FUZZY NUMERICAL SIMULATION OF WATER QUALITY

J. GANOULIS, P. ANAGNOSTOPOULOS, H. MPIMPAS

Hydraulics Laboratory, Department of Civil Engineering,
Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
Tel. +30-2310-99.56.82 Fax +30-2310-99.56.81 e-mail iganouli@civil.auth.gr

ABSTRACT

In mathematical modelling of water quality, uncertainties are due to limited information on the values of physical parameters. In order to quantify such uncertainties in unsteady advective dispersion phenomena, dispersion coefficients are represented as fuzzy numbers. By combining representation of fuzzy numbers as a discrete set of h-level cuts with reliable finite difference numerical algorithms and applying the extension principle in interval arithmetics, it is shown that reliable fuzzy numerical simulations are obtained. This is concluded by comparing numerical simulations with analytical solutions of one-dimensional unsteady advective dispersion equation.

INTRODUCTION

Deterioration of surface and groundwater water quality from natural and anthropogenic causes is of growing concern in many parts of the world. Increased water pollution, caused by urban, industrial and agricultural activities, is directly related to the constant growth in the number of people living in river basins and coastal areas. Huge amounts of nitrogen and phosphorus are released from sewage and leaching of fertilizers resulting in algal blooming, eutrophication, shellfish die-off and loss of habitat, together with severe economic, social and ecological losses. The situation becomes critical in lakes and semi-enclosed seas such as the Mediterranean, where water exchange with the outer ocean is quite small.

Natural pollutant sources, such as rivers and ditches, may also induce critical pollution situations in river deltas and estuaries. How non-controlled fluctuations of pollutant loads and imprecise knowledge on the values of physical parameters affect water quality simulations is an important question to be investigated.

Uncertainties in water pollution are of two main types:

- *Aleatory* Uncertainties: These are due to natural variability or randomness
- *Epistemic* Uncertainties: Different man-induced uncertainties such as those in (a) input data, (b) modelling and (c) technological applications.

Deterministic modelling introduces sharp values of physical parameters and boundary conditions. This approach is not adequate to incorporate imprecision on data and propagate uncertainties.

Two main methodologies are appropriate to quantify uncertainties: (1) stochastic simulation and (2) fuzzy modelling.

According to stochastic simulation, physical parameters and input loads are considered as *random variables*. This is a frequency-based approach to propagate uncertainties. Results of probability theory (stochastic arithmetic) may be introduced in stochastic modelling in the form of analytical functional relationships between random variables, or by simulating a large number of different realizations (Monte Carlo method).

In this paper, the fuzzy set theory in combination with mathematical modelling is proposed in order to assess uncertainties in estimating environmental water pollution. First, uncertainties in input loads and values of physical parameters are introduced as *fuzzy numbers*, and then, uncertainties are propagated by using fuzzy calculus. With fuzzy mathematical modelling it is possible to assimilate imprecise data and, in the form of fuzzy numbers, directly produce imprecise output without repeating a large number of computations.

Examples of application in simplified and complicated real cases illustrate the capabilities of the above methodology and the precautions to be taken for its successful implementation in water pollution problems.

FUZZY MATHEMATICAL MODELLING

Fuzzy modelling has not yet been developed extensively, although fuzzy numbers and fuzzy relations have found many applications in control engineering and industrial devices. Fuzzy set theory (Zadeh, 1965; Klir and Folger, 1988; Zimmerman, 1991), and its derivative fuzzy arithmetic (Kaufmann and Gupta, 1985), may be used to directly introduce imprecise data into a mathematical model with minimal input data requirements (Ferson et al., 1994).

In fuzzy modelling, only the range and the most confident values of the input variables are required. This means it can be successfully used when the available data is too sparse for a probabilistic method to be applied (Ganoulis, 1994; Silvert, 1997).

CRISP WATER QUALITY MODELLING

Simulation of environmental water quality is based on the well known advection-dispersion mathematical model. The rate of change of the concentration of n different pollutants under biochemical interactions can be expressed in two dimensions as:

$$\frac{\partial c_k}{\partial t} + u\frac{\partial c_k}{\partial x} + v\frac{\partial c_k}{\partial y} = D_x \frac{\partial^2 c_k}{\partial x^2} + D_y \frac{\partial^2 c_k}{\partial y^2} + S_k \qquad k=1,2,..n \tag{3}$$

where:

c_k is the concentration for the k^{th} pollutant;

u and v are the water velocities in the x and y directions (m/s);

D_x and D_y are the dispersion coefficients in the x and y directions (m^2/s);

S_k is the source term describing the biochemical reactions.

MODEL FUZZIFICATION

Let us now assume that only limited information is available on the input pollutant loads (boundary conditions) and the value of the dispersion coefficients. This type of uncertainty can be taken into account by considering both input loads and dispersion coefficients as fuzzy numbers. This implies that unknown pollutant concentrations at any point x, and at any time t, will also behave as fuzzy numbers. They will still follow the advective dispersion partial differential equations, which are written as:

$$\frac{\partial \tilde{c}_k}{\partial t} + u\frac{\partial \tilde{c}_k}{\partial x} + v\frac{\partial \tilde{c}_k}{\partial y} = \tilde{D}_x \frac{\partial^2 \tilde{c}_k}{\partial x^2} + \tilde{D}_y \frac{\partial^2 \tilde{c}_k}{\partial y^2} + S_k \qquad k=1,2,...n \tag{4}$$

where the symbol $\tilde{}$ is used to denote a fuzzy variable.

Although derivatives of fuzzy variables exist, there is no unique solution to equation (4), because fuzzy numbers take different values at different levels of confidence (Appendix). If, for every confidence level h, we are looking only for the lower and upper limiting values of the unknown fuzzy variables, then the non-uniqueness problem may be resolved as follows. Every fuzzy number $\tilde{X}$ may be represented by a discrete set of h-level cuts $\overline{X}(h)$ (Kaufmann and Gupta, 1988, Appendix). These are ordinary intervals that for the fuzzy variables C and D should also follow the advective dispersion equations written as:

$$\frac{\partial \overline{c}_k(h)}{\partial t} + u\frac{\partial \overline{c}_k(h)}{\partial x} + v\frac{\partial \overline{c}_k(h)}{\partial y} = \overline{D}_x(h)\frac{\partial^2 \overset{\mathbf{r}}{c}_k(h)}{\partial x^2} + \overline{D}_y \frac{\partial^2 \overline{c}_k(h)}{\partial y^2} + S_k$$

$$k=1,2,\ldots n \qquad (5)$$

where the symbol $^{-}$ is used to denote an ordinary interval.

Applying finite differences or finite elements to equation (5), a system of interval equations needs to be solved. This is difficult from a mathematical point of view, and has stimulated a lot of interest, because whatever possible technique is used, only enclosures for the range of the output function can be produced (Hansen, 1969; Moore, 1979; Neumaier, 1990). Finding the best possible enclosure for an unknown interval function, which is defined as the “hull” of the solution (Fig. 1), is a fundamental problem of interval analysis. It should be treated with care, as the solution accuracy depends on the shape of the interval function (Rall, 1986).

Shafike (1994) has used the finite element method to simulate a groundwater flow model with fuzzy coefficients. The algebraic system of interval equations was solved with an iterative algorithm (Moore, 1979). Dou et al. (1995) applied the fuzzy set theory into a steady-state groundwater flow model with fuzzy parameters combined with the finite difference method. A non-linear optimization algorithm was used in order to apply the extension principle for the solution of groundwater flow equations with fuzzy numbers as coefficients for the hydraulic heads. Because of the non-uniqueness of the max and min values, this may lead to great inaccuracies.

Ganoulis et al. (1995; 1996) used fuzzy arithmetic to simulate imprecise relations in ecological risk assessment and management. For the solution of the algebraic system of equations with fuzzy coefficients, direct interval operations were employed, instead of the iterative methods or non-linear optimization techniques used in previous studies.

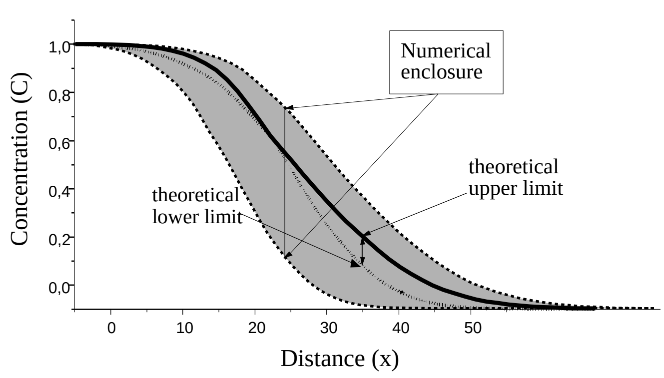


Figure 1: The “hull” of one-dimensional convective concentration with fuzzy dispersion coefficient and numerical enclosure at given confidence level.

FUZZY NUMERICAL SIMULATION

Finite differences and finite elements have been used to produce numerical enclosures of the solution of equation (5) (Ganoulis, 1994; Mpimpas, 1998). Enclosures may fit the exact range of the unknown function better, depending on two main factors:

1. The accuracy of the numerical algorithm, and
2. The correct application of the algebra of fuzzy arithmetic (Appendix)

The Eulerian-Lagrangian scheme based on the characteristics method has also been used. In this method the numerical integration of the parabolic part of equation (5) is conducted on the characteristics lines of the equation.

Equation (5) is split in two parts:

1. The hyperbolic part, which in one dimension is written as $dx/dt = u$, and
2. The parabolic part that takes the following form

$$d\overline{C}(h)/dt = \overline{D}(h)(d^2\overline{C}(h))/dx^2) \quad (6)$$

Using finite differences, a large number of particles are moved over a stationary grid to track the advection equation. At each moment in time $n\Delta t$ the position of each particle is x_p and its concentration is C_p. At the next moment in time $(n+1)\Delta t$ the new position of the particle becomes $x_p{}^{n+1} = x_p{}^n + u\Delta t$, and the change of concentration at the nodal points, due to the dispersion part (6), is computed by an explicit finite difference algorithm as follows:

$$\overline{C}_i^{n+1}(h) = \overline{C}_i^n(h) + \Delta t\,\overline{D}(h)\,(\overline{C}_{i-1}^n(h) - 2\overline{C}_i^n(h) + \overline{C}_{i+1}^n(h)) \quad (7)$$

The new particle concentrations are evaluated by adding the change due to the diffusive part (7). Then, the new nodal concentrations are calculated by accounting the corresponding number of particle concentrations.

EXAMPLES OF APPLICATION

First the effectiveness of the above Lagrangian-Eulerian algorithm in terms of numerical accuracy and precision was checked. When the dispersion coefficient takes deterministic values, and a point pollutant load is considered at the origin of coordinates with initial and boundary conditions written as:

$C=0$	$t=0,\ x>0$
$C_0=1$	$t>0,\ x=0,$

the analytical solution takes the form

$$C = C_o + \frac{1}{2}\{1 - \mathrm{erf}\,(\frac{x - Ut}{2\sqrt{Dt}})\} \quad (8)$$

As shown in Fig.2, numerical results are very close to the analytical solution. The same numerical algorithm has given satisfactory results in simple two-dimensional cases, where the analytical solution is also known (Mpimpas, 1998).

At confidence level h=0 and D(h=0)=[0.1, 0.5], comparison between numerical enclosure and the output concentration range is shown in Fig.3. For this computation, particular care should be taken with the application of expression (7). This is due to the fact that in this expression, the second derivative of the output concentration takes both positive and negative values. Interval multiplication should be done according to the max and min rules and also respect the subdistributivity enclosure (Appendix). Near the inflexion point, where the second derivative is zero, numerical imprecision reaches a maximum (Fig. 3).

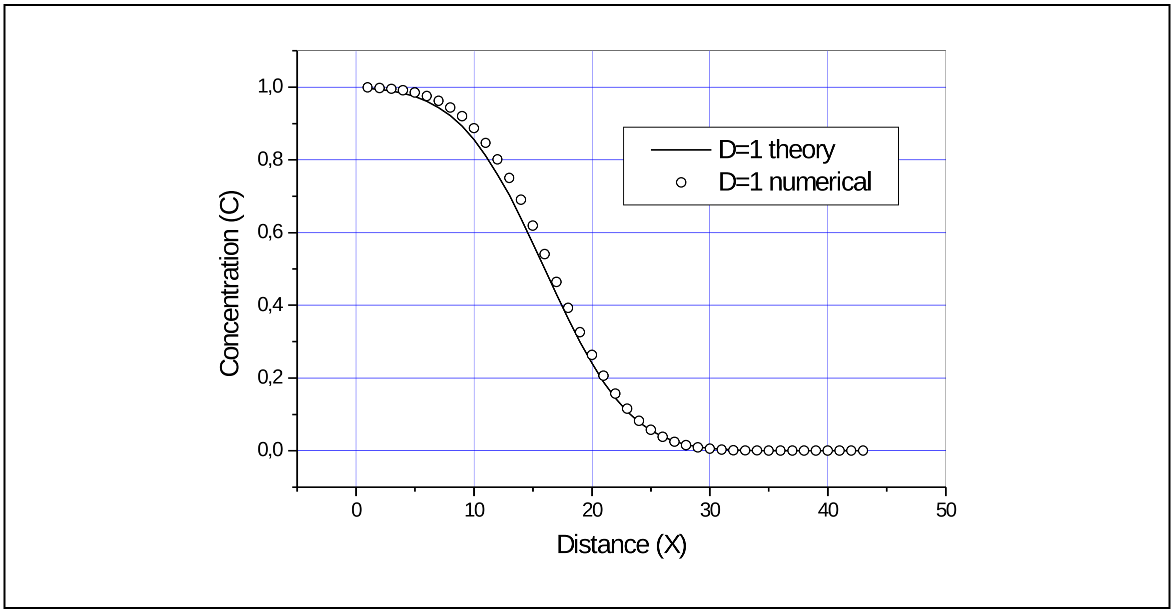


Figure 2: Comparison between numerical simulation and analytical solution of the one dimensional convective dispersion equation (D=1, Δx=1m, Δt=1s, u=1m/s, Pe=u Δx/D=1, t=16s) .

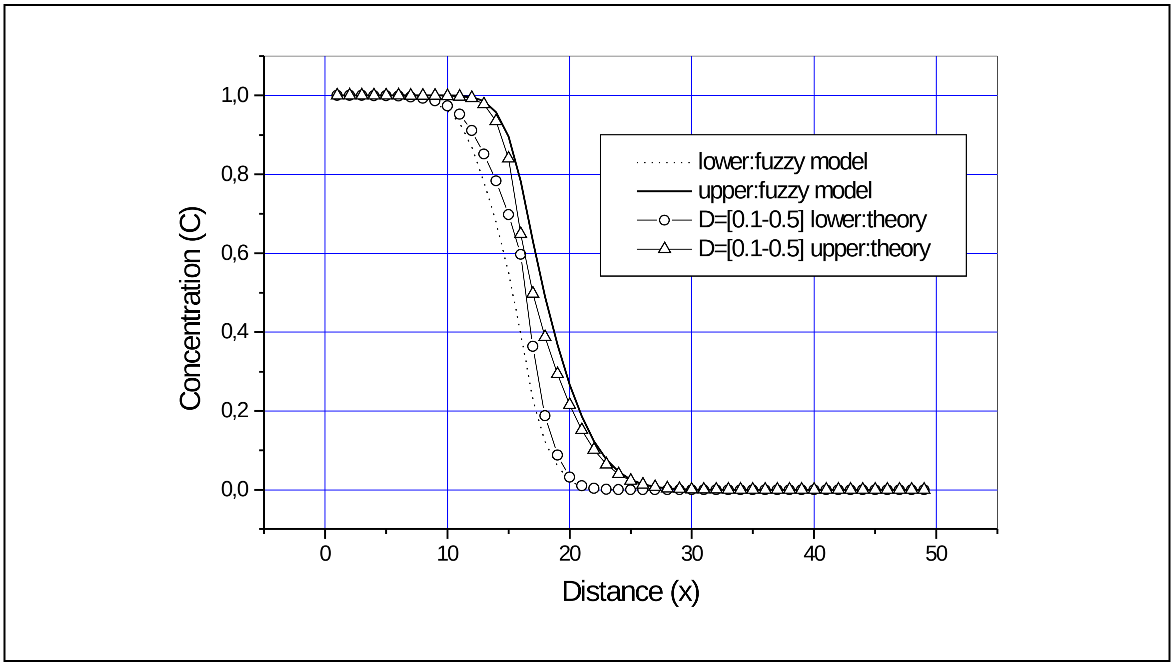


Figure 3: Comparison between theoretical concentration range and numerical fuzzy enclosure (Δx=1m, Δt=1s, u=1m/s, t=16s).

In order to further validate the accuracy of the proposed methodology, another unsteady, one-dimensional case was studied, for which the analytical solution is known (Mpimpas, 1998). This is the case of the one-dimensional convective dispersion equation (5) with a supplementary term at the right side representing a linear concentration decay of the form $-\tilde{k}\,c$, where $\tilde{k}$ is a fuzzy decay coefficient. The analytical solution has the following form

$$c = \frac{c_0}{2}\left[e^{T_1}[1 - \mathrm{erf}(S_1)] + e^{T_2}[1 - \mathrm{erf}(S_2)]\right] \quad (9)$$

with

$$F = \sqrt{u^2 + 4kD_x},\ S_1 = \frac{x - Ft}{\sqrt{4D_x t}},\ S_2 = \frac{x + Ft}{\sqrt{4D_x t}},\ T_1 = \frac{u - F}{2D_x}x,\ T_2 = \frac{u + F}{2D_x}x$$

Taking $\tilde{k}$ as a triangular fuzzy number ($1\cdot10^{-5}$, $3\cdot10^{-5}$, $5\cdot10^{-5}$), and using interval representations of fuzzy numbers for h=0, 0.25, 0.5, 0.75, comparison between numerical simulation and analytical solution shown in Figures 4 and 5 indicate a very good accuracy of the proposed methodology in unsteady water pollution fuzzy simulation.

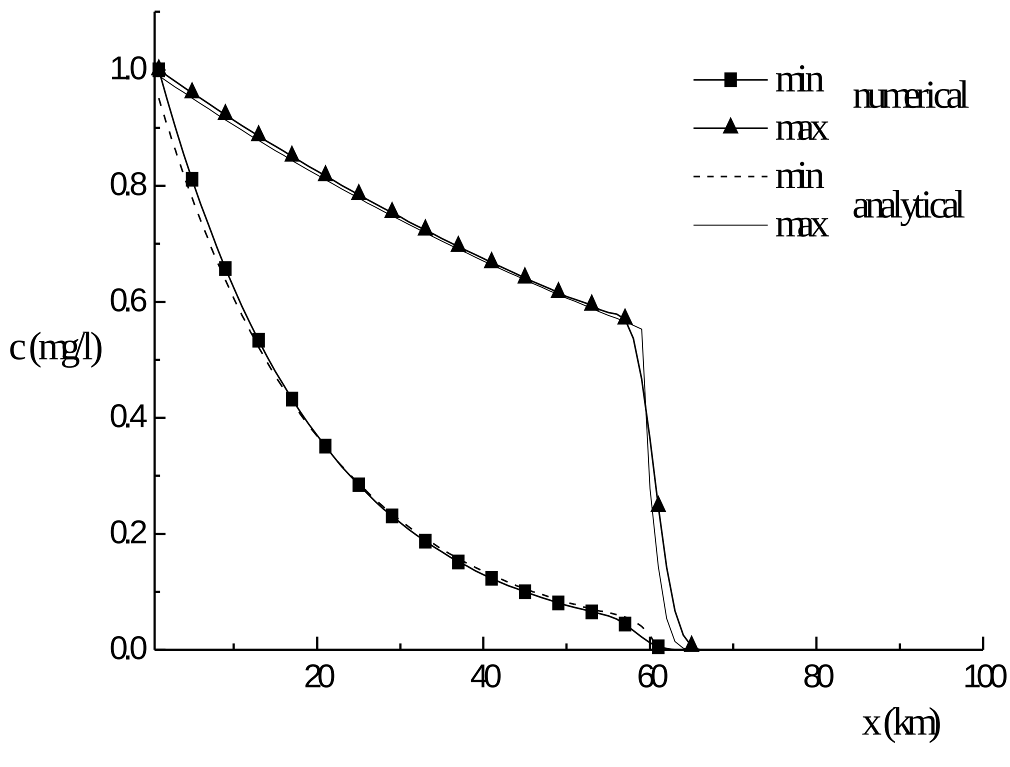
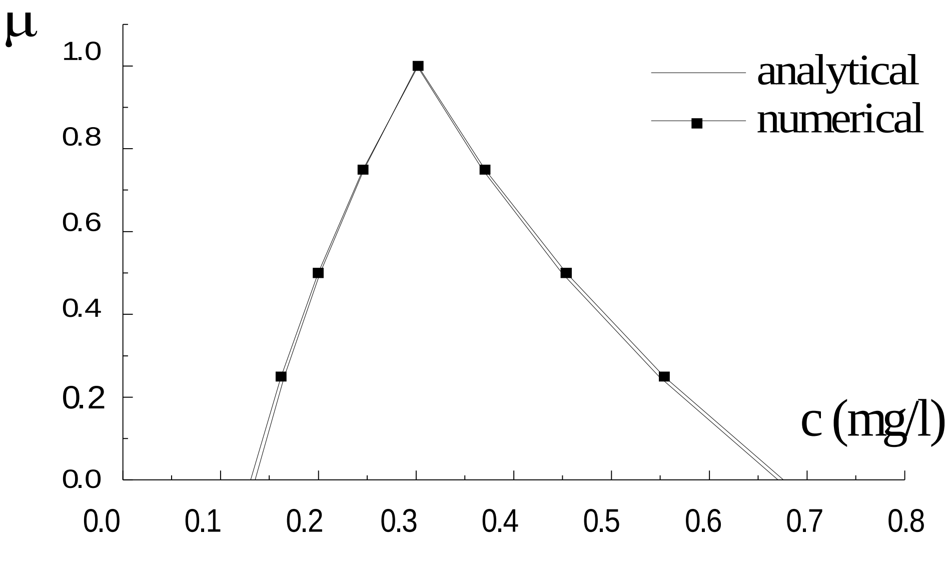
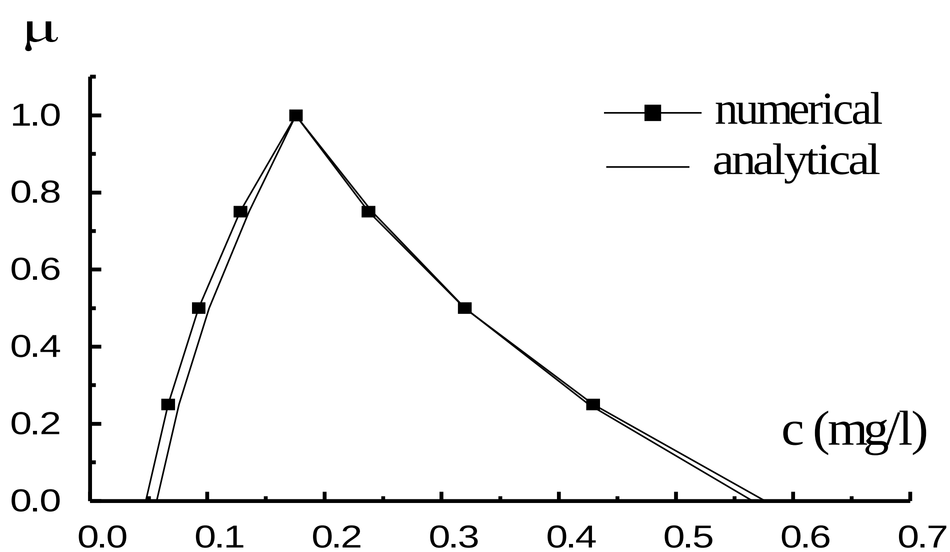


Figure 4: Variation of interval concentrations $\bar{c}$(h) for h=0, t=16h: comparison between numerical and analytical solutions.



(A)



(b)

Figure 5: Fuzzy concentrations $\tilde{c}$: (a) at x=40 km and (b) at x=57 km, t=16h.

In real geophysical flows, as in coastal circulation in the Bay of Thermaikos, a finite element grid can better describe irregular boundaries. The characteristic Galerkin numerical algorithm was applied for fuzzy simulation of coastal water quality. Velocities of currents were computed using a two-dimensional coastal circulation model (Mpimpas, 1998). Numerical simulation of output concentration of different pollutants at different locations within the bay in the form of fuzzy numbers (Mpimpas et al., 2001) incorporates imprecision, due to limited information on land-based pollutant sources and the value of the dispersion coefficients.

CONCLUSIONS

Fuzzy modelling is a methodology that can produce the output uncertainties in form of fuzzy numbers. Accuracy and precision of fuzzy logic-based numerical simulations depend on the quality of the numerical algorithm as well as on the proper application of rules of fuzzy calculus as related to the extension principle.

For the general unsteady convective dispersion model, a methodology is presented in this paper for an effective fuzzy logic–based simulation of water quality. The methodology is based on representing fuzzy numbers as discrete sets of h-level cuts, the use of reliable Lagrangien-Eulerian finite difference numerical algorithms and application of the extension principle in interval arithmetics.

Such a methodology is very useful when boundary pollutant loads and model parameters are inadequately known and a sensitivity analysis is necessary.. For example, in order to simulate the output range of the solution due to the variation range of *s* parameters, one needs to repeat at least 2^s computations. With the reliable methodology proposed here, one simulation is sufficient for quantifying the range of uncertainty in the output.

APPENDIX : FUZZY NUMBERS AND FUZZY ARITHMETIC

A fuzzy number $\tilde{X}$ may be formally defined as a set of ordered pairs

$$\tilde{X} = \{(x, \mu_{\tilde{X}}(x)) : x \in R; \mu_{\tilde{X}}(x) \in [0, 1]\} \qquad (A.1)$$

where x is a particular value of $\tilde{X}$ and $\mu_{\tilde{X}}(x)$ represents its membership function. Values of the membership function are located in the closed interval [0,1]. The closer $\mu_{\tilde{X}}(x)$ is to 1, the more "certain" one is about the value of x. A fuzzy number $\tilde{X}$ is normal and convex when its membership function takes one maximum value equal to 1 and always increases to the left of the peak, and decreases to the right.

The simplest type of fuzzy number is the triangular, that is one having linear membership functions on either side of the peak. A fuzzy triangular number can be characterized by three real numbers: two values of x i.e.x_1, x_2 where the membership function reaches zero, and one value $x_3$where it reaches a value of 1.

A triangular fuzzy number (TFN) may be described by the values of x at points x_1, x_2 and, i.e. $\tilde{X} = (x_1, x_2, x_3)$

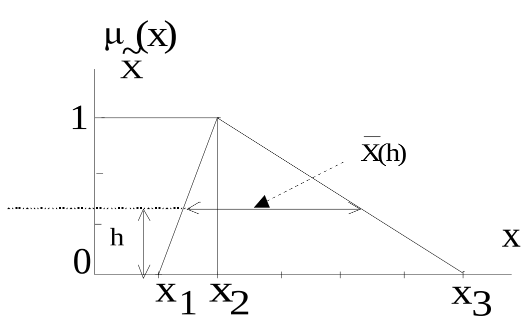


Figure A.1: A triangular fuzzy number $\tilde{X}=(x_1, x_2, x_3)$.

The h-level set of a fuzzy number $\tilde{X}$ (Fig. A.1) is the ordinary set or interval $\overline{X}(h)$, defined as

$$\overline{X}(h)=\{x: \mu_{\tilde{X}}(x) \geq h)\} \qquad \text{(A.2)}$$

Let us consider two triangular fuzzy numbers $\tilde{A}$ and $\tilde{B}$ given by the triplets $\tilde{A}=(a_1, a_2, a_3)$ and $\tilde{B}=(b_1, b_2, b_3)$. We have

i) *Addition:* $\tilde{A}+\tilde{B} = (a_1+b_1, a_2+b_2, a_3+b_3)$

ii) *Subtraction:* $\tilde{A}-\tilde{B}=(a_1-b_3, a_2-b_2, a_3-b_1)$

The multiplication or division of two fuzzy numbers does not always produce a fuzzy number. These operations can be defined as follows:

iii) Multiplication

$$\tilde{A}*\tilde{B}=[\min(a_1*b_1, a_1*b_3, a_3*b_1, a_3*b_3), a_2*b_2, \max(a_1*b_1, a_1*b_3, a_3*b_1, a_3*b_3)]$$

iv) *Division*

$$\tilde{A}/\tilde{B}=[\min(\frac{a_1}{b_1}, \frac{a_1}{b_3}, \frac{a_3}{b_1}, \frac{a_3}{b_3}), \frac{a_2}{b_2}, \max(\frac{a_1}{b_1}, \frac{a_1}{b_3}, \frac{a_3}{b_1}, \frac{a_3}{b_3})]$$

<u>Algebraic Properties of Fuzzy Numbers</u>

Let us assume that $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ are fuzzy numbers. The following laws hold:

Cummutativity $\quad \tilde{A}+\tilde{B}=\tilde{B}+\tilde{A} \qquad \tilde{A}\,\tilde{B}=\tilde{B}\,\tilde{A}$

Associativity $\quad (\tilde{A}+\tilde{B})\pm\tilde{C} = \tilde{A} +(\tilde{B}\pm\tilde{C}) \qquad (\tilde{A}\,\tilde{B})\tilde{C}=\tilde{A}(\tilde{B}\,\tilde{C})$

However, subdistributivity and subcancellation are not always valid. In particular we have

Subdistribution

$$\tilde{A}(\tilde{B}\pm\tilde{C}) \subseteq \tilde{A}\,\tilde{B} \pm \tilde{A}\,\tilde{C} \qquad (\tilde{A}\pm\tilde{B})\tilde{C} \subseteq \tilde{A}\,\tilde{C} \pm \tilde{B}\,\tilde{C}$$

Subcancellation

$$\tilde{A}-\tilde{B} \subseteq (\tilde{A}+\tilde{C})-(\tilde{B}+\tilde{C}) \qquad \tilde{A}/\tilde{B} \subseteq \tilde{A}\,\tilde{C}/\tilde{B}\,\tilde{C}$$

Overestimation usually occurs because of the failure of the distributive and cancellation laws.

REFERENCES

Dou, C., Woldt, W., Bogardi, I. and Dahab, M., 1995. Steady State Groundwater Flow Simulation with Imprecise Parameters. Water Resour. Res., 31: 2709-2719.

Duckstein, L., A. Bardossy, T. Barry and I.Bogardi, 1990. Health Risk Assessment under Uncertainty: a Fuzzy-Risk Methodology In: Y. Haimes, D. Moser and E. Stakhiv (Eds.), Risk Based Decision Making in Water Resources V, ASCE, NY.

Ferson, S., Millstein, J. and Mauriello, D., 1994. Environmental Risk Assessment Using Interval Analysis and Fuzzy Arithmetic. In: Association with 1994 Annual Meeting of the Society for Risk Analysis, Baltimore, Maryland.

Ganoulis, J. 1994. Risk Analysis of Water Pollution, WILEY-VCH, Weinheim, Cambridge, New York, Tokyo, 306 pp.

Ganoulis, J., Mpimpas, H., and P. Anagnostopoulos, 1995. Coastal Water Assessment Using Fuzzy Arithmetic in Proc. IEEE Workshop on Non-linear Signal Image Processing, Vol. II, 1015-1018.

Ganoulis, J., Mpimpas, H., Duckstein, L. and Bogardi, I., 1996. Fuzzy Arithmetic for Ecological Risk Management. In: Y. Haimes, D. Moser and E. Stakhin (Eds.), Risk Based Decision Making in Water Resources VII, ASCE, NY, pp. 401-415.

Ganoulis, J., P. Anagnostopoulos, and Mpimpas, H., 2001. Fuzzy Logic-based Risk Analysis of Water Pollution In Proc. XXIX IAHR Congress, Theme B, pp. 169-175, Beijing, China

Hansen, E., 1969. Topics in Interval Analysis. Claperton Press, Oxford.

Kaufmann, A. and Gupta, M., 1985. Introduction to Fuzzy Arithmetic: Theory and Applications. Van Nostrand Reinhold, New York.

Klir, G. and Folger, T., 1988. Fuzzy Sets, Uncertainty and Information. Prentice Hall.

Moore, R.E., 1979. Methods and Applications of Interval Analysis. SIAM, Philadelphia, PA.

Mpimpas, H. (1998) Numerical Investigation of Pollutant Dispersion Using Fuzzy Logic. Ph.D. Thesis, AUTH, Greece (in Greek).

Mpimpas H, Anagnostopoulos P., and Ganoulis J. (2001) Modelling of Water Pollution in the Thermaikos Gulf with Fuzzy Parameters, Ecological Modelling, 142, 91-104

Neumaier, A., 1990. Interval Methods for Systems of Equations. Cambridge University Press, Cambridge.

Rall, L.B, 1986. Improved Interval Bounds for Range of Functions. In: Interval Mathematics 1985, Springer-Verlag, Berlin.

Shafike, N.G., 1994. Groundwater Flow Simulations and Management under Imprecise Parameters. Ph.D thesis, Dept. of Hydrology and Water Resources, University of Arizona, Tucson, Arizona.

Shresta, B.P., K.R. Reddy and L. Duckstein (1990) Fuzzy Reliability inHydraulics. In: Proc. First Int. Symp. on Uncertainty Mod. and Analysis, Univ. of Maryland, College Park

Silvert, W., 1997. Ecological Impact Classification with Fuzzy Sets. Ecol. Model., 96: 1-10.

Zadeh, L.A., 1965. Fuzzy Sets. Information and Control, 8: 338-353.

Zimmerman, H.J., 1991. Fuzzy Set Theory and its Applications (2nd Ed.). Kluwer Academic Publishers, Dordrecht, The Netherlands.