10th International Conference on Hydroinformatics
HIC 2012, Hamburg, GERMANY

FUZZY POSSIBILISTIC APPROACH TO THE ESTIMATION OF UNCERTAINTY IN HYDROLOGICAL DATA AND MODELS

WASSIM ABU ABED (1), PETER MILBRADT (2)
(1): Institute of Computer Science in Civil Engineering, Leibniz Universität Hannover, Callinstr. 34, Hanover, 30167, Germany
(2): smile consult GmbH, Vahrenwalderstr. 4, Hanover, 30165, Germany

Impreciseness, incomplete knowledge, linguistic vagueness and randomness are the main sources of uncertainty in hydrological engineering. The enhancement of the information content of hydrological data by including a quantification of the resulting uncertainties is very important to help the engineer make better informed decisions. A fuzzy possibilistic computational framework, in which the estimation and propagation of the uncertainty that is induced by the inherent impreciseness and incomplete knowledge in hydrological data and simulation techniques are consistently fulfilled, is the subject of this paper.

INTRODUCTION

The proper evaluation of the different uncertainties in hydrological observations and the modelling process in a scientifically rigorous way has been and still is a challenging arena for research and development. Different concepts and techniques with different philosophical backgrounds exist for handling uncertainty in hydrological modelling and simulation. The estimation and propagation of the kind of uncertainty resulting from impreciseness and incomplete knowledge are the main themes of this paper. The theory of fuzzy sets and the underlaid possibility theory offer the appropriate concepts and techniques for the purpose of dealing with such uncertainty. Fuzzy sets were proposed as a new concept to deal with the uncertainty induced by impreciseness or vagueness rather than by randomness, see [1]. Possibility theory is related to the theory of fuzzy sets by defining the concept of a possibility distribution as a fuzzy restriction on the values that may be assigned to a variable, see [2]. In the following section modeling and processing the kind of uncertainty resulting from impreciseness and incomplete knowledge will be introduced. Fuzzy possibilistic interpolation and numerical methods will be visited in the section that follows. The applicability of these is then demonstrated with the help of two case studies in coastal morphodynamics and groundwater modelling.

POSSIBILISTIC PROCESSING OF UNCERTAINTY

The notions of gradual possibility and necessity presented by possibility theory extend fuzzy set theory to be able to model partial belief. Possibility theory is an uncertainty calculus based on the idea of gradual membership offered by fuzzy set theory. Modelling partial truth is the essence of fuzzy sets. The definitions of fuzzy sets and the associated membership function, as given in the following, deliver the sense of gradualness in the compatibility of a statement with a state of fact, hence the degree of the truth of the statement.

Fuzzy Set over R

Let R be the universal set of the real numbers. A fuzzy set $\tilde{X}$ over R is defined as a set of ordered pairs $(x, \mu_{\tilde{X}}(x))$ of real numbers, $x \in R$, and assigned values of a membership function, $\mu_{\tilde{X}}(x) \in [0,1]$, representing the degree, in which the real number x satisfies certain properties.

In this context fuzzy set theory based only on the multi-valued logic, which deals with partial truth only, has nothing to say about partial belief and uncertainty. Fuzzy sets placed in the constellation of possibility theory have a possibility theoretic interpretation different from the one taken from the preliminary definition that is based on a multi-valued logic. Given the case, in which the value of a variable x is not known precisely and all that is known about the value of x is that "x is X", where X is a set of possible values of x, this corresponds to a situation of incomplete information and impreciseness. In possibility theory the membership function of the fuzzy set is replaced by a possibility distribution function. Given a specific state of knowledge the possibility distribution function expresses the level of possibility that a value in a set of different values is in fact the value of the variable x, see [3].

Possibility Distribution Function

A possibility distribution function $\mu_{\tilde{x}}$ of an unknown variable x is a mapping $\mu_{\tilde{x}} : X \subset R \rightarrow [0,1]$ from the set X of the possible values of a variable x to the unit interval [0,1], where

$\mu_{\tilde{x}}(\xi) = 0$ means that $x = \xi$ is rejected as impossible, and
$\mu_{\tilde{x}}(\xi) = 1$ means that $x = \xi$ is totally possible.

Partial belief aroused by the knowledge embodied in the possibility distribution function $\mu_{\tilde{x}}(\xi)$ is quantified and conveyed by two dual measures in possibility theory, namely possibility and necessity measures. Possibility measure evaluates the consistency of the statement $x \in A$, where A is a non empty set, with the available knowledge represented by the possibility distribution function $\mu_{\tilde{x}}(\xi)$. In a dual manner to possibility measure, necessity measure evaluates the certainty of the statement $x \in A$. This certainty is implied by the available knowledge embodied in the possibility distribution function $\mu_{\tilde{x}}(\xi)$. For more information on possibility and necessity measures see [3].

Calculus of Fuzzy Numbers

A fuzzy number $\tilde{x}$ is a convex, normal fuzzy set over the set of real numbers R with an upper semi-continuous possibility distribution function, a bounded support set and exactly one element $m \in R$ with $\mu_{\tilde{x}}(m) = 1$. The real number m is called the modal value of the fuzzy number. Examples of fuzzy numbers are given in Figure 1. For more information on fuzzy sets and numbers see [3].

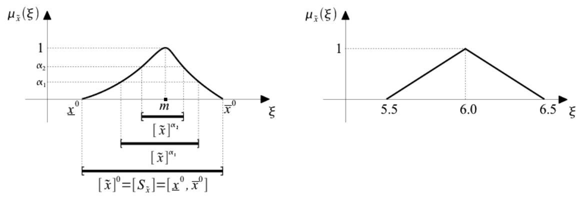

Figure 1. Examples of fuzzy numbers and the so-called α-Cut sets $[\tilde{x}]^{\alpha}$ of a fuzzy number

The basis for a fuzzy mathematical framework was established on ZADEH's genuine idea of extending functions of real numbers to functions of fuzzy sets by means of the so-called extension principle [4]. This has opened the door for the development of a fuzzy calculus that enabled the integration of the fuzzy possibilistic estimation of uncertainty in the classical mathematical concepts used in system and process modelling.

Using computational procedures for evaluating functions with an arithmetic notion on the set of fuzzy numbers, without dropping the basic idea of the extension principle as a possibility propagation mechanism consistent with possibility and necessity measures, is very essential. Since the introduction of ZADEH's extension principle many approaches have been proposed. The so-called arithmetic of LR-fuzzy numbers is one of the first of these approaches. It presumes special membership functions of the fuzzy operands. This restriction confines its application to a very small area of simple problems. The formulation of the extension principle in terms of a mapping of the α-cut sets of the fuzzy operands to an α-cut set of the fuzzy result has stimulated the adoption of interval arithmetic [3]. This approach results in an inconsistent propagation of possibility when applied to problems with a non-trivial mathematical complexity. In [5] the so-called vertex method was proposed. This method is also based on the α-cut set formulation of the extension principle. For the evaluation of monotonic functions the vertex method is a consistent method for the correct propagation of possibility. For non-monotonic functions further development is still needed.

FUZZY POSSIBILISTIC INTERPOLATION AND NUMERICAL METHODS

Fuzzy interpolation methods can be understood as procedures for interpolating data, in which the impreciseness or fuzziness is a non-separable part of the sampling points, as well as procedures for interpolating either crisp or fuzzy data quantifying the fuzziness induced by the reduction of the actual state of the interpolated analytical or empirical function.

Fuzziness in the sampling points

Given is a set of fuzzy data such that at various crisp sampling points x_i there is a fuzzy sampling quantity $\tilde{y}_i$ with $i = 1,\cdots,m$. The interpolation methods, used for interpolating such discrete fuzzy data in terms of relatively simple functions, can be based on the simple form of an interpolating function

$$\tilde{\tilde{f}}(x) = \sum_{i=1}^{m} \tilde{y}_i \cdot \varphi_i(x) \tag{1}$$

Considering a one-dimensional linear interpolating function the basis functions take the form of:

$$\varphi_1(x) := \frac{x_2 - x}{x_2 - x_1} \tag{2}$$

$$\varphi_2(x) := \frac{x - x_1}{x_2 - x_1} \tag{3}$$

Fuzziness of the interpolating function

The set of sampling data is an incomplete information item, because it does not completely account for the original analytical or empirical function. Interpolation methods are used to get a totally unknown piece of information from this incomplete information item. Thus, the values of the interpolating function are implicitly uncertain, regardless of whether the sampling data are precise or not. Moreover, the choice of the method for interpolating specific data is a subjective decision. It depends on the expert opinion on which interpolating function best suits the data set. Hence, an extra amount of uncertainty is also induced by this decision. This kind of implicit uncertainty can be quantified, for example, in the case of linear interpolation, by substituting x in Eq. (1) with a fuzzy function $\tilde{q}(x)$ that maps the coordinate x of the interpolation position to a fuzzy number. The possibility distribution function $\mu_{\tilde{q}}$ of the resulting fuzzy number $\tilde{q}(x)$ should reflect the subjectivity in the choice of the method of interpolation and satisfy the interpolation conditions, i.e, $\tilde{q}(x_1) = x_1$ and $\tilde{q}(x_2) = x_2$. Such a possibility distribution function can be given for the case of one-dimensional linear interpolating function by

$$\mu_{\tilde{q}}(\xi) = \begin{cases} L\left(\dfrac{x - \xi}{x - x_1}\right) & for \quad x_1 \le \xi < x \\ 1 & for \quad \xi = x \\ R\left(\dfrac{\xi - x}{x_2 - x}\right) & for \quad x < \xi \le x_2 \end{cases} \tag{4}$$

where L and R are given in the following, where φ_L and φ_R are given by the subjectively chosen basis functions φ_1 and φ_2, respectively.

$$L(\bar{\xi}) := (1 - \bar{\xi})^{\alpha_L}(1 - \varphi_L) + \varphi_L \quad ; \quad \alpha_L := \left(\frac{1 - \varphi_L}{\varphi_L}\right) \quad ; \quad \bar{\xi} = \left(\frac{x - \xi}{x - x_1}\right) \tag{5}$$

$$R(\bar{\xi}) := (1 - \bar{\xi})^{\alpha_R}(1 - \varphi_R) + \varphi_R \quad ; \quad \alpha_R := \left(\frac{1 - \varphi_R}{\varphi_R}\right) \quad ; \quad \bar{\xi} = \left(\frac{\xi - x}{x_2 - x}\right) \tag{6}$$

Numerical methods for solving fuzzy differential equations

Differential equations are usually used to model real natural phenomena. The uncertainty about this kind of models in its reflection of reality can hardly be quantified, but the uncertainty of the measured data that give an account of the regarded phenomenon, can still be integrated into the model. Measured data appear in such models as initial and boundary conditions as well as in the parameters that control the behaviour of the model. The fuzzy possibilistic interpolating functions, suggested above, are the basis of possibilistic numerical methods, such as a fuzzy possibilistic FVM. Appling such a numerical method to solve a fuzzy initial and boundary problem will result in a fuzzy approximate solution. Consequently, the uncertainty that results from impreciseness in the initial and/or boundary conditions, or in the model parameters, can be integrated in the resulting solution. Moreover, the fuzziness resulting from the discretisation of the domain of the problem can also be incorporated in the solution.

FUZZY POSSIBILISTIC DIGITAL TEMPORAL BATHYMETRIC MODEL

Bathymetric data resulting from several surveying campaigns with different survey goals are the basic data sets that are used in building Digital Bathymetric Models. The fact that the measurements of the altitude of the seafloor are taken at discrete and distributed points in time and space reduces the actual state of the measured continuous bathymetry. In addition to this inadequate representation of the seafloor, the genuine impreciseness in the basic bathymetric data resulting from errors and inaccuracy of the acquisition method introduces, if neglected, an uncertainty in the bathymetric model. Moreover, as bathymetric models are based on data sets and the associated interpolation procedures, the assumption that the function resulting from the interpolation procedure represents the real morphology of the seafloor is flawed with great uncertainty.

Figures 2, 3 and 4 show in plan view the fuzzy spatial-temporal interpolation of the altitude of the seafloor in the summer of the year 2002 for a study area that covers the near shore and surf zones north of the “Pirolatal” (Island of Langeoog, North Sea, Germany). In Figure 2 the least possible and most necessary distributions of the minimum $\underline{z}^0$ and the maximum $\overline{z}^0$ altitude of the seafloor with the possibility level of α = 0.0 are shown. Figure 3 shows the distributions of the minimum $\underline{z}^{0.5}$ and the maximum $\overline{z}^{0.5}$ altitude of the seafloor with the possibility level of α = 0.5. Figure 4 shows the most possible and least necessary distribution of the altitude of the seafloor m_z with the possibility level of α = 1.0, where m_z is the modal value of the fuzzy altitude $\tilde{z}(x, y, t)$.

In this academic case study attention is paid to the spatial-temporal interpolation of fuzzy bathymetric data points with fuzzy altitudes ($\tilde{z}$) of the seafloor at crisp positions (x, y, t). The used interpolation procedure is a fuzzy bilinear interpolation method. The uncertainty induced by the interpolating function, the density of the data set and the characteristic of the investigated area are considered in the resulting fuzzy possibilistic digital temporal bathymetric model. The relationship between the data density and the characteristics of the study area was expressed in a characteristic-density factor, with which a modification of the resulting possibility distribution can be conducted. In the case of spatial-temporal interpolation methods one has to differentiate between two cases. The first one is the morphologic characteristic associated with the spatial density and the second one is the morphodynamic characteristic associated with the temporal density.

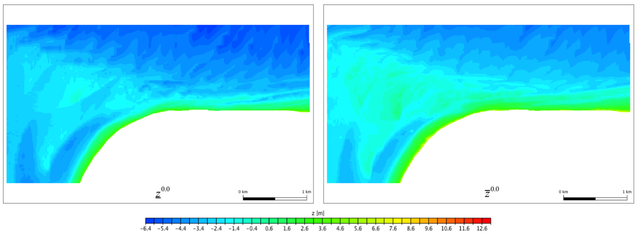

Figure 2. The minimum $\underline{z}^{0}$ and the maximum $\overline{z}^{0}$ altitude with the possibility level of 0.0

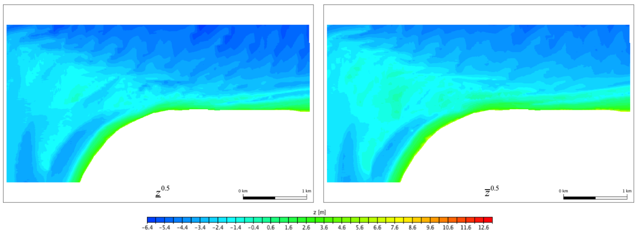

Figure 3. The minimum $\underline{z}^{0.5}$ and the maximum $\overline{z}^{0.5}$ altitude with the possibility level of 0.5 .

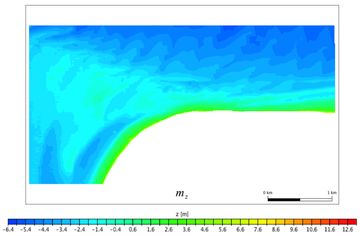

Figure 4. The modal value m_z of the fuzzy altitude with the possibility level of 1.0 .

FUZZY POSSIBILISITC GROUNDWATER MODEL

In the process of groundwater exploration, information should be collected in locations as close as possible and at times as frequent as feasible. This can be the case in an ideal situation. However, the inaccessibility of the geological formations of a groundwater system, the difficulties associated with the field investigations and considerations of cost make it impossible to collect thorough and densely distributed sets of data in space and time that reflect the real situation of the groundwater system. Moreover, the fact that all the measured quantities are the result of estimations based on subjective interpretations and experience results in a non-ignorable uncertainty in the collected data sets. In addition, uncertainty will be induced by the successive conceptualization and modelling of the groundwater system based on subjective interpolations and extrapolations.

Figure 5, Figure 6 and Figure 7 show the results of a simulation of the flow of groundwater and the transport of a contaminant in a hypothetical groundwater system with hypothetical fuzzy hydrogeological parameters and initial and boundary conditions. In Figure 5 the least possible and most necessary distribution of the maximum values of the hydraulic head $\overline{h}^0(t=8)$ and concentration $\overline{c}^0(t=8)$ with the possibility level of α = 0.0 are shown. Figure 6 shows the most possible and least necessary distribution of the values of the hydraulic head $m_h(t=8)$ and the concentration $m_c(t=8)$ with the possibility level of α = 1.0, where m_h and m_c refer to the modal values of the fuzzy functions $\tilde{h}(x,y,t)$ and $\tilde{c}(x,y,t)$, respectively. In Figure 7 the least possible and most necessary distribution of the minimum values of the hydraulic head $\underline{h}^0(t=8)$ and concentration $\underline{c}^0(t=8)$ with the possibility level of α = 0.0 are shown.

In this academic case study the flow in the system is assumed to be only horizontal as an example of a regional groundwater flow. The flow of groundwater and the transport of the contaminant are simulated by a numerical fuzzy model based on the fuzzy finite volume method. The time of the simulation is 15 days. Abstraction of groundwater from a well in the middle of the simulated region and injection of a contaminant at the upper corner of the region are assumed.

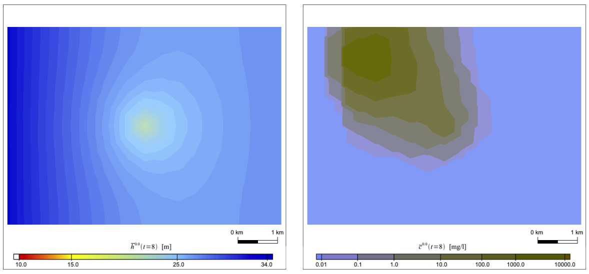

Figure 5. The distribution of $\overline{h}^0(t=8)$ and $\overline{c}^0(t=8)$ after eight days.

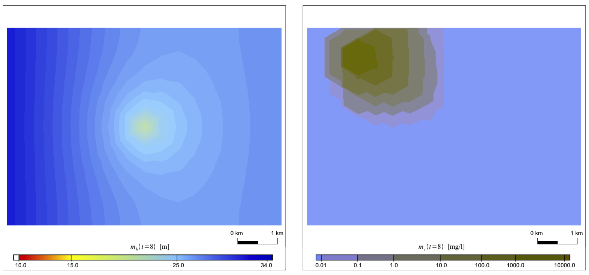

Figure 6. The distribution of $m_h(t=8)$ and $m_c(t=8)$ after eight days.

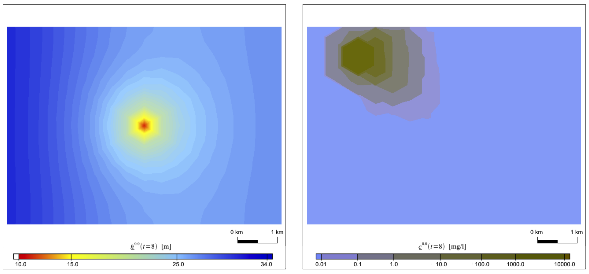

Figure 7. The distribution of $\underline{h}^0(t=8)$ and $\underline{c}^0(t=8)$ after eight days

CONCLUSION

In conclusion, using fuzzy sets to deal with partial belief problems and uncertainty requires the adoption of the concepts of possibility theory. Fuzzy sets in their possibility theoretic interpretation confine and restrict the uncertainty that might have been otherwise stemmed by omitting parts of important information, namely the degree of the lack of information and impreciseness. A computational framework of fuzzy numbers based on their interpretation as possibility distributions was presented. New fuzzy interpolation methods are used to meet the need of estimating the uncertainty resulting from the discrete and distributed state of the measured data sets as well as the intrinsic impreciseness in the measurements. On this basis fuzzy numerical methods for the solution of fuzzy initial and boundary value problems are also addressed.

REFERENCES

[1] Zadeh L., "Fuzzy sets", *Information and Control*, Vol. 8, No. 3, (1965), pp 338-353.
[2] Zadeh L., "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, Vol. 1, No. 1, (1978), pp 3-28.
[3] Dubois D. and Prade H., editors, "*Fundamentals of Fuzzy Sets. The Handbooks of Fuzzy Sets Series.",* Kluwer Academic Publishers, (2000).
[4] Zadeh L., "The concept of a linguistic variable and its application to approximate reasoning - I", *Information Sciences*, Vol. 8, No. 3, (1975), pp 199-249.
[5] Dong W. and Shah H., "Vertex method for computing functions of fuzzy variables", *Fuzzy Sets and Systems*, Vol. 24, No. 1, (1987), pp 65-78.