ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Approximate Reasoning 38 (2005) 99–111

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

www.elsevier.com/locate/ijar

A fuzzy-based tool for modelization and analysis of the vulnerability of aquifers: a case study

Ferdinando Di Martino [a], Salvatore Sessa [a,*], Vincenzo Loia [b]

[a] *Dipartimento di Costruzioni e Metodi Matematici in Architettura, Università degli Studi di Napoli "Federico II", Via Monteoliveto 3, 80134 Napoli, Italy*

[b] *Dipartimento di Matematica e Informatica, Università degli Studi di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Salerno), Italy*

Received 1 August 2002; received in revised form 1 August 2003; accepted 1 May 2004
Available online 8 July 2004

Abstract

A fuzzy-based tool, called FUZZY-SRA (Fuzzy Spatial Reliability Analysis), is used for realizing a more "reliable" study of the values of the final parameters concerning the vulnerability of aquifers located in the territory of Cava de' Tirreni, city in the district of Salerno (Italy). The SINTACS method is adopted for evaluating the involved parameters and these evaluations are modelled from attributes represented from triangular fuzzy numbers which supply the overall final information if combined with suitable algebraic operations. The tool FUZZY-SRA is implemented inside a GIS (Geographical Information Systems) software.

Keywords: Geographic Information Systems; Fuzzy number; Spatial analysis

1. Introduction

Since 1988 (European Year of the Environment) the European Community (EU) requires to architects, urbanists and Local Administrations to make explicitly more clear and transparent the environmental impact of their projects. Unfortunately in

* Corresponding author. Tel.: +39-81-7903211; fax: +39-81-7903222.
E-mail addresses: fdimarti@unina.it (F. Di Martino), sessa@unina.it (S. Sessa), loia@unisa.it (V. Loia).

0888-613X/$ - see front matter
doi:10.1016/j.ijar.2004.05.001

these studies the know-how is not fully developed and, on the other hand, the laws of the States forming EU do not impose the usage of specific models of simulation and environmental impact evaluation. Generally speaking, these studies are strongly problem-dependent.

In 1987 National Water Well Association (NWWA) and US Environmental Protection Agency (EPA) invented the method DRASTIC [1] in order to solve the problem of the intrinsic vulnerability of the aquifers.

Recent data show that the aquifers satisfy the necessities of the 53% world population. Their increasing pollution obliges the Public Administrations to project a good maintenance, which can be modelled via the Geographic Information Systems (GIS) (for instance, see [2,4,5,7,8,10,12,15,17–20,22,24]).

Unfortunately the model of the pollution of a aquifer in a GIS can meet a low impact because of either possible errors during the analytical procedure or of partial information on input data. Indeed, in the study of the vulnerability indices of the aquifers of the territory of Cava de' Tirreni (city in the district of Salerno, Italy), we used the SINTACS method [6] based on seven hydro-geological parameters, some of which were partially available. Then the treatment of this partial information was based on the Fuzzy Logic, used also for implementation of a tool called FUZZY-SRA (Fuzzy Spatial Reliability Analysis) which incorporates a procedure of calculation whose algebraic operations are given in Section 2.

2. The algebraic operations

In accordance to [9,16] and [19], we first define triangular fuzzy number (for short, TFN) a given map $\mu : \mathbf{R}$ (Reals) $\to [0,1]$ such that $\mu(x) = 0.0$ if $x \leqslant \mathrm{LE}$, $\mu(x) = (x - \mathrm{LE})/(\mathrm{BE} - \mathrm{LE})$ if $\mathrm{LE} \leqslant x \leqslant \mathrm{BE}$, $\mu(x) = (\mathrm{RE} - x)/(\mathrm{RE} - \mathrm{BE})$ if $\mathrm{BE} \leqslant x \leqslant \mathrm{RE}$ and $\mu(x) = 0.0$ if $x \geqslant \mathrm{RE}$, being LE and RE the left and right, respectively, extremes of the real interval (LE, RE) (which is the support of μ, that is (LE, RE) $= \{x \in \mathbf{R} : \mu(x) \neq 0.0\}$) and BE (Between) is any value of (LE, RE). In the sequel the above TFN is represented by $\mu = [\mathrm{LE}_\mu, \mathrm{BE}_\mu, \mathrm{RE}_\mu]$ and let $\lambda = [\mathrm{LE}_\lambda, \mathrm{BE}_\lambda, \mathrm{RE}_\lambda]$ be another TFN. We recall [16] that the addition of λ and μ is the TFN $\mu + \lambda = [\mathrm{LE}_{\lambda+\mu}, \mathrm{BE}_{\lambda+\mu}, \mathrm{RE}_{\lambda+\mu}]$ such that $\mathrm{LE}_{\lambda+\mu} = \mathrm{LE}_\lambda + \mathrm{LE}_\mu, \mathrm{BE}_{\lambda+\mu} = \mathrm{BE}_\lambda + \mathrm{BE}_\mu$ and $\mathrm{RE}_{\lambda+\mu} = \mathrm{RE}_\lambda + \mathrm{RE}_\mu$. We also remember [16] that the multiplication "$*$" of μ by $k \in \mathbf{R}$ is the TFN $k * \mu = [\mathrm{LE}_{k\mu}, \mathrm{BE}_{k\mu}, \mathrm{RE}_{k\mu}]$ such that $\mathrm{LE}_{k\mu} = k \cdot \mathrm{LE}_\mu$, $\mathrm{BE}_{k\mu} = k \cdot \mathrm{BE}_\mu$ and $\mathrm{RE}_{k\mu} = k \cdot \mathrm{RE}_\mu$, denoting with "$\cdot$" the usual arithmetical multiplication. Furthermore, N stands for the set of non-negative integers and when no misunderstanding can arise, we omit the subscript μ in $[\mathrm{LE}_\mu, \mathrm{BE}_\mu, \mathrm{RE}_\mu]$.

Let U be the universe of discourse and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a set of $n \in \mathbf{N}$ linguistic labels, each composed from linguistic modifiers (as, for instance, "False", "More", "More or Less", "Very") and a variable (as, for instance, "Good"). In accordance to their meaning, we assume that $\alpha_1 \leqslant \alpha_2 \ldots \leqslant \alpha_n$ as, for instance, "$\alpha_1 =$ False", "$\alpha_2 =$ More or Less Good",…,"$\alpha_i =$ Good", "$\alpha_{i+1} =$ Very Good",…,"$\alpha_n =$ Completely Good" and each linguistic label is represented from a suitable TFN, denoted also with α_i, $i = 1, 2, \ldots, n$. Let A be a fuzzy attribute, that is a map $A : U \to \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, represented by a string of the following type:

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

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

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

where the β_i's have a similar meaning to the α_i's and $m \in \mathbf{N}$. Following [7], we define the operation Δ between A and B by setting

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

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

$$c_i = \begin{cases} \cup_{j=1,\ldots,i} & (a_{i-j+1} \cap b_j) \quad \text{if } 1 \leqslant i \leqslant m-1 \\ \cup_{j=1,\ldots,m} & (a_{i-j+1} \cap b_j) \quad \text{if } m \leqslant i \leqslant n-1 \\ \cup_{j=i-n+1,\ldots,m} & (a_{i-j+1} \cap b_j) \quad \text{if } n \leqslant i \leqslant m+n-1 \end{cases}$$

As suggested in [7], the c_i's can be calculated by using a simple rule based on the usual arithmetical multiplication. The TFN γ_i's (for $i = 1, \ldots, m+n-1$) are given by

$$\gamma_i = \begin{cases} \frac{1}{k1+k2} \cdot \sum_{j=1}^{i} d2_j \cdot d1_{i-j+1} \cdot (k1 * \alpha_{i-j+1} + k2 * \beta_j) & \text{if } 1 \leqslant i \leqslant m-1 \\ \frac{1}{k1+k2} \cdot \sum_{j=1}^{m} d2_j \cdot d1_{i-j+1} \cdot (k1 * \alpha_{i-j+1} + k2 * \beta_j) & \text{if } m \leqslant i \leqslant n-1 \\ \frac{1}{k1+k2} \cdot \sum_{j=i-n+1}^{m} d2_j \cdot d1_{i-j+1} \cdot (k1 * \alpha_{i-j+1} + k2 * \beta_j) & \text{if } n \leqslant i \leqslant m+n-1 \end{cases}$$

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

$$d_i = \begin{cases} \sum_{j=1}^{i} d2_j \cdot d1_{i-j+1} & \text{if } 1 \leqslant i \leqslant m-1 \\ \sum_{j=1}^{m} d2_j \cdot d1_{i-j+1} & \text{if } m \leqslant i \leqslant n-1 \\ \sum_{j=i-n+1}^{m} d2_j \cdot d1_{i-j+1} & \text{if } n \leqslant i \leqslant m+n-1 \end{cases}$$

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

The following example shall clarify the above concepts and definitions. Let $\{A, B, A'\}$ be a set of three physical parameters (fuzzy attributes), $U = \{O1, O2\}$ be two geographic zones and $\alpha_4 = \beta_4 = Cv$, $\alpha_3 = \beta_3 = V$, $\alpha_2 = \beta_2 = Mv$, $\alpha_1 = \beta_1 = F$ be four TFNs describing the reliability measure of each parameter in both zones. Their linguistic labels are shown in Table 1.

Table 1
The TFNs of the linguistic labels

Linguistic label	Description	LE	BE	RE
Cv	Optimum reliability	0.80	1.00	1.20
V	Good reliability	0.60	0.70	0.80
Mv	Sufficient reliability	0.30	0.40	0.60
F	Scanty reliability	0.00	0.10	0.20

We suppose to have the following strings:

$$A = [O1]^{Cv}[-]^{V}[-]^{Mv}[O2]^{F}$$

$$B = [-]^{Cv}[O1, O2]^{V}[-]^{Mv}[-]^{F}$$

$$A' = [O2]^{Cv}[-]^{V}[-]^{Mv}[O1]^{F}$$

This means that in the zone $O1$ (respectively, $O2$) the parameter A (respectively, A') has a optimum reliable measure but not A' (respectively, A) which has a scanty reliable measure, whereas the parameter B has good reliable measure in both zones. By taking in account that $m = n = 4$ (hence $m + n - 1 = 7$), now we calculate the seven subsets c_i

$$\begin{array}{ccccccc}
 & & & [O1] & [-] & [-] & [O2] \\
 & & & [-] & [O1,O2] & [-] & [-] \\
\hline
 & & & [-] & [-] & [-] & [-] \\
 & & [-] & [-] & [-] & [-] & \\
 & [O1] & [-] & [-] & [O2] & & \\
[-] & [-] & [-] & [-] & & & \\
\hline
[-] & [O1] & [-] & [-] & [O2] & [-] & [-]
\end{array}$$

In the calculation of the related seven TFN γ's, we put for brevity

$$a_{i,j} = d2_j \cdot d1_{i-j+1} * (k1 * \alpha_{i-j+1} + k2 * \beta_j)$$

We observe that $k1 = k2 = 1$, $d2_i = d1_i = 1$ for every $i = 1, \ldots, 4$ and thus we have

$$\begin{array}{cccc}
[0.80, 1.00, 1.20] & [0.60, 0.70, 0.80] & [0.30, 0.40, 0.60] & [0.00, 0.10, 0.20] \\
[0.80, 1.00, 1.20] & [0.60, 0.70, 0.80] & [0.30, 0.40, 0.60] & [0.00, 0.10, 0.20]
\end{array}$$

$$\begin{array}{ccccccc}
\hline
 & & & a_{4,1} & a_{3,1} & a_{2,1} & a_{1,1} \\
 & & a_{4,2} & a_{3,2} & a_{2,2} & a_{1,2} & \\
 & a_{4,3} & a_{3,3} & a_{2,3} & a_{1,3} & & \\
a_{4,4} & a_{3,4} & a_{2,4} & a_{1,4} & & & \\
\hline
\gamma_7 & \gamma_6 & \gamma_5 & \gamma_4 & \gamma_3 & \gamma_2 & \gamma_1
\end{array}$$

where

$$a_{1,1} = 1 \cdot 1 * (1 * [0.00, 0.10, 0.20] + 1 * [0.00, 0.10, 0.20]) = [0.00, 0.20, 0.40]$$
$$a_{2,1} = 1 \cdot 1 * (1 * [0.30, 0.40, 0.60] + 1 * [0.00, 0.10, 0.20]) = [0.30, 0.50, 0.80]$$
$$a_{3,1} = 1 \cdot 1 * (1 * [0.60, 0.70, 0.80] + 1 * [0.00, 0.10, 0.20]) = [0.60, 0.80, 1.00]$$
$$a_{4,1} = 1 \cdot 1 * (1 * [0.80, 1.00, 1.20] + 1 * [0.00, 0.10, 0.20]) = [0.80, 1.10, 1.40]$$
$$a_{1,2} = 1 \cdot 1 * (1 * [0.00, 0.10, 0.20] + 1 * [0.30, 0.40, 0.60]) = [0.30, 0.50, 0.80]$$
$$a_{2,2} = 1 \cdot 1 * (1 * [0.30, 0.40, 0.60] + 1 * [0.30, 0.40, 0.60]) = [0.60, 0.80, 1.20]$$
$$a_{3,2} = 1 \cdot 1 * (1 * [0.60, 0.70, 0.80] + 1 * [0.30, 0.40, 0.60]) = [0.90, 1.10, 1.40]$$
$$a_{4,2} = 1 \cdot 1 * (1 * [0.80, 1.00, 1.20] + 1 * [0.30, 0.40, 0.60]) = [1.10, 1.40, 1.80]$$
$$a_{1,3} = 1 \cdot 1 * (1 * [0.00, 0.10, 0.20] + 1 * [0.60, 0.70, 0.80]) = [0.60, 0.80, 1.00]$$
$$a_{2,3} = 1 \cdot 1 * (1 * [0.30, 0.40, 0.60] + 1 * [0.60, 0.70, 0.80]) = [0.90, 1.10, 1.40]$$
$$a_{3,3} = 1 \cdot 1 * (1 * [0.60, 0.70, 0.80] + 1 * [0.60, 0.70, 0.80]) = [1.20, 1.40, 1.60]$$
$$a_{4,3} = 1 \cdot 1 * (1 * [0.80, 1.00, 1.20] + 1 * [0.60, 0.70, 0.80]) = [1.40, 1.70, 2.00]$$
$$a_{1,4} = 1 \cdot 1 * (1 * [0.00, 0.10, 0.20] + 1 * [0.80, 1.00, 1.20]) = [0.80, 1.10, 1.40]$$
$$a_{2,4} = 1 \cdot 1 * (1 * [0.30, 0.40, 0.60] + 1 * [0.80, 1.00, 1.20]) = [1.10, 1.40, 1.80]$$
$$a_{3,4} = 1 \cdot 1 * (1 * [0.60, 0.70, 0.80] + 1 * [0.80, 1.00, 1.20]) = [1.40, 1.70, 2.00]$$
$$a_{4,4} = 1 \cdot 1 * (1 * [0.80, 1.00, 1.20] + 1 * [0.80, 1.00, 1.20]) = [1.60, 2.00, 2.40]$$

Furthermore we have that

$$d_1 = d1_1 \cdot d2_1 = 1 \cdot 1 = 1$$
$$d_2 = d1_1 \cdot d2_2 + d1_2 \cdot d2_1 = 1 \cdot 1 + 1 \cdot 1 = 2$$
$$d_3 = d1_3 \cdot d2_1 + d1_2 \cdot d2_2 + d1_1 \cdot d2_3 = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 3$$
$$d_4 = d1_4 \cdot d2_1 + d1_3 \cdot d2_2 + d1_2 \cdot d2_3 + d1_1 \cdot d2_4 = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 4$$
$$d_5 = d1_4 \cdot d2_2 + d1_3 \cdot d2_3 + d1_2 \cdot d2_4 = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 3$$
$$d_6 = d1_4 \cdot d2_3 + d1_3 \cdot d2_4 = 1 \cdot 1 + 1 \cdot 1 = 2$$
$$d_7 = d1_4 \cdot d2_4 = 1 \cdot 1 = 1$$

Therefore we have the following TFNs:

$$\gamma_1 = (1/2) * a_{1,1} = [0.00, 0.10, 0.20]$$
$$\gamma_2 = (1/4) * (a_{2,1} + a_{1,2}) = [0.15, 0.25, 0.40]$$
$$\gamma_3 = (1/6) * (a_{3,1} + a_{2,2} + a_{1,3}) = [0.30, 0.40, 0.53]$$
$$\gamma_4 = (1/8) * (a_{4,1} + a_{3,2} + a_{2,3} + a_{1,4}) = [0.42, 0.55, 0.70]$$
$$\gamma_5 = (1/6) * (a_{4,2} + a_{3,3} + a_{2,4}) = [0.56, 0.70, 0.86]$$
$$\gamma_6 = (1/4) * (a_{4,3} + a_{3,4}) = [0.70, 0.85, 1.00]$$
$$\gamma_7 = (1/2) * a_{4,4} = [0.80, 1.00, 1.20]$$

Then we deduce the following fuzzy attribute C

$$C = A\Delta B = [-]^{\gamma_7}[O1]^{\gamma_6}[-]^{\gamma_5}[-]^{\gamma_4}[O2]^{\gamma_3}[-]^{\gamma_2}[-]^{\gamma_1}$$

In other words, interpreting the fuzzy attribute C, we can say that in the zone $O1$ (respectively, $O2$) there is a global "between good and optimum" (respectively, "sufficient") reliable measure of both parameters A and B. By considering the composition $(A\Delta B)\Delta A'$, we have that $k1 = 2$, $d1_1 = d1_7 = 1$, $d1_2 = d1_6 = 2$, $d1_3 = d1_5 = 3$, $d1_4 = 4$ and $k2 = 1$, $d2_1 = d2_2 = d2_3 = d2_4 = 1$. In this case we observe that $n = 7$ and $m = 4$, hence the fuzzy attribute $(A\Delta B)\Delta A'$ has $m + n - 1 = 10$ classes to which 10 TFNs are associated and built with the above formulas. It is possible to show [13] that $(A\Delta B)\Delta A' = A'\Delta(A\Delta B)$ but we omit this fact for brevity.

3. From DRASTIC to SINTACS

DRASTIC is the most widely used spatial method [7] for estimating the pollution of aquifers. The acronym DRASTIC draws origin from the initials of the seven parameters listed in Table 2. To each parameter is assigned a weight which is an integer between 1 and 5 indicating its importance with respect to the remaining ones in accordance to the problem under study: industrial pollution or agricultural pollution. The weights are given in the following Table 2:

In order to adequate the DRASTIC method to the Italian territory, on 1997 it was invented the SINTACS method [6], whose acronym derives from the Italian denomination of seven hydro-geological parameters listed in column 1 of Table 3. They essentially concern the movement of the surface water in the aquifer and therefore its vulnerability to pollutants (for instance, pesticides). Also here to each parameter is assigned a weight which is an integer between 1 and 5.

In accordance to the SINTACS method, generally speaking, in Italy five typologies of hydro-geological areas can be distinguished. They are reported in the titles of columns 3–7 of Table 3 and the user decides the weights by underlying the importance of a parameter with respect to the remaining ones in accordance to the typology of area under study, with the restriction that the sum of these weights must be always equal to 26.

Every parameter is identified with its ID number given in column 2 of Table 3. By studying the aquifers of the territory of Cava de' Tirreni, the experts can essentially distinguish the three typologies of hydro-geological areas reported in the titles of columns 3, 4, 7 of Table 3 and, following [20], make two important assumptions:

Table 2
Weights for parameters in DRASTIC

Parameters	Industrial weight	Agricultural weight
Depth	5	5
Recharge	4	4
Aquifer media	3	3
Soil media	2	5
Topography	1	3
Impact vadose zone	5	4
Conductivity	3	2

Table 3
Weights for parameters in SINTACS

Parameters	ID	Area with normal impact	Area with relevant impact	Area submitted to drainage	Area submitted to Karsticism	Area with textured rocks
Soggiacenza	P1	5	5	4	2	3
Infiltrazione efficace	P2	4	5	4	5	3
Litologia del Non-saturo	P3	5	4	4	1	3
Tipologia del suolo	P4	4	5	2	3	4
Litologia dell'Acquifero	P5	3	3	5	5	4
Conducibilità idraulica	P6	3	2	5	5	5
Acclività della Superficie	P7	2	2	2	5	4

- the required data on the parameters P_i, $i = 1, \dots, 7$, must be effectively available;
- the data must possess precision and accuracy in accordance to the guidelines of [1].

Unfortunately all the data were partially available and moreover the estimation of some parameters, deduced from geographical archives, was judged uncertain or false in some parts of the territory under study. Strictly speaking, then the experts took the decision of dividing the whole territory of Cava de' Tirreni in six zones, that in zones in which if the information is judged reliable, the weight of each parameter assumes a constant value. Each zone is called "iso-reliable zone" and the related geographic map is given in Fig. 1.

The treatment of this partial information was based on the Fuzzy Logic. Precisely speaking, we assume as reliability index of a parameter P_i $(i = 1, \dots, 7)$, in each iso-reliable zone denoted with O_j $(j = 1, \dots, 6)$, one of the linguistic labels representing the TFNs reported in Table 1, defined primary TFNs. Note that the minimum of these TFNs is pointwise zero and the related intersection of all the supports is empty. This assumption is necessary for facilitating the process of linguistic approximation described in Section 6. The inputs of the linguistic labels and of the weights W_{ij} of the parameters P_i $(i = 1, \dots, 7)$ for the zone O_j $(j = 1, \dots, 6)$ are reported in the following Table 4.

In other words, every hydro-geological parameter P_i $(i = 1, \dots, 7)$ is considered as a fuzzy attribute and represented from a string. For instance, we have for $P1$:

$$P1 = [O4, O6]^{Cv}[O1, O2, O3, O5]^{Mv}$$

and similarly for $P2, \dots, P7$.

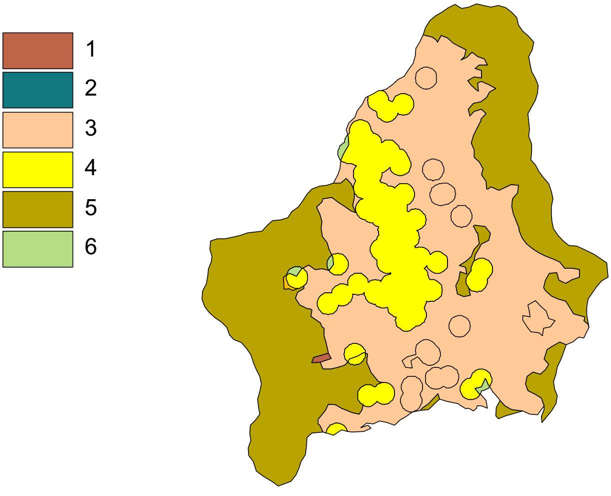

Fig. 1. Geographic map of the iso-reliable zones.

Table 4
Values of the inputs to FUZZY-SRA

ID zone	P1	P2	P3	P4	P5	P6	P7	W_{1j}	W_{2j}	W_{3j}	W_{4j}	W_{5j}	W_{6j}	W_{7j}
*O*1	*Mv*	*Mv*	*V*	*V*	*Mv*	*Mv*	*Cv*	5	4	5	4	3	3	2
*O*2	*Mv*	*Mv*	*V*	*V*	*Cv*	*Cv*	*Cv*	5	4	5	4	3	3	2
*O*3	*Mv*	*Mv*	*Mv*	*V*	*Mv*	*Mv*	*Cv*	5	5	4	5	3	2	2
*O*4	*Cv*	*Mv*	*Mv*	*V*	*Cv*	*Cv*	*Cv*	5	5	4	5	3	2	2
*O*5	*Mv*	*Mv*	*V*	*V*	*Mv*	*Mv*	*Cv*	3	3	3	4	4	5	4
*O*6	*Cv*	*Mv*	*V*	*V*	*Cv*	*Cv*	*Cv*	3	3	3	4	4	5	4

4. Decomposition of the classes

The first step, which precedes the algebraic operations over the strings, consists in an algorithm of decomposition of the classes in classes with finer linguistic labels by taking in account of the weights of each parameter.

4.1. Analysis of the weights

Strictly speaking, we calculate the mean $W_{i,\mu}$ of the weights W_{ij} of the parameter P_i for the iso-reliable zones which have the same linguistic label represented by the TFN μ. If we denote with $[q] = \min\{p \in \mathbf{N} : p \leqslant q\}$ where $q \in \mathbf{R}$, then we shall consider a number of new classes equal to $N_{i,\mu} = [W_{i,\mu} \cdot \mathrm{BE}_\mu]$, denoted by $[-]_{i,\mu}^{\mu,N}, \ldots, [-]^{\mu,1}$, which shall be inserted at the right of the class $[\ldots]^\mu$ in the string *Pi*.

Indeed, in our case we first consider the linguistic label *Cv*. Then we have that $W_{1,Cv} = (W_{14} + W_{16})/2 = (5+3)/2 = 4$ and hence $N_{1,Cv} = [W_{1,Cv} \cdot 1.0] = 4$ since

$BE_{Cv} = 1.0$ (cf. Table 1). Furthermore, we have that $W_{1,Mv} = (W_{11} + W_{12} + W_{13} + W_{15})/4 = (5+5+5+3)/4 = 4.5$ and hence $N_{1,Mv} = [W_{1,Mv} \cdot BE_{Mv}] = [4.5 \cdot 0.40] = [1.8] = 1$. By inserting the new classes, we obtain the following new string representative for the attribute $P1$:

$$S(P1) = [O4, O6]^{Cv}[-]^{Cv,4}[-]^{Cv,3}[-]^{Cv,2}[-]^{Cv,1}[O1, O2, O3, O5]^{Mv}[-]^{Mv,1}$$

4.2. TFNs of the new linguistic labels

The calculation of the values LE, BE and RE for the TFNs representing the new linguistic labels is made in accordance to the following considerations: let β be the TFN of the linguistic label present in the attribute P_i and let $N_{i,\beta}$ be the number of the new linguistic labels obtained with the procedure of Section 5. Let α be the TFN of the linguistic label immediately following in the attribute P_i. For every $t = 1, \ldots, N_{i,\beta}$, we put $LE_{\beta,t} = LE_{\alpha} + t \cdot (LE_{\beta} - LE_{\alpha})/(N_{i,\beta} + 1)$ and similarly for $BE_{\beta,t}$ and $RE_{\beta,t}$. Then we define the TFN $\beta, t = [LE_{\beta,t}, BE_{\beta,t}, RE_{\beta,t}]$ representative of the same linguistic label β, t. Note that if β is the TFN of the last linguistic label present in the attribute P_i, then we assume $\alpha = F$ since $P_i^{-1}(F) = \emptyset$.

Returning to our case study discussed in Section 4.1, we have that $LE_{Cv,4} = LE_{Mv} + 4 \cdot (LE_{Cv} - LE_{Mv})/(4+1) = 0.3 + 4 \cdot (0.8 - 0.3)/5 = 0.7$ and similarly we deduce $BE_{Cv,4} = 0.88$, $RE_{Cv,4} = 1.08$. With analogous computations we obtain that $Cv, 3 = [0.6, 0.76, 0.96]$, $Cv, 2 = [0.5, 0.64, 0.84]$ and $Cv, 1 = [0.4, 0.52, 0.72]$. Since Mv is the TFN of the last linguistic label present in the attribute $P1$, we assume $\alpha = [0.0, 0.1, 0.2]$ (cf. Table 1) and hence we deduce that $LE_{Mv,1} = LE_F + 1 \cdot (LE_{Mv} - LE_F)/(1+1) = 0.0 + 1 \cdot (0.3 - 0.0)/2 = 0.15$. Similarly we have that $BE_{Mv,1} = 0.25$ and $RE_{Mv,1} = 0.4$.

4.3. The interpretation of the new TFNs

On the basis of the primary TFNs, we must interpret the new linguistic labels introduced with the procedure of Section 5. In order to achieve this aim, we suppose that the string $S(P_i)$ contains a class with the new TFN $\mu = [LE_{\mu}, BE_{\mu}, RE_{\mu}]$. Let $\alpha = [LE_{\alpha}, BE_{\alpha}, RE_{\alpha}]$ and $\beta = [LE_{\beta}, BE_{\beta}, RE_{\beta}]$ be two TFNs representative of two primary linguistic labels α and β (present in the string P_i) such that $BE_{\alpha} = BE_{\mu} = BE_{\beta}$ (this assumption is without loss of generality). Then, by putting $d = BE_{\beta} - BE_{\alpha}$, we give the following algorithm:

if $BE_{\alpha} \leqslant BE_{\mu} \leqslant BE_{\alpha} + d/10$, then $\mu = \alpha$;
if $BE_{\alpha} + d/10 < BE_{\mu} \leqslant BE_{\alpha} + 3d/10$, then $\mu = NT[\alpha]$ (i.e., μ "is NEXT TO" α);
if $BE_{\alpha} + 3d/10 < BE_{\mu} \leqslant BE_{\alpha} + 7d/10$, then $\mu = IB[\alpha, \beta]$ (i.e., μ "is IN BETWEEN" α and β);
if $BE_{\alpha} + 7d/10 < BE_{\mu} \leqslant BE_{\alpha} + 9d/10$, then $\mu = BT[\beta]$ (i.e., μ "is BEFORE TO" β);
if $BE_{\alpha} + 9d/10 < BE_{\mu} \leqslant BE_{\beta}$, then $\mu = \beta$.

Note that if β is the TFN of the last linguistic label present in the attribute P_i, then we assume $\alpha = F$ as in Section 4.2.

In our case, by putting $Cv, 4 = \mu$, $Mv = \alpha$, $Cv = \beta$, we have that $d = \mathrm{BE}_{Cv} - \mathrm{BE}_{Mv} = 1 - 0.4 \leqslant 0.6, 3d/10 = 0.18, 7d/10 = 0.82$ and $9d/10 = 0.94$. Then $0.82 = \mathrm{BE}_{Mv} + 7d/10 < \mathrm{BE}_{\mu} = 0.88 \leqslant 0.94 = \mathrm{BE}_{Mv} + 9d/10$, that is $Cv, 4 = \mathrm{BT}[Cv]$. Analogously we deduce $Cv, 3 = Cv, 2 = \mathrm{IB}[Mv, Cv]$ and $Cv, 1 = Mv$.

Since Mv is the TFN of the last linguistic label present in the attribute $P1$, we assume $\alpha = F$ and hence $d = \mathrm{BE}_{Mv} - \mathrm{BE}_F = 0.4 - 0.1 = 0.3$. In this case it is easily seen that $Mv, 1 = \mathrm{BT}[Mv]$. Then we can rewrite the string $S(P1)$ as follows:

$$S(P1) = [O4, O6]^{Cv}\ [-]^{\mathrm{BT}[Cv]}[-]^{\mathrm{IB}[Mv,Cv]}[O1, O2, O3, O5]^{Mv}[-]^{\mathrm{BT}[Mv]}$$

and we obtain analogous expressions for $S(P2), \ldots, S(P7)$ here omitted for brevity.

5. Concomitance and final classification

After the first step which consists in the procedure of Section 4, we apply the algebraic operations defined in Section 2 to the strings $S(P1), \ldots, S(P7)$ by obtaining the long string

$$S = S(P1)\Delta S(P2) \ldots S(P7)$$

containing many empty classes with the related linguistic labels, i.e. classes of type $[-]^{\alpha}$ which we can eliminate because they have no meaning. Indeed, we have the following string:

$$S = [O2, O6]^{\mathrm{IB}[V,Cv]}[O4, O3]^{V}[O1, O5]^{\mathrm{BT}[V]}$$

where the linguistic labels are represented from the following TFNs:

$$\mathrm{IB}[V, Cv] = [0.74, 0.91, 0.94]; \quad V = [0.6, 0.7, 0.8]; \quad \mathrm{BT}[V] = [0.5, 0.63, 0.72].$$

Since the string S contains the global information on each iso-reliable zone O_j $(j = 1, \ldots, 6)$ coming from all the hydro-geological parameters P_i $(i = 1, \ldots, 7)$, it is natural to define as reliability index of the zone O_j its linguistic label, that is the TFN associated to the class $[O_j]$ in the string S. By overlapping the geographic map of the iso-reliable zones with the thematic map of the aquifers of the territory of Cava de' Tirreni, we define the vulnerability index of an aquifer as the iso-reliability index of the zone O_j to which the aquifer belongs. After this operation, the experts pointed out that the aquifers could receive an high value of the vulnerability index in some iso-reliable zones O_j due mainly to the fact that some fuzzy attribute P_i is represented in those zones either from a TFN μ with high values of its membership function or from an high evaluation of its weight W_{ij}. Hence it arises the problem of solving the "concomitant" presence of all the fuzzy attributes P_i in those zone O_j by recalculating a global equal weight W in the string S for each parameter P_i. The experts suggested of calculating W by taking into account of the number m_{μ} of zones O_j which have μ as linguistic label and of the quantity $|1 - \mathrm{BE}_{\mu}|$. In other words, if

$\alpha, \beta, \ldots, \mu$ are the TFNs appearing in the string S and if n is the total number of classes, the experts suggested to use the following formula:

$$W = \{1/n \cdot [|1 - \mathrm{BE}_\alpha| \cdot m_\alpha + |1 - \mathrm{BE}_\beta| \cdot m_\beta + \cdots + |1 - \mathrm{BE}_\mu| \cdot m_\mu)]\}^{-1}$$

and afterwards we again apply the procedure of Section 4 in the string S for getting the final classification.

Indeed, in our case, we have $m_{\mathrm{BT}[V]} = m_V = m_{\mathrm{IB}[V,Cv]} = 2$. Since $n = 6$, we deduce that

$$W = [1/6 \cdot (0.37 \cdot 2 + 0.30 \cdot 2 + 0.09 \cdot 2)]^{-1} = 3.95$$

with which each linguistic label μ is again weighted. By adopting the same procedure of Section 4, we must insert $N_\mu = [W \cdot \mathrm{BE}_\mu]$ new classes at the right of the class $[\ldots]^\mu$ (cf. Section 4.1) and calculate the membership functions of the new TFNs (cf. Section 4.2) which represent the ultimate linguistic labels (cf. Section 4.3) appearing in the final string, also denoted with S. By omitting these calculations for brevity, we finally obtain that

$$S = [O2, O4, O6]^{\mathrm{IB}[V,Cv]}[O3]^{\mathrm{IB}[Mv,V]}[O1, O5]^{\mathrm{BT}[V]}$$

where the linguistic labels are represented from the following TFNs:

$$\mathrm{IB}[V, Cv] = [0.74, 0.91, 0.94]; \quad \mathrm{BT}[V] = [0.5, 0.63, 0.72];$$
$$\mathrm{IB}[Mv, V] = [0.47, 0.60, 0.69]$$

with the respective description "Very Very Good", "Quasi Good" and "Very Very Sufficient". Fig. 2 gives a geographic representation of the final classification of the vulnerability index of the aquifers in the territory of Cava de' Tirreni.

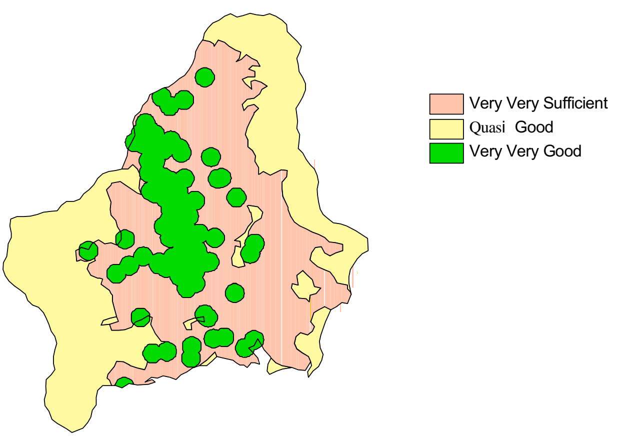

Fig. 2. Final classification of the vulnerability index.

6. Conclusion

By realizing the tool FUZZY-SRA implemented inside the MapObjects OCX GIS [11] from Environmental Systems Research Institute (ESRI, Redlands, CA), we have shown how fuzzy logic can be considered a good support for modelization and analysis of water resources management (see also [3,4,14,21,23,25]).

References

[1] L. Aller, T. Bennett, J.H.Lehr, R.J. Petty, DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings, EPA/600/2-85/018, R.S. Kerr Environmental Research Laboratory, US Environmental Protection Agency, Ada, Oklahoma, 1985.

[2] S.F. Atkinson, J.R. Thomlison, An examination of ground water pollution potential through GIS modelling, University of Texas, Report No. 76203, 1994.

[3] A. Bardossy, M. Disse, Fuzzy rule-based models for infiltration, Water Resources Research 29 (2) (1993) 373–382.

[4] A. Bardossy, L. Duckstein, Fuzzy rule-based modelling with applications to geophysical, biological and engineering systems, CRC Press, Bocca Raton, FL, 1995.

[5] H. Bouwer, Agricultural chemicals and groundwater quality, Journal of Soil and Water Conservation 2 (1990) 184–189.

[6] M. Civita, M. Di Maio, Sintacs, in: S. Pitagora (Ed.), Un Sistema Parametrico per la Valutazione e Cartografia della Vulnerabilità degli Acquiferi all'Inquinamento, Metodologia & Automazione, Bologna, Italy, 1997.

[7] Committee on Techniques for Assessing Groundwater Vulnerability, Groundwater Vulnerability Assessment: Contamination Potential under Conditions of Uncertainty, National Academy Press, Washington, USA, 1993.

[8] Conservation Foundation: Groundwater Protection, Groundwater: Saving the Unseen Resource and a Guide to Groundwater Pollution: Problems, Causes and Government Responses, Washington, DC, 1987.

[9] D. Dubois, H. Prade, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Academic Press, New York, 1980.

[10] A.M. Duda, R.J. Johnson, Targeting to protect groundwater quality, Journal of Soil and Water Conservation 5 (1987) 325–330.

[11] ESRI, MapObjects GIS and Mapping Components, Environmental Systems Research Institute, Inc. 2001.

[12] J.E. Estes, K.C. McGuire, G.A. Fletcher, Coordinating hazardous waste Management Activities Using Geographical Information Systems, International Journal of Geographical Information Systems 1 (4) (1987) 359–377.

[13] A. Gisolfi, V. Loia, A complete flexible fuzzy based approach to the classification problem, International Journal of Approximation and Reasoning 13 (1995) 151–183.

[14] U. Haberlandt, V. Krysanova, A. Bardossy, Assessment of nitrogen leaching from arable land in large river basins, Part II: regionalisation using fuzzy rule based modelling, Ecological Modelling 150 (2002) 277–294.

[15] J.M. Harlin, K.J. Lanfear (Eds.), Proceedings of the Symposium on Geographic Information Systems and Water Resources, American Water Resources Association, Bethesda, MD, 1993.

[16] A. Kaufmann, M.M. Gupta, Fuzzy Mathematical Model in Engineering and Management Science, North Holland, Amsterdam, 1985.

[17] K. Kilborn, H.S. Rifai, P.B. Bedient, Connecting groundwater models and GIS, Geoinformation Systems 2 (1992) 9–30.

[18] D.R. Maidment, GIS and hydrologic modelling, in: M. Goodchild, B. Parks, L. Steyaert (Eds.), Environmental Modelling with GIS, Oxford University Press, New York, USA, 1994, pp. 147–167.

[19] J.M. Mansur, Fuzzy Sets and Economics: Applications of Fuzzy Mathematics to Non-Cooperative Holigopoly, Edward Elgar, London, 1995.
[20] J.W. Merchant, GIS-based groundwater pollution hazard assessment: a critical review of the drastic model, Photogrammetric Engineering and Remote Sensing 60 (9) (1994) 1117–1127.
[21] B. Mitra, H.D. Scott, J.C. Dixon, J.M. McKimmey, Applications of fuzzy logic to the prediction of soil erosion in a large watershed, Geoderma 86 (3–4) (1998) 183–209.
[22] D.W. Moody, Groundwater contamination in the United States, Journal of Soil and Water Conservation 2 (1990) 170–179.
[23] R. Pongracz, I. Bogardi, I.L. Duckstein, Application of fuzzy rule-based modeling technique to regional drought, Journal of Hydrology 224 (1999) 100–114.
[24] F.A. Schoolmaster, P.G. Marr, Geographic information systems as a tool in water use data management, Water Resources Bulletin 28 (2) (1992) 331–336.
[25] B.P. Shrestha, L. Duckstein, E.Z. Stakhiv, Fuzzy rule-based modeling of reservoir operation, Journal of Water Research and Planning Management 122 (4) (1996) 262–269.