European Water 35: 13-22, 2011.

Fuzzy Logic Model Development for Groundwater Pollution Risk Assessment

D. Caniani[1], D.S. Lioi[1], I. M. Mancini[1], S. Masi[1] and F. Sdao[2]
[1] *Dept. Of Engineering and Physics of the Environment, University of Basilicata, Potenza, Italy, donatella.caniani@unibas.it, donataserafina.lioi@unibas.it, ignazio.mancini@unibas.it, salvatore.masi@unibas.it*
[2] *Dept. of Structural Engineering, Geotechnical Engineering, Engineering Geology, University of Basilicata, Potenza, Italy, francesco.sdao@unibas.it*

Abstract: The following paper presents a methodology for a rapid and effective assessment of groundwater pollution risk connected with the presence of uncontrolled landfills. Moreover, this methodology supplies an environmental planning tool to define a groundwater risk hierarchy index in function of some intrinsic characteristics of landfills and territory, in which they are located. The risk assessment was carried out with a fuzzy model, which includes the groundwater vulnerability and the intrinsic hazard of landfills. In particular, for the assessment of groundwater vulnerability the GNDCI-CNR method was used. This method examines the different geological, hydrogeological and morphological characteristics of the territory. Whereas, the hazard of landfills was evaluated together with the risk of contamination of the aquifers in the fuzzy model, taking into account the different descriptor parameters of the landfills. In addition, the fuzzy model was combined with the sensitivity analysis in order to overcome the problem of subjectivity and uncertainty associated with both input data and fuzzy conceptual model.

Keywords: Fuzzy model, Groundwater pollution risk, Groundwater vulnerability, Hazard of landfills, Sensitivity analysis

1. INTRODUCTION

Contamination of groundwater from diffuse or point sources of pollution affects much of the industrialized countries. In particular, water pollution by nitrates or landfills, in other European countries is becoming an environmental catastrophe.

For an early assessment of the risk of groundwater contamination a model of hierarchical classification was developed, in order to identify sites requiring urgent remediation actions.

The developed conceptual model was applied to the Basilicata Region territory for the assessment of pollution risk of the aquifers that underlie some uncontrolled landfills that were identified in the surveys carried out in 2002 on the national territory by the "Corpo Forestale dello Stato" ("Italian Forest Rangers").

In particular, the developed conceptual model is a fuzzy model that assesses the risk of aquifers pollution, by considering groundwater vulnerability and intrinsic hazard of landfills.

The fuzzy logic theory, developed for the very first time by Zadeh in 1965, is today used in several fields; at the beginning it was adjusted to modify the concept of binary logic by approaching it to the human way of thinking and abandoning, in particular, the true/false bivalence, and eliminating the well-known paradoxes.

The earliest practical fuzzy logic applications, however, only date back to the beginning of the Nineties, mainly to control industrial processes and household electrical appliances, and means of transport. Later on, this approach has been used in several fields, such as the environment, to identify the hazard connected to oil drilling waste (Sadiq et al., 2003), botany to identify an index of habitat suitability for Tugai forests (Rüger et al., 2004), in the hydraulics or water management to classify water quality and its subsequent treatment (Chang et al., 2001) or to assess the water quality in function of algae growth (Marsili-Libelli, 2004), in the assessment of the transport of dangerous materials on roads and ducts (Bonvicini et al., 1998), in pedology to model and simulate soil processes (McBratney et al, 1997) and finally in hydrogeology for the identification of the vulnerability of aquifers (Gemitzi et al., 2006).

Given the wide use of this methodology in the environmental field, we decided to apply it to the problem of waste, which has been for years at the core of political and environmental discussions on an international and European scale. Intervening punctually and comprehensively on all the sites involved in illegal waste disposal is often impossible due to the lack of proper economic availability and human resources. It is thus necessary to identify a methodology which, cost-effectively, quickly and immediately identifies the most critical situations for the environment. This is a typical characteristic of the fuzzy logic which is used in this work as the basis in the data elaboration, thus being an extremely important decisional tool to optimise the technical and economic resources.

The fuzzy approach was used to assess the risk of contamination of the aquifers that are below some uncontrolled landfills in the Basilicata territory (Southern Italy), and which are spread between the districts of Potenza and Matera, each of which is characterised by morphological, hydrogeological and environmental parameters, such as: water table depth, leachate production, waste volume and type, soil coverage, landfill activity and proximity to superficial water. Some of these parameters were obtained by means of GIS applications. In order to determine the risk of contamination of aquifers, the intrinsic vulnerability of aquifers was added to the parameters characterizing the hazard of landfills. The intrinsic vulnerability of aquifers was assessed by zoning for homogeneous areas (GNDCI-CNR method). In particular, from the fuzzy inference an environmental risk index in the range 0-1 was derived. Subsequently the risk index was reclassified, supplying 5 risk classes (Very low, Low, Medium, High, Very high). Moreover, in order to face the risk of subjectivity and overcome the problem of uncertainty, linked both to the starting data and the developed model, we had to resort to a sensitivity analysis through which we analysed several fuzzy patterns. The results obtained from the simulations underwent several statistical analyses, in order to identify which fuzzy scheme could represent in the best possible way the distribution of input data. The fuzzy model integrated with the sensitivity analysis allowed us the classification of such sites on a risk scale in order to assess which of these require deeper investigations and consequent environmental remediation.

2. MATERIALS AND METHODS

2.1 Fuzzy logic

Fuzzy logic, developed for the very first time by Zadeh in 1965, is used today in several fields. Initially, it was created in order to modify the binary logic concept, by approaching it to the human way of thinking, abandoning, in particular, the true/false bivalence, and eliminating the well-known paradoxes. Binary logic is characterized by the Aristotelian principles of non-contradiction and the excluded middle. In particular, the principle of non-contradiction (1) affirms that if X is a generic set under a generic element, then x may belong or may not belong to the whole X.

$$X \cap X^c = \emptyset \tag{1}$$

This means that, if x is an element of X, it cannot also belong at the same time to X^c (complementary set).

The principle of the excluded middle affirms that the union of X with its complementary set, X^c, constitutes the universe (2):

$$X \cup X^c = 1 \tag{2}$$

This system of representation of reality is rather rigid and does not take into account its natural transformation. Hence, into the fuzzy logic the classical Aristotelian principles decay. In practice, a variable can assume values in the range 0 and 1. This value represents the degree of membership of a real data to a specific class. The fuzzy class is a fuzzy set, namely a set of ordered pairs: real data, degree of membership.

$$A = \{x, \mu_a | x \in X\} \quad (3)$$

where μ_a is the membership function (Fig. 1), namely the function that associates to any real value a degree of membership to that particular fuzzy class.

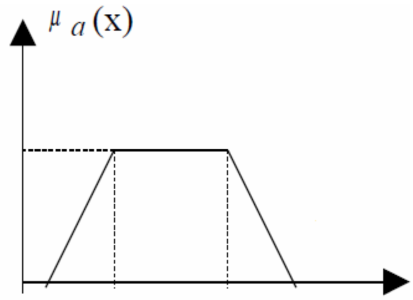

Figure 1. Example of membership function.

Fuzzy logic is therefore primarily the means by which try to make a quantitative description of natural language. Systems based on fuzzy logic are capable of processing fuzzy rules like "if ... then ... else", in opposition to traditional systems.

The fuzzy systems are the actual implementation of fuzzy logic for the solution of complex problems that are difficult to be formalized with a traditional algorithm. The construction of a fuzzy system has inputs and outputs, so it is necessary to define fuzzy sets of input variables and rules that associate inputs to outputs.

The first step of the analysis concerns the identification of fuzzy input variables to the system. The next step regards the definition of fuzzy sets and fuzzy rules that will allow to proceed to fuzzification.

The fuzzification is the process by which input variables are converted to fuzzy measures as a function of their membership to specific fuzzy classes. Eventually, defuzzification converts fuzzy sets into a given numeric value (crisp).

Therefore, the general architecture of a fuzzy system contains an internal database that describes fuzzy sets, a data base for fuzzy rules, a block for fuzzification, a block for inference and one for defuzzification (Fig.2).

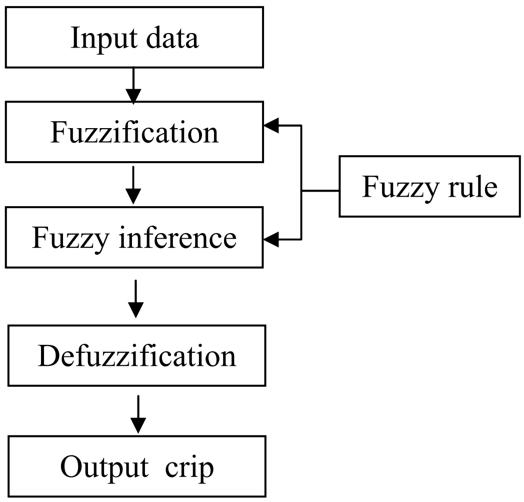

Figure 2. Flow chart of the fuzzy inference system.

2.2 Groundwater intrinsic vulnerability: GNDCI-CNR method.

The assessment of the aquifers vulnerability was carried out through a hydrogeological zoning method made for homogeneous areas: GNDCI-CNR method (Civita M., 1990). This method permits a qualitative assessment of groundwater vulnerability to pollution. The method does not require any numerical parameter but it is a method of hydrogeological zoning made for homogeneous areas. It uses a protocol in which each hydrogeological scenario is associated to a degree of vulnerability. The hydrogeological scenarios of reference are distinguished according to geometry of aquifers, lithologies and their characteristics in terms of effective porosity.

The method is very flexible and can be adapted, even in those situations not listed in the specific methodological protocol. The protocol is characterized by all of the typical situations. Such situations are identified through the main factors that influence the vulnerability of aquifers: water table depth, fracturing and carsification, position of piezometric level. Actually the method does not require any numerical input parameter and it is adapted for the area in question. In fact, the area under study is the whole Basilicata region that is vast and complex under geological, hydrogeological and morphological points of view. Moreover, information about the area of interest is generic and not evenly distributed in the territory. In particular, the vulnerability map has been obtained by taking into account different hydrological scenarios in reference and by assigning degrees of intrinsic vulnerability in the method: extremely high E_E, high E, high A, medium, and M, bottom, B; very low, B_B.

The application of the GNDCI-CNR method is based on technical cartographic overlay of different levels of information and it provides some operational phases:

- identification and definition of hydrological and hydrodynamic characteristics of the area of interest;
- creation of an hydrogeological map and identification of scenario or hydrological and impact scenarios in the area of interest;
- identification, on the basis of the GNDCI-CNR protocol, of the reference scenario or scenarios adaptable to the hydrogeological situation of interest by the consequent assignment of the relative degrees of vulnerability;
- drafting of the aquifer intrinsic vulnerability map accompanied by an appropriate legend.

2.3 Assessment of groundwater pollution risk with fuzzy model.

The sources of groundwater pollution are associated with anthropic activities which include also uncontrolled landfills. In particular, the present study assessed the risk of pollution due to uncontrolled landfills.

The risk assessment was carried out through a fuzzy conceptual model that provides the intrinsic vulnerability of the aquifers and the hazard of landfills. The intrinsic vulnerability of the aquifers was assessed as described in section 2.2 and it is represented in Figure 3. The hazard of landfills was determined through the use of fuzzy logic (Caniani et al., 2010), by considering descriptors parameters of landfills.

The parameters considered were estimated or measured by the census of the State Forestry ("Italian Forest Rangers"); in particular, leachate production (Fig. 6), volume, acclivity and proximity to river index (Fig. 5) were calculated on the basis of other known information, by using GIS tools and historical hydrological data. In order to reduce the number of rules and manage the algorithm easily, the parameters previously indicated were used to define three different fuzzy inferences, as shown in the conceptual diagram in Fig. 4. The results obtained through the first two fuzzy inferences, defined as site vulnerability and landfill potentiality, respectively, were then aggregated to the crisp parameter called landfill conditions, by obtaining the hazard index of each landfill.

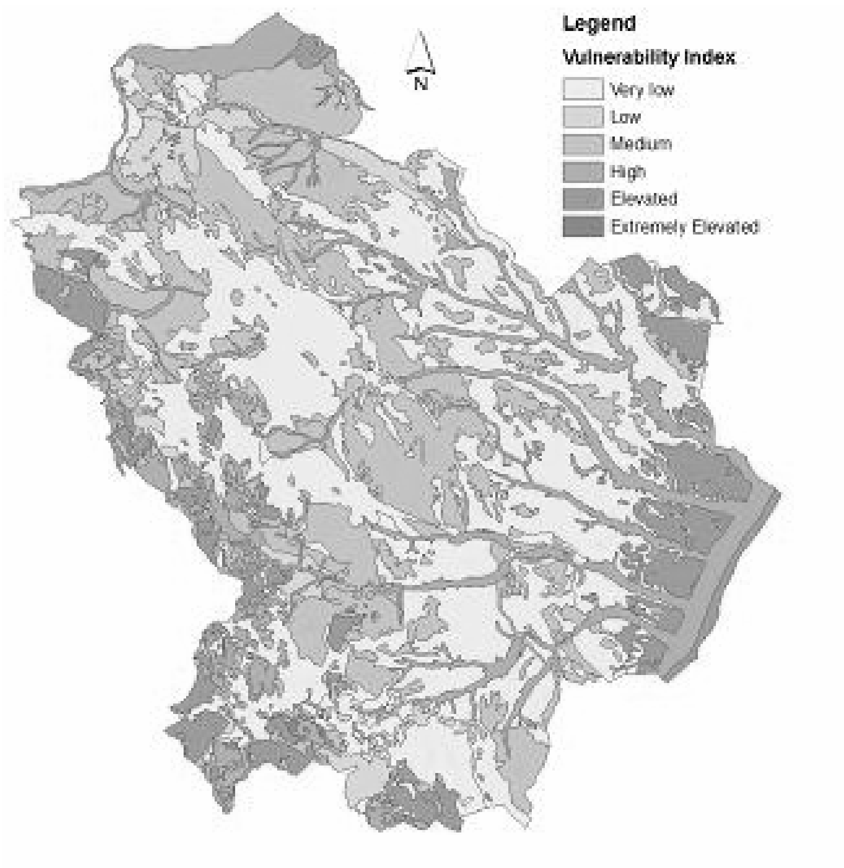

Figure 3: Map of aquifer intrinsic vulnerability

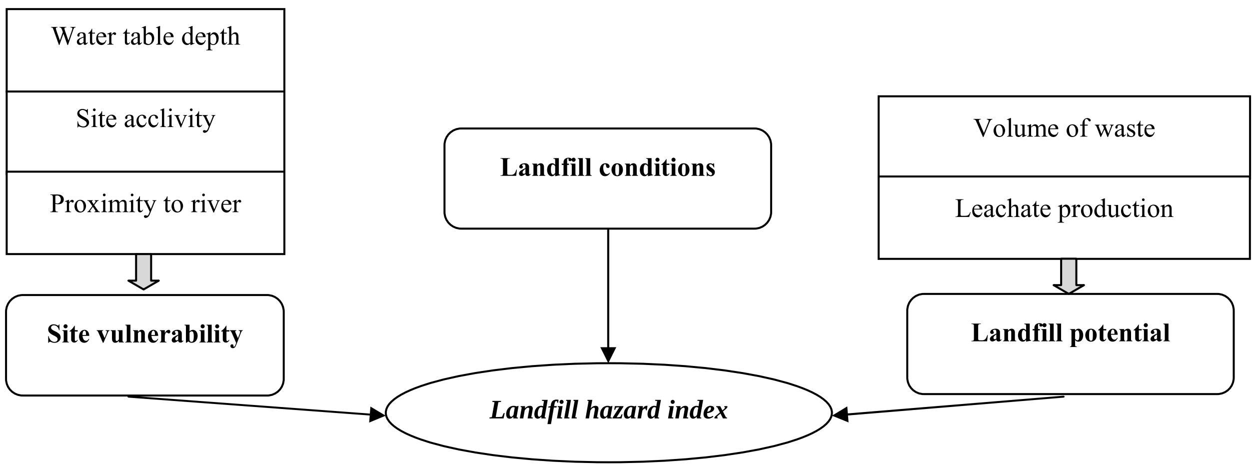

Figure 3. Conceptual diagram of the implemented fuzzy model.

Site vulnerability, defined through acclivity, water table depth and watercourse proximity, allowed us to obtain the site's propensity to be contaminated because of a possible leachate infiltration. In the assignment of the rules, we assumed that the increase in vulnerability is favoured by low slopes, proximity to surface watercourses (meant as index, so the higher the index the higher the proximity to rivers) and reduced depth of the water table. On the contrary, potentiality of landfill evaluates the probability of a landfill to release contaminants depending on the waste volume and the leachate production. Therefore, with the increase of these two factors, this potentiality will increase as described by the fuzzy rules. The procedure to determine the landfill hazard index combines the results obtained by the two previous fuzzy diagrams with the addition of the landfill conditions. The basis of the fuzzy inference has been drawn up keeping in mind that with increasing values of the three parameters of the subset, increases the hazard of landfills.

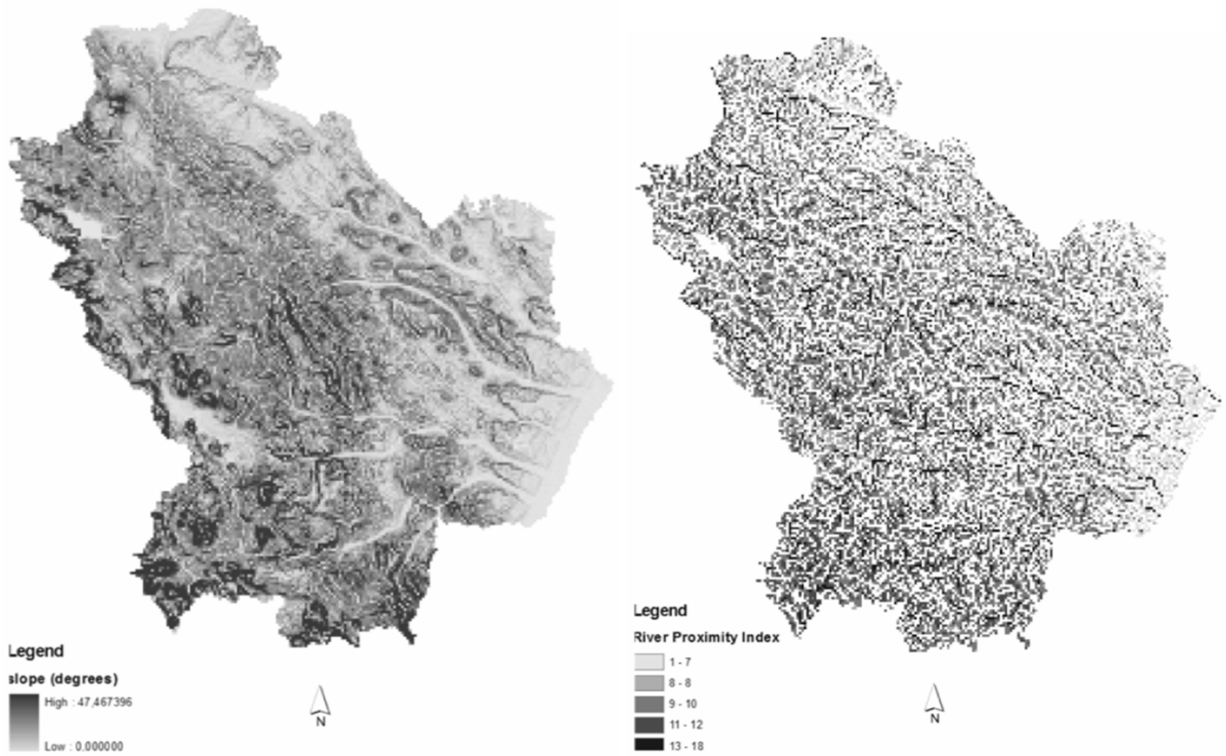

Figure 4. Maps of site acclivity (on the left) and river proximity index (on the right).

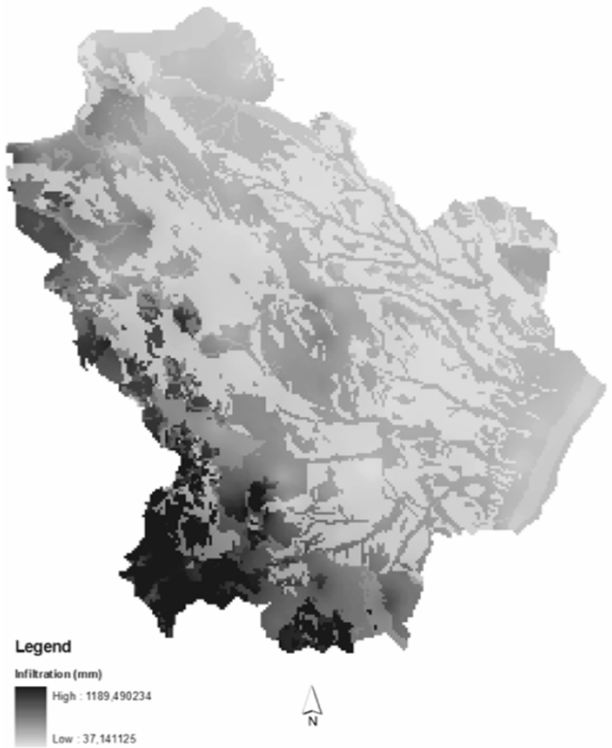

Figure 5. Map of leachate production.

After the assessment of the intrinsic vulnerability of aquifers, shown in Figure 3, the fuzzy conceptual model was applied (Fig. 7). Subsequently, it was possible to rank the sites at risk of pollution and, at the same time, draw up a hierarchy of the sites that require more urgent reclamation actions.

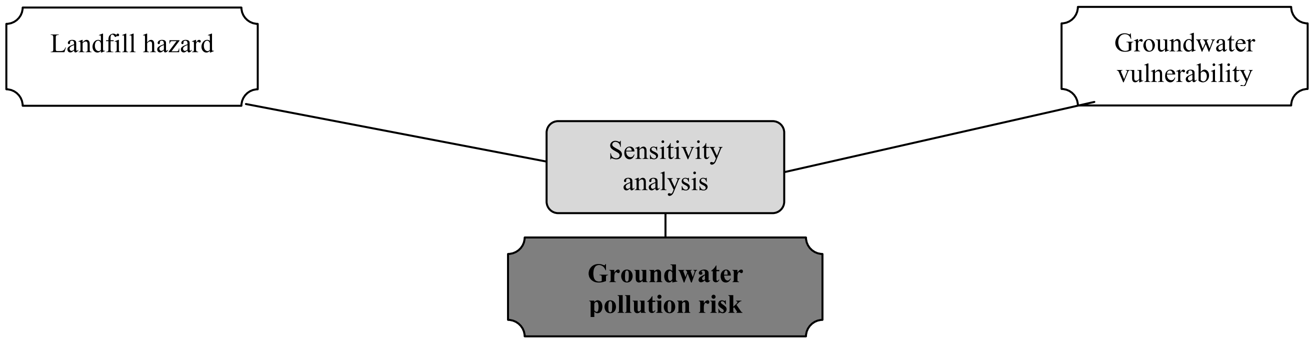

Figure 7. Fuzzy conceptual schema for groundwater pollution assessment.

Moreover, in order to assess risk of contamination of aquifers and, at the same time, to face the risk of subjectivity and overcome the problem of uncertainty, linked both to the input data and the developed model, we elaborated the sensitivity analysis through which we analysed several fuzzy patterns. The fuzzy patterns differ in membership functions and methods of defuzzification. For each of the fuzzy inferences we defined the membership function, the membership classes (very high, high, medium, low, very low), the fuzzy rules "if-then" and the defuzzification method. Fuzzy rules were defined whereas aquifer pollution risk increases with intrinsic vulnerability of aquifers and hazard of landfills.

A membership function is a function that associates a value (usually numerical) with the level of membership to a set. By convention, the real number which represents the level of membership takes the 0 value when the element does not belong to the set, the 1 value when the element belongs to it completely, and an intermediate real number when the element belongs to the set in a measure. Membership functions can be of several types: the simplest are made up by straight lines, while the most used are the triangular and trapezoidal functions. There are other more complex functions, i.e. the Gauss function made up of a simple Gaussian curve and the Gauss2 function given by the fusion of two different Gaussian functions. Moreover, among this type of functions, there is the bell membership function (Gbell) which is a hybrid of the Gaussian function, it is mainly used to manage non-fuzzy sets. In order to face a possible asymmetry, in fact, we can use another type of function, such as the sigmoid function which may have both left and right asymmetry, and with a horizontal asymptote. In addition to this one, we have further asymmetric functions, the Dsigm and Psigm membership functions. Three more membership functions correlated with them are the functions Z, S and Pi; the first one is an asymmetric function open to the left, the second is open to the right, while the third one is asymmetric but closed to both its ends.

The rules are not represented by complex mathematical models, but by simple linguistic expressions, which are turned into mathematical formalism by the language "if-then" of the fuzzy logic. The rules are usually made up of an "if-then" structure that is made up of an antecedent, which defines the conditions, and a consequent, which defines the action. Fuzzy output obtained from the inference is often unusable, and therefore must be converted back into a deterministic value through an operation called defuzzification. Different methods of calculation have been proposed to achieve this result; the mostly used of which are the following: Centroid method, in which the chosen numerical value for the output is calculated as the centre of mass of the fuzzy set; Bisector method, in which the output is the abscissa of the bisector of the area subtended to the fuzzy data set; Middle of maximum method, in which the output value is determined as the average of maximum values (*Mom*: middle of maximum); Largest of maximum method, in which the output numerical value is calculated as the maximum of the maximum (*Lom*: Largest of maximum); Smallest of maximum method, in which the output value is represented by the output minimum value (*Som*: Smallest of maximum).

3. RESULTS AND DISCUSSION

This study enabled the assessment of contamination risk of the aquifers underlying some uncontrolled landfills in the Basilicata Region territory (Southern Italy), by developing and applying a fuzzy conceptual model. Through this model, taking into account both intrinsic vulnerabilities of aquifers and hazards of landfills, we developed an environmental index risk. In order to further reduce subjectivity and refine the model, we carried out a sensitivity analysis that allowed us to perform a series of simulations characterised by different input parameters of the model. Sensitivity analysis was performed by repeating the whole fuzzy inference procedure, modifying one only membership function or defuzzification method at a time, keeping unvaried all the other variables, thus performing 55 different simulations.

The results obtained from the simulations underwent several statistical analyses, in order to identify which fuzzy scheme could represent in the best possible way the distribution of input data (depth to water table, site acclivity, proximity to watercourses, waste volume, leachate production, and landfill conditions). Initially, an analysis related to the variability of the starting data sample was performed, by adopting afterwards a criterion to make a comparison with the results obtained from fuzzy simulations. To this end, in order to make a comparison in terms of data dispersion, we deemed it necessary to identify, for each uncontrolled landfill, an average value obtained through the average of the parameters that characterise it, opportunely standardised, thus obtaining a data set characteristic of input data, directly comparable with the results of simulations.

In order to choose the most appropriate fuzzy scheme to represent input data we compared two dispersion indices (standard deviation and variance) of all the distributions with those of the input data, in order to identify the distributions whose values of deviation and variance can be compared with those related to input data.

In the case under analysis, the distributions which showed, in several fuzzy schemes, a higher resemblance with respect to that of input data, are those obtained through implementing in the model either the membership functions Gauss2, Pi, Trapezoidal together with the Centroid defuzzification method, or the membership functions Trapezoidal and Pi, together with the Bisector defuzzification method, moreover membership functions Dsig and Psig integrated with the Mom (Middle-of-maximum) defuzzification method and membership function Z with Lom defuzzification method.

To compare more accurately the variances of the distributions corresponding to each fuzzy model with the input data, an F test was carried out (Table 1).

Table 1. F test results

	Centroid	Bisector	Mom	Lom	Som
Gauss2	0.793	0	0	0	0
Triangular	0.02	0.032	0.003	0.002	0.035
Trapezoidale	0.631	0.379	0.021	0.021	0.076
Gbell	0	0	0	0	0
Gauss	0	0	0	0	0
Disig	0	0	0.672	0	-
Psig	0	0	0.649	-	0
Pi	0.784	0.429	0.021	0.128	0.087

The results obtained from the analysis of deviation were confirmed by the F test, which highlighted that the Centroid defuzzification method associated with the Trapezoidal, Gauss2 or Pi membership functions and the Mom defuzzification method associated with the Psig membership function, gave comparable results (Table 1). The fuzzy scheme characterized by the Gauss2 membership function together with the Centroid defuzzification method was selected.

The environmental risk index, obtained through the latest above described fuzzy scheme, was successively revised for a classification of groundwater pollution risk in linguistic terms (Very low, Low, Medium, High, Very high), by using a cumulated frequency distribution curve (Figure 8), and choosing as limits of the classes the points of the curve where a clear slope variation is observed.

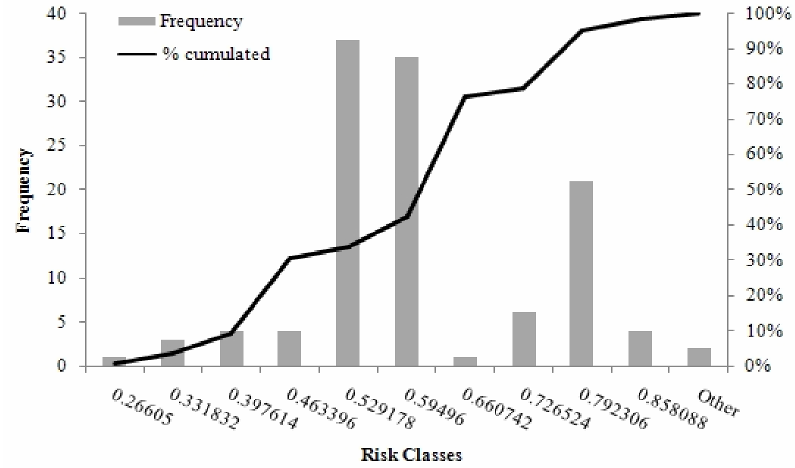

Figure 8: Cumulated frequency curve.

The graphs in Figures 9 and 10 show respectively the percentage and spatial distribution of groundwater pollution risk.

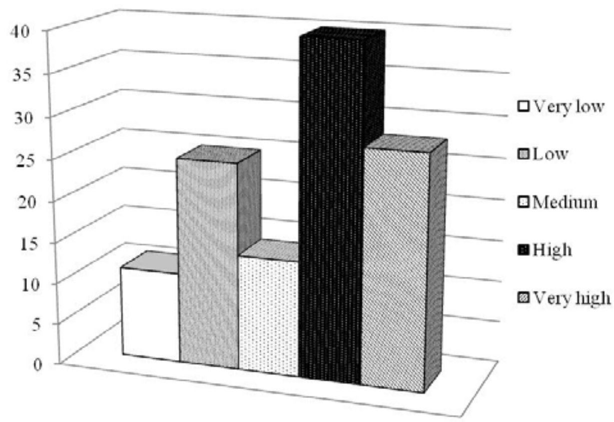

Figure 9: Percentage distribution of aquifers in the pollution risk classes.

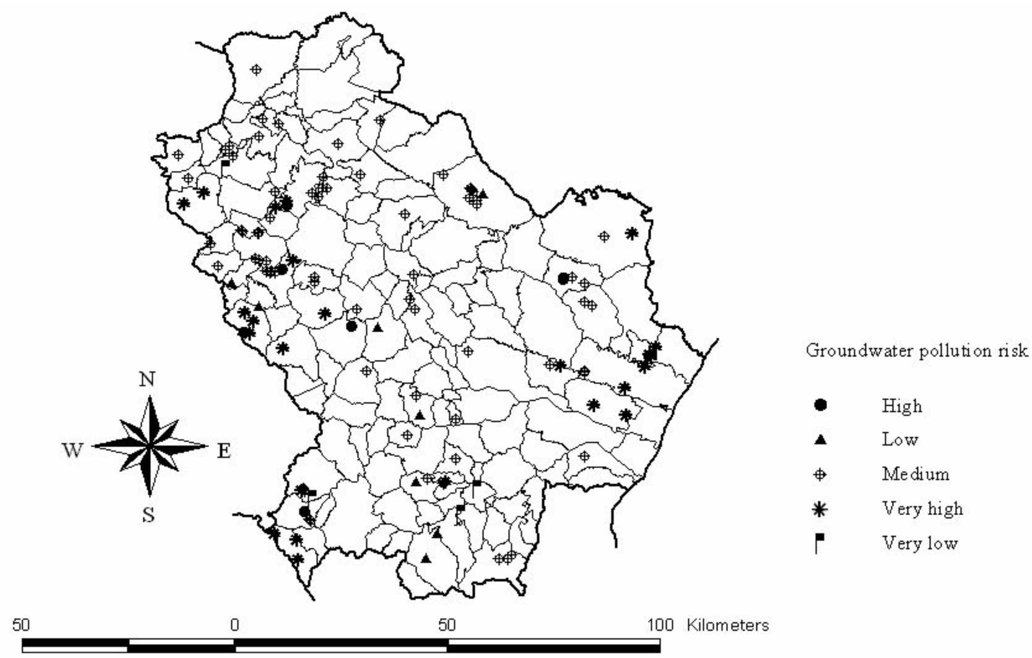

Figure 10: Map of contamination risk of the aquifers.

From the choice of the fuzzy scheme previously described we highlighted that the methodology has identified a small percentage of groundwater pollution risk with a score very low (3%), low (7%) and high (6%); in fact, the cake diagram and the histogram show that the majority of groundwater has a medium risk (61%) and a very high risk (23%).

4. CONCLUSIONS

The fuzzy model is a fast and economic method for the assessment of groundwater pollution risk of aquifers underlying some uncontrolled landfills in the Basilicata Region territory. Moreover it is a useful support to the decisions concerning the reclamation measures to adopt in the sites classified as hazardous.

Through this method it has been possible the identification of environmental risk indexes for each of the examined sites. These indexes have been turned into nominal values, in order to perform a risk classification and generate a scale of priority. The results show that a considerable percentage of sites, around 30%, needs urgent reclamation actions.

This study suggests an innovative methodology characterized by the integration between the fuzzy approach and the sensitivity analysis, able to mitigate the problems connected to the subjectivity and arbitrariness of the evaluations based on the fuzzy approaches normally present in literature, especially for the choice of the membership functions or the defuzzification methods. The proposed study, in fact, highlights how, by varying the choice of the membership functions or the defuzzification methods, it is possible to obtain extremely different results.

A future development of research will aim at refining the model with its validation, by carrying out field analyses, or developing a fuzzy neural network to overcome the difficulties connected with the assignment of the fuzzy rules.

REFERENCES

Bonvicini S., Leonelli P., Spadoni G., 1998. Risk analysis of hazardous materials transportation: evaluating uncertainty by means of fuzzy logic. Journal of Hazardous Materials; 62, 59–74.

Caniani D., Lioi D. S., Mancini I.M., Masi S., 2010. Application of fuzzy logic and sensitivity analysis for soil contamination hazard classification. Waste Management 31, 583–594.

Civita M., 1988. Una metodologia per la definizione e il dimensionamento delle are di salvaguardia delle opere delle opere di presa delle sorgenti normali. Boll. Ass. Min. SUBALPINA; 25, 4, 423-440.

Civita M., 1990. Legenda unicata per le Carte della vulnerabilità dei corpi idrici sotterranei/ Unified legend for the aquifer pollution vulnerability Maps. Studi sulla Vulnerabilità degli Acquiferi, 1 Pitagora, Bologna.

Gemitzi A., Petalas C., Tsihrintzis V.A., Pisinaras V., 2006. Assessment of groundwater vulnerability to pollution: a combination of GIS, fuzzy logic and decision making techniques. Environmental Geology; 49, 653–673.

Marsili-Libelli S. 2004. Fuzzy prediction of the algal blooms in the Orbetello lagoon, Environmental Modelling & Software; 19, 799-808.

McBratney A. B., Odeh Inakwu O.A., 1997. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions. Geoderma; 77, 85–113.

Nadja Rüger, Maja Schlüter, Michael Matthies, 2005. A fuzzy habitat suitability index for Populus euphratica in the Northern Amudarya delta (Uzbekistan). Ecological Modelling; 184, 313-328.

Ni-Bin Chang, Chen H. W. and Ning S. K, 2001. Identification of river water quality using the Fuzzy Synthetic Evaluation approach. Journal of Environmental Management; 63, 293-305.

Rehan Sadiq, Tahir Husain, 2005. A fuzzy-based methodology for an aggregative environmental risk assessment: a case study of drilling waste. Environmental Modelling & Software; 20, 33-46.

Zadeh L.A., 1965; Information and control; Fuzzy sets 338-353.