelization and analysis of the vulnerability of aquifers: a case study," *International Journal of Approximate Reasoning*, vol. 38, no. 1, pp. 99–111, 2005.

[6] A. Gisolfi and V. Loia, "A complete, flexible fuzzy-based approach to the classification problem," *International Journal of Approximate Reasoning*, vol. 13, no. 3, pp. 151–183, 1995.

[7] F. Di Martino, V. Loia, and S. Sessa, "By fuzzy rules for applications to risk analysis in the Vesuvian area," *Contemporary Engineering Sciences*, vol. 4, no. 2, pp. 55–78, 2011.