Uncertainty of model parameters in stream pollution using fuzzy arithmetic

H. Mpimpas, P. Anagnostopoulos and J. Ganoulis

ABSTRACT

Fuzzy arithmetic is employed for the analysis of uncertainties in water-stream pollution, when the various model parameters involved are imprecise. The one-dimensional advection–dispersion equation, for both a conservative and a non-conservative substance, was solved analytically for point and Gaussian-hill input loads of pollution, considering the dispersion and decay coefficients involved as fuzzy numbers. The solution of the advection–dispersion equation was also conducted numerically for the same input loads with the finite-difference method, employing a Lagrangian–Eulerian scheme. The good agreement between analytical and numerical results presented in the form of fuzzy numbers confirms the reliability of the numerical scheme. The advection–dispersion equation of a non-conservative substance was then solved numerically for ten different water quality parameters, in order to study the water pollution in a water stream. The dispersion coefficient, the source terms and the input loads were expressed as fuzzy numbers, and the concentration of each quality parameter was obtained in fuzzy-number form. With fuzzy modeling, imprecise data can be represented and imprecise output produced, with minimal input data requirements and without the need of a large number of computations.

Key words | advection–dispersion, fuzzy number, input load, interval number, quality parameter, water pollution

H. Mpimpas
P. Anagnostopoulos (corresponding author)
J. Ganoulis
Department of Civil Engineering,
Division of Hydraulics and Environmental Engineering,
Aristotle University of Thessaloniki,
Thessaloniki 54124,
Greece
Tel.: +30 2310 995675
Fax: +30 2310 995680
E-mail: *anagnost@civil.auth.gr*

INTRODUCTION

The problems associated with water quality of rivers and coastal areas are of growing concern, since the increasing pollution in these areas has serious environmental impact. Numerical models have been extensively used to predict water quality, and to provide reliable tools for water quality management of rivers and coastal zones. However, uncertainties in water pollution models are introduced by the imprecise data related to the values of both the input loads of pollution and the biochemical and physical model parameters. The values of these parameters depend on factors which change randomly in space and time, like turbulence, temperature and composition of the organic mass. The vagueness in the values of these coefficients may also arise from indirect measurements and the limited number of samples (Ducstein & Plate 1987; Ganoulis 1991). For these reasons, the use of the deterministic approach, where values of the input loads and the biochemical coefficients are expressed in the form of sharp (or crisp) numbers, is not adequate.

Various probabilistic methods, like the Monte Carlo simulation or the Bayesian approach, have been used for the study of uncertainties in water pollution problems. These methods consider the input loads of pollution and the biochemical parameters as random variables. Guymon & Yen (1990) applied a deterministic–probabilistic model for regional groundwater flow with uncertain parameters, based on a two-layer deterministic model cascaded with a two-point probability method. Compared with the Monte Carlo simulation, this method gives valid results for linear problems where the variation of the coefficients of the uncertain parameters is small. Plate (1991) described a model for river

doi: 10.2166/hydro.2008.037

pollution with point and non-point pollutant sources considered as random variables. Perk (1996) presented a water quality model for the removal of nutrients from a river; the model's eight parameters were calibrated by means of the Monte Carlo method using field data. Numerical simulations were conducted for the 390, 625 possible combinations of the various parameters. It is important to mention that stochastic methods require a large number of data, as well as information about the entire probabilistic distribution for all input variables contained in the mathematical model, and a great number of simulations (Ang & Tang 1984).

On the other hand, fuzzy set theory (Zadeh 1965; Klir & Folger 1988; Zimmerman 1991) and its derivative fuzzy arithmetic (Kaufmann & Gupta 1985), may be used in order to introduce imprecise data into a mathematical model in a direct way with minimal input data requirements. In fuzzy modeling only the range and the most confident values of the input variables are required, so it can be used successfully when the available data is too sparse for a probabilistic method to be applied (Ganoulis 1994; Silvert 1997).

Shafike (1994) introduced the fuzzy set theory coupled with the finite-element method into a groundwater flow model. The algebraic system of equations with fuzzy coefficients was solved with an iterative algorithm by Moore (1979). Dou *et al.* (1995) applied the fuzzy set theory to a steady-state groundwater flow model with fuzzy parameters combined with the finite-difference method. A nonlinear optimization algorithm was used for the solution of the groundwater flow equations, with fuzzy numbers as coefficients for the hydraulic heads. Ganoulis *et al.* (1995) used fuzzy arithmetic to simulate imprecise data and model coefficients uncertainty in ecological risk assessment and management. This technique was applied to a simplified domain with coastal circulation in order to evaluate the risk of coastal pollution.

In the present study the solution of the one-dimensional advection–dispersion equation combined with fuzzy arithmetic was obtained, when the various model parameters involved are imprecise. The one-dimensional advection–dispersion equation, for both a conservative and a non-conservative substance, was solved analytically for point and Gaussian-hill input pollutant loads. The dispersion and decay coefficients were expressed as fuzzy numbers. The solution of the advection–dispersion equation was then conducted numerically with the finite-difference method. A Lagrangian–Eulerian model was applied in order to suppress the numerical diffusion introduced by the convective terms. Both numerical and analytical results are presented in the form of fuzzy numbers, and their comparison confirms the reliability of the numerical scheme.

This technique was used for the solution of a one-dimensional water pollution model that contained ten different water quality variables, namely: chlorophyll-a, coliforms, organic nitrogen, ammonia nitrogen, nitrite nitrogen, nitrate nitrogen, organic and inorganic phosphorus, BOD and dissolved oxygen deficit. The dispersion coefficient, the source terms and the input loads were considered as fuzzy numbers, and the concentration of each quality parameter was obtained in fuzzy-number form.

THE MATHEMATICAL MODEL AND THE SOLUTION PROCEDURE

The one-dimensional advection–dispersion equation of a non-conservative substance with fuzzy parameters is written as

$$\frac{\partial \tilde{c}}{\partial t} + u\frac{\partial \tilde{c}}{\partial x} = \tilde{D}_x \frac{\partial^2 \tilde{c}}{\partial x^2} - \tilde{k}\,\tilde{c} \tag{1}$$

where $\tilde{c}$ is the fuzzy concentration of a pollutant, u is the water velocity in the x direction (m/s), $\tilde{D}_x$ is the fuzzy dispersion coefficient in the x direction (m^2/s) and $\tilde{k}$ is a fuzzy decay coefficient. For $\tilde{k} = 0$ the equation corresponds to a conservative substance.

An explicit finite-difference scheme was used for the solution of Equation (1). To overcome the numerical diffusion–dispersion introduced by the advection term and improve the solution accuracy, a Lagrangian–Eulerian method was employed in the finite-difference formulation (Ganoulis 1994). The hyperbolic part (convection term) and the parabolic part (diffusion term) are treated independently. Initially, a large number of particles of equal concentrations is introduced in the computational domain, in such a way that the initial concentration is correctly represented. At each time level the solution of the advection part of Equation (1) is accomplished from the convection of each particle. The finite-difference method is then used for the solution of the diffusion part of the equation.

At the time level n the position of a particle is x_p and its concentration $\tilde{c}_p$. At the next time level $n+1$ the new position of the particle is $x_p^{n+1} = x_p^n + u\Delta t$. The concentration at node i is computed as the sum of concentrations of the m particles which correspond to node i as $\tilde{c}_i^{n+1} = \sum_{p=1}^{m} \tilde{c}_p^{n+1}$. The change of concentration in the interval Δt between levels n and $n+1$ due to the diffusion part, $\Delta\tilde{c}_i^n$, is derived from

$$\Delta\tilde{c}_i^n = \tilde{D}_x(\tilde{c}_{i+1}^n + \tilde{c}_{i-1}^n - 2\tilde{c}_i^n)\Delta t/\Delta x^2 - \tilde{k}\,\tilde{c}_i^n\Delta t. \quad (2)$$

The concentration difference $\Delta\tilde{c}_i^n$ is distributed equally amongst the m particles corresponding to node i. In this way the concentration difference of each particle, $\Delta\tilde{c}_p^n = \Delta\tilde{c}_i^n/m$, is obtained. The particle concentration at time level $n+1$ is $\tilde{c}_p^{n+1} = \tilde{c}_p^n + \Delta\tilde{c}_p^n$.

Since the model coefficients are considered as fuzzy numbers the output concentrations $\tilde{c}_i$ are also fuzzy; thus the algebraic Equation (2) becomes a fuzzy equation. The solution of a fuzzy equation can be performed with the use of fuzzy arithmetic (Appendix A). Since triplets cannot be used for the multiplication and division operations as explained by Kaufmann & Gupta (1985), the mathematical operations should be performed at various α-level cuts (see Appendix A and Figure 1), by the use of the confidence interval at each α-level (Dou *et al.* 1995). It is also important to mention that, from the solution of an interval equation using interval operations (Appendix B), only enclosures for the range of the output function can be produced (Hansen 1969; Moore 1979; Neumaier 1990; Ganoulis *et al.* 2003). The estimation of the best possible enclosure, which is called the "hull" of the solution, is a fundamental problem of Interval Analysis (Rall 1986). This is related to the criterion used to define the reliability limits of a given simulation and has no unique solution. For the

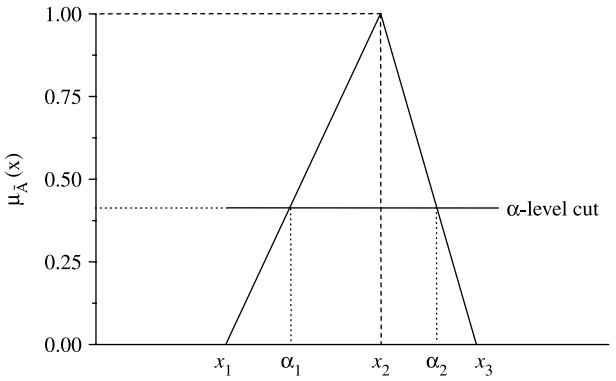

Figure 1 | A triangular fuzzy number $\tilde{A}$ defined by the triplet (x_1, x_2, x_3). The α-level cut defining the interval number $\overline{A} = [\alpha_1, \alpha_2]$ is also shown.

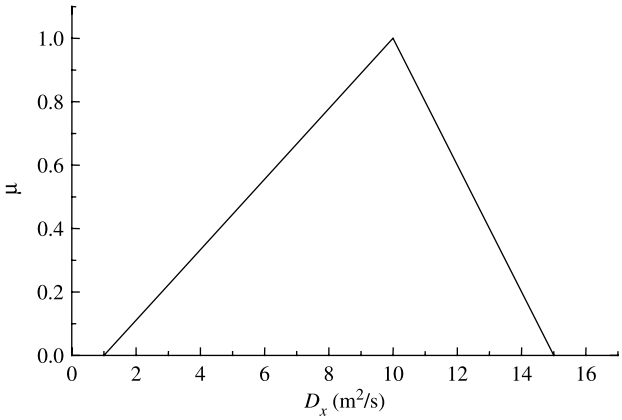

Figure 2 | The fuzzy dispersion coefficient $\tilde{D}_x = (1, 10, 15)$.

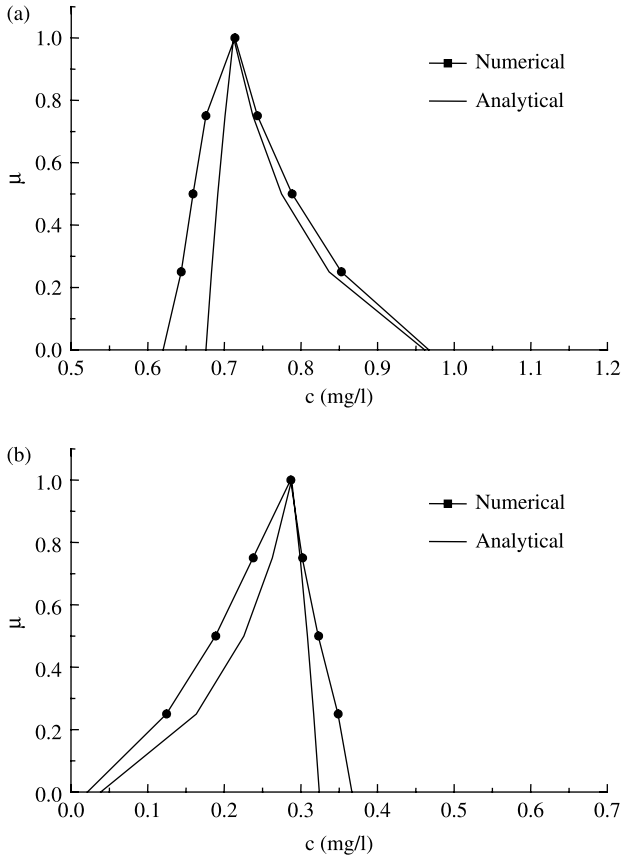

Figure 3 | Fuzzy concentration derived both analytically and numerically for a point-source input load, at $t = 10{,}000$ s. (a) $x = 4{,}750$ m; (b) $x = 5{,}250$ m. The solution corresponds to a conservative substance.

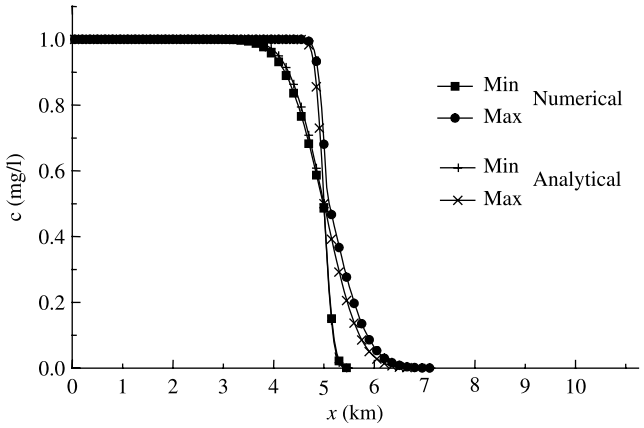

Figure 4 | Comparison between analytical and numerical results of concentration along the x axis at α-level cut = 0 for $t = 10{,}000$ s for a point-source input load. The solution is for a conservative substance.

advection–dispersion equation the resulting concentration distribution is a function of the model parameters, and since these parameters are interval numbers, the function is also interval. Thus the fuzzy Equation (2) becomes an interval equation at each membership α-level for each node i:

$$\Delta\bar{c}_i^n = \bar{D}_x(\bar{c}_{i+1}^n + \bar{c}_{i-1}^n - 2\bar{c}_i^n)\Delta t/\Delta x^2 - \bar{k}\,\bar{c}_i^n\Delta t \qquad (3)$$

where $\bar{c}$ means interval concentration.

Many methods have been proposed for the solution of interval equations. Iterative methods (Moore 1979) are simple but normally give solutions much wider than the hull, although they do contain the hull. In the present study direct interval operations (Appendix B) were employed, instead of the iterative methods or nonlinear optimization

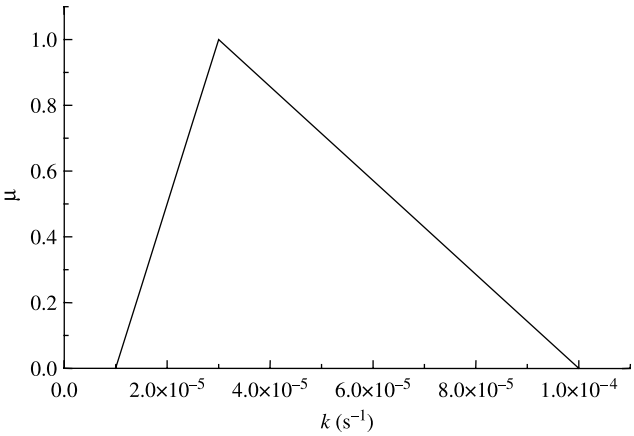

Figure 5 | Fuzzy decay coefficient $\bar{k} = (1 \times 10^{-5}, 3 \times 10^{-5}, 1 \times 10^{-4})$.

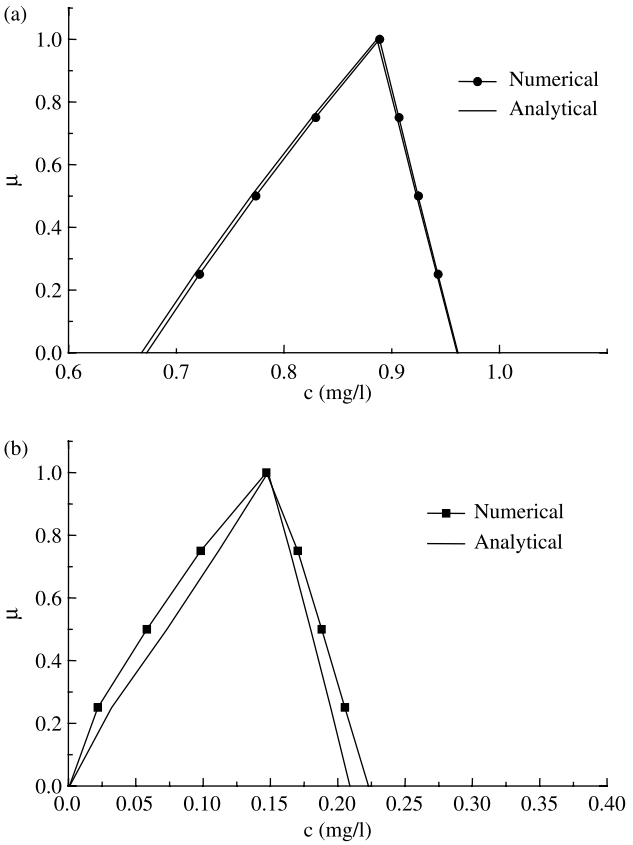

Figure 6 | Fuzzy concentration derived both analytically and numerically for a point-source input load at $t = 10{,}000$ s. (a) $x = 4{,}000$ m; (b) $x = 5{,}250$ m. The solution corresponds to a non-conservative substance.

techniques. For the solution of Equation (3) the interval calculations were carried out at five different membership α-level cuts (0, 0.25, 0.50, 0.75 and 1) of fuzzy numbers (Nickel 1986; Neumaier 1990). At each membership level of

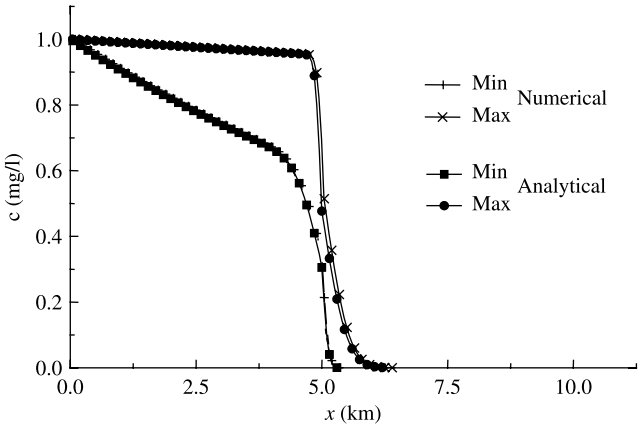

Figure 7 | Concentration derived both analytically and numerically along the x axis at α-level cut = 0 for $t = 10{,}000$ s for a non-conservative point-source input load.

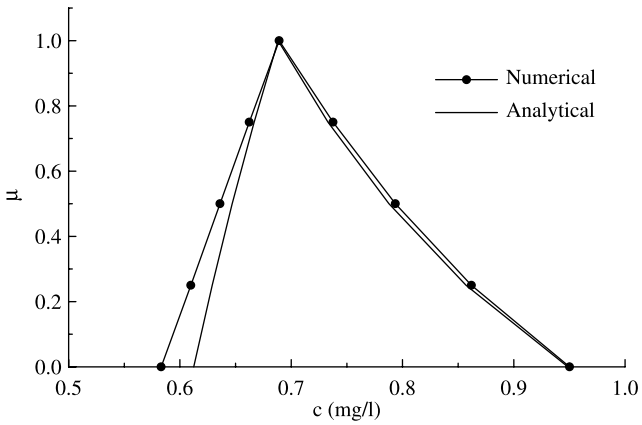

Figure 8 | Fuzzy concentration derived both analytically and numerically for a Gaussian-hill input load at $x = 6{,}500$ m for $t = 5{,}000$ s. The solution is for a conservative substance.

the model parameters an interval value of concentration was obtained. From the results at all membership levels considered, the fuzzy concentrations were estimated. It is important to mention that the lower (minimum) and upper (maximum) bounds of the values of concentrations cannot be calculated by using simply the lower and upper values of the coefficients of the model parameters.

ANALYTICAL VERSUS NUMERICAL RESULTS FOR SIMPLE TYPES OF INPUT POLLUTION LOADS

To validate the numerical technique the computational results were compared to those obtained analytically for two simple types of input pollution loads, namely:

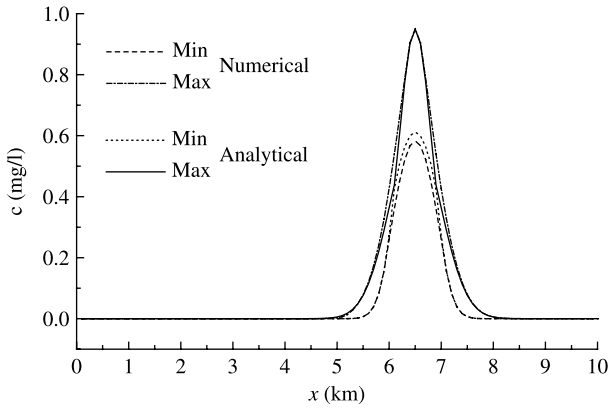

Figure 9 | Comparison between analytical and numerical results of concentration along the x axis at α-level cut $= 0$ for $t = 5{,}000$ s for a Gaussian-hill input load. The solution is for a conservative substance.

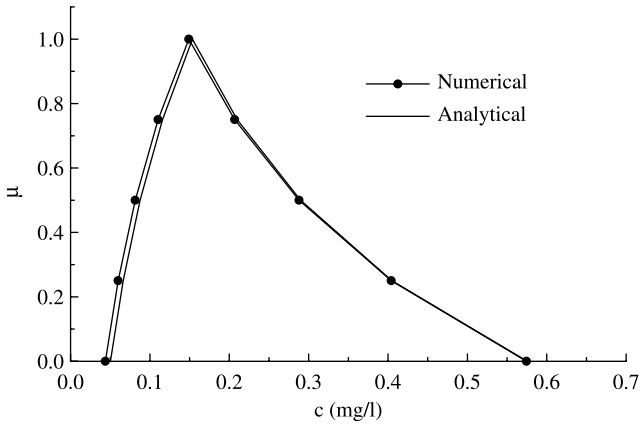

Figure 10 | Fuzzy concentration derived both analytically and numerically for a Gaussian-hill input load at $x = 6{,}500$ m for $t = 5{,}000$ s. The solution corresponds to a non-conservative substance.

(a) The first was a point source, c_o, at the origin of coordinates. The initial and boundary conditions were

$$\begin{aligned} &t = 0 : c = 0 \text{ for } x \geq 0 \\ &t > 0 : c = c_o \text{ for } x = 0 \end{aligned} \tag{4}$$

(b) The second was a pollution load in the form of a Gaussian hill, expressed as

$$c_o(x) = \exp\left[\frac{-(x - x_o)^2}{2s_o^2}\right] \tag{5}$$

where x_o is the centre of mass and s_o is the standard deviation of the initial concentration field. The initial

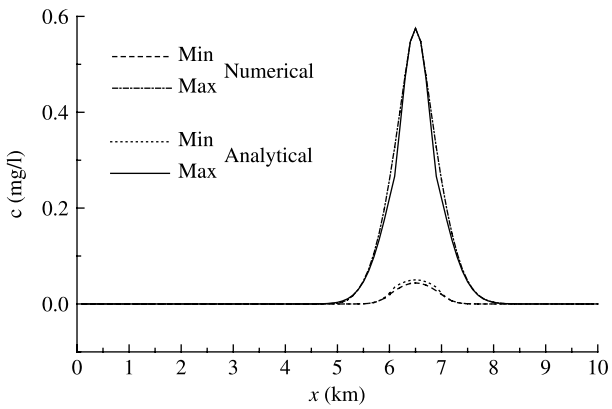

Figure 11 | Concentration derived both analytically and numerically along the x axis at α-level cut $= 0$ for $t = 5{,}000$ s for a Gaussian-hill input load. The solution is for a non-conservative substance.

Table 1 | Definition of source terms of Equation (9)

$\mu = \mu_{20,\max} 1.027^{(T-20)} FF[\mathrm{e}/(\lambda d)][\mathrm{e} - (\mathrm{e}^{-(\lambda d \mathrm{I})/\mathrm{I_L}}) - e^{\mathrm{I}/\mathrm{I_L}}][(c_{\mathrm{NH}} + c_{\mathrm{NAN}})/(c_{\mathrm{NH}} + c_{\mathrm{NAN}} + F_{\mathrm{N}})][c_{\mathrm{IP}}/(c_{\mathrm{IP}} + F_{\mathrm{IP}})]$
$S_{\mathrm{CL}} = S_1 = \mu c_{\mathrm{CL}} - C_{\mathrm{KDA},20} 1.047^{(T-20)} c_{\mathrm{CL}} - \rho c_{\mathrm{CL}} - [(\sigma_{1,20} 1.024^{(T-20)})/d] c_{\mathrm{CL}}$
$S_{\mathrm{ON}} = S_2 = \alpha_1 \rho c_{\mathrm{CL}} - \beta_{3,20} 1.047^{(T-20)} c_{\mathrm{ON}} - \sigma_{3,20} 1.024^{(T-20)} c_{\mathrm{ON}}$
$S_{\mathrm{NH}} = S_3 = \beta_{3,20} 1.047^{(T-20)} c_{\mathrm{ON}} + \sigma_{2,20} 1.074^{(T-20)}/d - \alpha_1 \mu c_{\mathrm{CL}} [c_{\mathrm{NH}}/(c_{\mathrm{NH}} + c_{\mathrm{NAN}})] - \beta_{1,20} 1.047^{(T-20)} c_{\mathrm{NH}}$
$S_{\mathrm{NIN}} = \mathrm{S}_4 = \beta_{1,20} 1.047^{(T-20)} c_{\mathrm{NH}} - \beta_{2,20} 1.047^{(T-20)} c_{\mathrm{NIN}}$
$S_{\mathrm{NAN}} = S_5 = \beta_{2,20} 1.047^{(T-20)} c_{\mathrm{NIN}} - \alpha_1 \mu c_{\mathrm{CL}} [c_{\mathrm{NAN}}/(c_{\mathrm{NH}} + c_{\mathrm{NAN}})]$
$S_{\mathrm{OP}} = S_6 = -\beta_{4,20} 1.047^{(T-20)} c_{\mathrm{OP}} + \alpha_2 \rho c_{\mathrm{CL}} + C_{\mathrm{KDA},20} 1.047^{(T-20)} \alpha_2 c_{\mathrm{CL}} - \sigma_{4,20} 1.027^{(T-20)} c_{\mathrm{OP}}$
$S_{\mathrm{IP}} = S_7 = \beta_{4,20} 1.047^{(T-20)} c_{\mathrm{OP}} - \alpha_2 \mu c_{\mathrm{CL}}$
$S_{\mathrm{BOD}} = S_8 = Y_3 C_{\mathrm{KDA},20} c_{\mathrm{CL}} - k_{1,20} 1.047^{(T-20)} c_{\mathrm{BOD}} - k_{3,20} 1.027^{(T-20)} c_{\mathrm{BOD}}$
$S_{\mathrm{DO}} = S_9 = (k_{2,20}/d) 1.024^{(T-20)} c_{\mathrm{DOD}} + (\alpha_3 \mu - \alpha_4 \rho) c_{\mathrm{CL}} - k_{1,20} 1.047^{(T-20)} c_{\mathrm{BOD}} - \alpha_5 \beta_{1,20} 1.047^{(T-20)} c_{\mathrm{NH}} - \alpha_6 \beta_{2,20} 1.047^{(T-20)} c_{\mathrm{NIN}} - (S_{\mathrm{VOD}})/d$
$S_{\mathrm{COL}} = S_{10} = -k_{4,20} 1.047^{(T-20)} c_{\mathrm{COL}} - 1.17 \times 10^{-5} (FF) I \mathrm{e}^{-\lambda d}$

condition ($t = 0$) is

$$c(x, 0) = c_{\mathrm{o}}(x). \tag{6}$$

The analytical solution of the advection–dispersion equation for a non-conservative point source of pollution as given by the boundary and initial conditions (4) is

$$\tilde{c} = \frac{c_{\mathrm{o}}}{2}\left[\mathrm{e}^{\tilde{T}_1}[1 - \mathrm{erf}(\tilde{S}_1)] + e^{\tilde{T}_2}[1 - \mathrm{erf}(\tilde{S}_2)]\right] \tag{7}$$

where

$$\tilde{S}_1 = \frac{x - \tilde{F}t}{\sqrt{4\tilde{D}_x t}}, \quad \tilde{S}_2 = \frac{x + \tilde{F}t}{\sqrt{4\tilde{D}_x t}}, \quad \tilde{T}_1 = \frac{u - \tilde{F}}{2\tilde{D}_x} x, \quad \tilde{T}_2 = \frac{u + \tilde{F}}{2\tilde{D}_x} x,$$

$$\tilde{F} = \sqrt{u^2 + 4\tilde{D}_x \tilde{k}}.$$

For a conservative substance $\tilde{k} = 0$, thus $\tilde{F} = u$.

The analytical solution of the advection–dispersion equation for the Gaussian-hill pollution source as given by the initial condition (6) is

$$\tilde{c} = \frac{s_{\mathrm{o}}}{\sqrt{2\tilde{D}_x t + s_{\mathrm{o}}^2}} \exp\left[\frac{-(x - x_{\mathrm{o}} - ut)^2}{4\tilde{D}_x t + 2s_{\mathrm{o}}^2 - \tilde{k}t}\right]. \tag{8}$$

For $\tilde{k} = 0$, the solution for a conservative substance is obtained.

For the two input loads of pollution quoted previously the solution of Equation (1) for a conservative substance was conducted first. Then the procedure was repeated for a non-conservative substance. The solution was performed at five different α-level cuts (0, 0.25, 0.5, 0.75, 1) of the fuzzy input parameters, namely $\tilde{D}_x$ for the conservative, and $\tilde{D}_x$ and $\tilde{k}$ for the non-conservative substance.

(a) For the point source at the origin of coordinates, the number of nodes was 200, the spacing between adjacent nodes was 50 m, the time step was 20 s and the number of iterations was 500. The velocity u at all nodes was considered constant equal to 0.5 m/s. The fuzzy dispersion coefficient, $\tilde{D}_x = (1, 10, 15)$, is depicted in Figure 2. For the conservative substance, the fuzzy concentrations derived numerically for $x = 4{,}750$ m and $x = 5{,}250$ m at $t = 10{,}000$ s, are superimposed on the analytical values in Figures 3(a, b). It is interesting to note that, while the input parameters are triangular fuzzy numbers, the output concentrations are fuzzy numbers but not triangular, owing to the multiplication and division operations between fuzzy numbers. Numerical and analytical results at α-level cut = 0 along the x axis for $t = 10{,}000$ s are shown in Figure 4. The left bound represents the minimum possible value, whereas the right bound represents the maximum possible value of concentration. It can be clearly seen that the numerical results agree closely with the analytical. It should be stressed that $\tilde{D}_x$ is the one and only fuzzy input parameter, thus the minimum and maximum value of concentration at each α-level cut can be obtained without the use of fuzzy analysis. Two different solutions, one with the minimum and the other with the maximum D_x at the corresponding

Table 2 | Description of the parameters used in the model

Symbol	Unit	Description
μ	$(d)^{-1}$	Growth rate of chlorophyll-a
$\mu_{20,max}$	$(d)^{-1}$	Maximum growth rate of chlorophyll-a (20°C)
T	°C	Temperature
FF	–	Factor for solar radiation
λ	m^{-1}	Light extinction coefficient (20°C)
d	m	Water depth
I	W/m^2	Surface solar light intensity
I_L	W/m^2	Solar light intensity for the maximum growth speed
F_N	mg/l	Monod constant for nitrogen (ammonia plus nitrate)
F_{IP}	mg/l	Monod constant for inorganic phosphorus
$C_{KDA,20}$	$(d)^{-1}$	Decay rate coefficient for chlorophyll-a (20°C)
ρ	$(d)^{-1}$	Respiration rate for chlorophyll-a
α_1	mg ON/mg CL	Fraction of algae biomass that is organic nitrogen
α_2	mg P/mg CL	Phosphorous content of chlorophyll-a
α_3	mg O/mg CL	Rate of oxygen production per unit of algae photosynthesis
α_4	mg O/mg CL	Rate of oxygen uptake per unit of algae respired
α_5	mg O/mg NH_3	Rate of oxygen uptake per unit of ammonia nitrogen oxidation
α_6	mg O/mg NO_2	Rate of oxygen uptake per unit of nitrite nitrogen oxidation
$\beta_{1,20}$	$(d)^{-1}$	Constant for the biological oxidation of ammonia nitrogen (20°C)
$\beta_{2,20}$	$(d)^{-1}$	Constant for the biological oxidation of nitrite nitrogen (20°C)
$\beta_{3,20}$	$(d)^{-1}$	Organic nitrogen hydrolysis rate (20°C)
$\beta_{4,20}$	$(d)^{-1}$	Rate constant for the decay of organic phosphorous to inorganic phosphorous (20°C)
$\sigma_{1,20}$	(m/d)	Local settling rate for algae (20°C)
$\sigma_{2,20}$	mg $NH_3/(m^2\ d)$	Benthos source rate for ammonia nitrogen (20°C)
$\sigma_{3,20}$	$(d)^{-1}$	Organic nitrogen settling rate (20°C)
$\sigma_{4,20}$	$(d)^{-1}$	Organic phosphorus settling rate (20°C)
$k_{1,20}$	$(d)^{-1}$	Deoxygenation rate coefficient (20°C)
$k_{2,20}$	(m/d)	Raeration rate coefficient (20°C)
$k_{3,20}$	$(d)^{-1}$	Rate of loss of BOD due to settling (20°C)
$k_{4,20}$	$(d)^{-1}$	Decay rate coefficient for coliforms(20°C)
Y_3	–	Stoichiometric constant
S_{VOD}	mg $DO/(m^2\ d)$	Benthic oxygen demand rate (20°C)

α-level cut, would be sufficient. This technique, although computationally more time-consuming, was also applied trivially, in order to verify the results of the fuzzy analysis.

The solution of the advection–dispersion equation for a non-conservative substance was conducted with $\tilde{D}_x$ depicted in Figure 2 and fuzzy decay coefficient $\tilde{k}$ depicted in Figure 5. The fuzzy concentration distributions obtained numerically for $x = 4{,}000$ m and $x = 5{,}250$ m at $t = 10{,}000$ s are superimposed on the analytical values in Figures 6(a, b). Numerical and analytical results at α-level cut = 0 along the x axis for $t = 10{,}000$ s are depicted in Figure 7. Again, the numerical results agree closely with the analytical.

Table 3 | Triplets of the fuzzy parameters used in the model

Symbol	x_1	x_2	x_3
D_x	1	3	5
$\mu_{20,\max}$	0.432	0.864	5.184
T	23	25	27
λ	1.2	1.5	2.5
I	70	180	240
I_L	200	270	300
F_N	1×10^{-3}	5×10^{-3}	1.4×10^{-2}
F_{IP}	1×10^{-3}	7×10^{-3}	1×10^{-2}
$C_{KDA,20}$	0.302	0.475	0.864
ρ	0.086	0.259	0.605
α_1	0.1	0.12	0.15
α_2	6×10^{-3}	1.5×10^{-2}	2×10^{-2}
α_3	2	2.5	2.8
α_4	1	1.5	1.8
α_5	1	1.5	2.5
α_6	0.5	0.7	1
$\beta_{1,20}$	0.086	0.259	0.432
$\beta_{2,20}$	0.086	0.302	0.432
$\beta_{3,20}$	1.728×10^{-2}	0.302	0.993
$\beta_{4,20}$	8.64×10^{-3}	6.48×10^{-2}	8.64×10^{-2}
$\sigma_{1,20}$	0.518	0.864	1.728
$\sigma_{3,20}$	8.64×10^{-3}	3.456×10^{-2}	8.64×10^{-2}
$\sigma_{4,20}$	8.64×10^{-3}	6.48×10^{-2}	8.64×10^{-2}
$\sigma_{5,20}$	0.086	1.728	6.048
$k_{1,20}$	3.456×10^{-2}	5.18×10^{-2}	8.64×10^{-2}
$k_{2,20}$	8.64	17.28	43.2
$k_{3,20}$	0.864	2.592	5.184
$k_{4,20}$	0.259	0.864	5.184
Y_3	1	4	10
S_{VOD}	0.086	0.562	0.864

(b) For the Gaussian-hill initial load, the center of mass of the initial concentration, x_o, was 1,500 m and the standard deviation, s_o, was 300. The number of nodes was 100, the spacing between adjacent nodes was 100 m, the time step was 20 s and the number of iterations was 250. The velocity u at all nodes was considered constant equal to 1 m/s. The solution was conducted for a fuzzy dispersion coefficient $\tilde{D}_x = (2.5, 15, 25)$. The fuzzy concentration distribution derived numerically for $x = 6{,}500$ m at $t = 5{,}000$ s is superimposed on the analytical in Figure 8. The numerical and analytical distributions at α-level cut $= 0$ along the x axis for $t = 5{,}000$ s are depicted in Figure 9. Again the numerical and the analytical results agree closely. Computation was also conducted for a non-conservative substance, for the same initial conditions and for a fuzzy decay coefficient $\tilde{k} = (1 \times 10^{-4},\ 3 \times 10^{-4},\ 5 \times 10^{-4})$. The fuzzy concentrations obtained numerically and analytically at $x = 6{,}500$ m for $t = 5{,}000$ s are shown in Figure 10, while the numerical and analytical results along the x axis at α-level cut $= 0$ for $t = 5{,}000$ s are superimposed in Figure 11. It can be clearly seen that the results obtained using the Lagrangian–Eulerian method display a remarkable similarity with those of the analytical fuzzy solution.

APPLICATION IN WATER-STREAM POLLUTION

Rivers are often receptors of wastewater from various sources, such as urban sewerage networks, industrial wastewater, and agricultural wastewater containing fertilisers and pesticides. Pollution of river waters is associated with:

(a) oxygen depletion, caused by loads of organic wastes,
(b) eutrophication, caused by oversupply of nutrients (N, P),
(c) contamination, caused by pathogenic organisms like coliforms.

A river quality model should take into account several interactions between the main quality parameters, which are: (1) chlorophyll-a (CL), (2) organic nitrogen (ON),

Table 4 | Triplets of fuzzy input loads of the pollutants used in the model

Pollutant	x_1	x_2	x_3
CL	1.0	1.5	2.0
ON	1.0	2.0	4.0
NH	0.6	0.8	1.0
NIN	1.0	1.5	2.0
NAN	0.5	0.7	1.0
OP	1.0	1.2	1.7
IP	0.5	1.0	1.3
BOD	5.0	7.0	10.0
DOD	4.0	4.5	5.0
COL	1,000	5,000	10,000

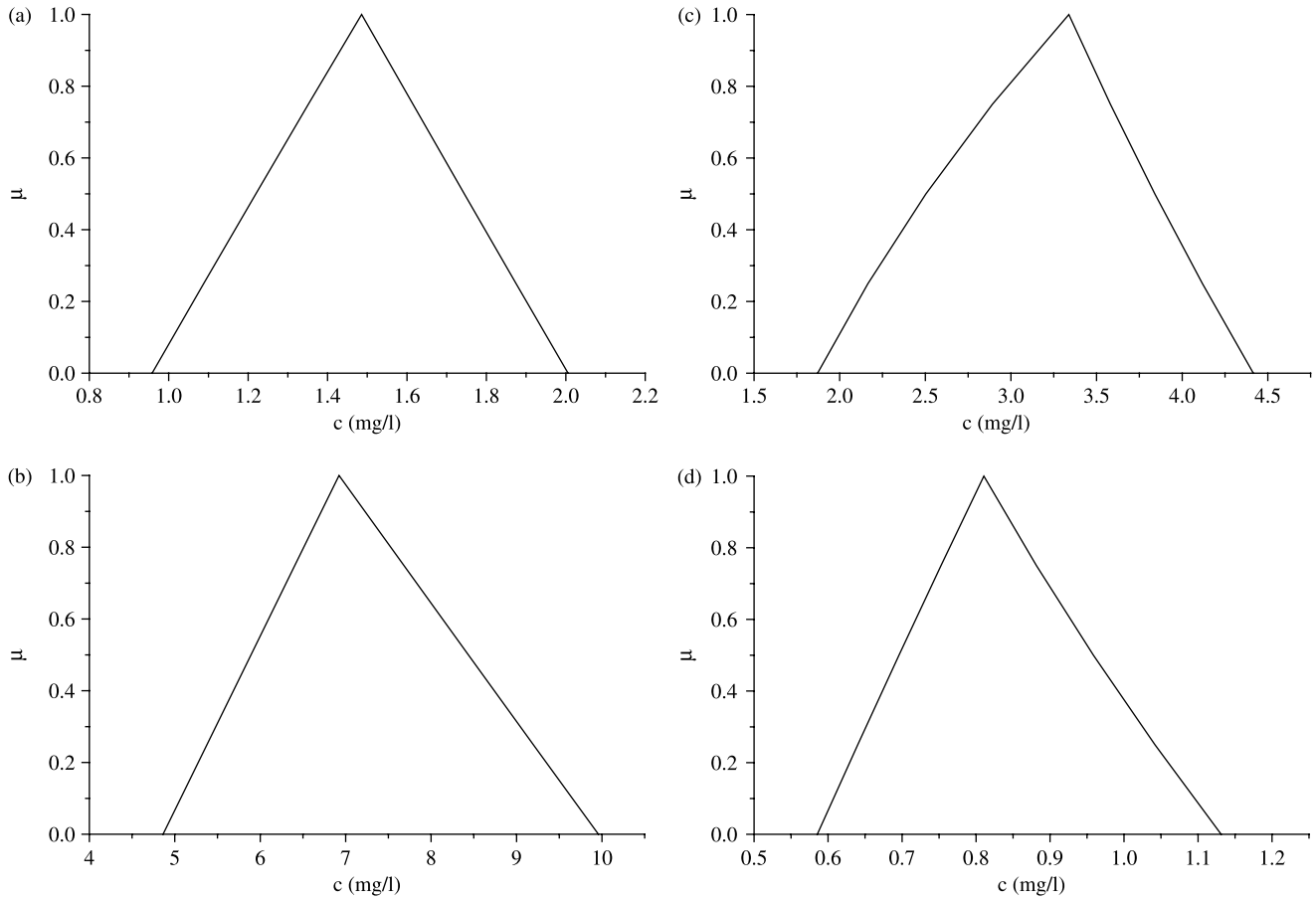

Figure 12 | Membership functions of concentrations of various pollutants at $x = 1{,}500$ m for $t = 5{,}000$ s. (a) Chlorophyll-a (CL); (b) biochemical oxygen demand (BOD); (c) dissolved oxygen deficit (DOD); (d) ammonia nitrogen (NH).

(3) ammonia nitrogen (NH), (4) nitrite nitrogen (NIN), (5) nitrate nitrogen (NAN), (6) organic phosphorus (OP), (7) inorganic phosphorus (IP), (8) biochemical oxygen demand (BOD), (9) dissolved oxygen deficit (DOD) and (10) coliforms (COL). The equations of the model presented herein describe the variation of the fuzzy concentrations $\tilde{c}_i$ of all these parameters with respect to time, under biochemical interactions. The full set of equations is composed of the well-known advection–dispersion equations for a non-conservative substance:

$$\frac{\partial \tilde{c}_i}{\partial t} + u\frac{\partial \tilde{c}_i}{\partial x} = \tilde{D}_x \frac{\partial^2 \tilde{c}_i}{\partial x^2} + \tilde{S}_i \tag{9}$$

where $\tilde{c}_i$ is the concentration for the ith parameter and $\tilde{S}_i$ are the source and sink fuzzy terms for each equation describing the biochemical reactions. The mathematical form of the source terms is given in Table 1, whereas the various coefficients used in the model are explained in Table 2. All parameters of Table 2 are fuzzy coefficients, except for the water depth d and the factor for solar radiation, FF. It should be noted that the model contains dissolved oxygen (DO), but the dissolved oxygen deficit (DOD) is used for the description of pollution. The DOD value was calculated by subtracting the computed DO value from the saturation value of DO. The exact saturation DO value depends on water temperature, and is approximately 10 mg/l.

The limits of the various coefficients contained in the model were obtained from various sources, such as that by Brown & Barnwell (1987), and were compiled by Mpimpas (1998). The triplets of the fuzzy coefficients used in the

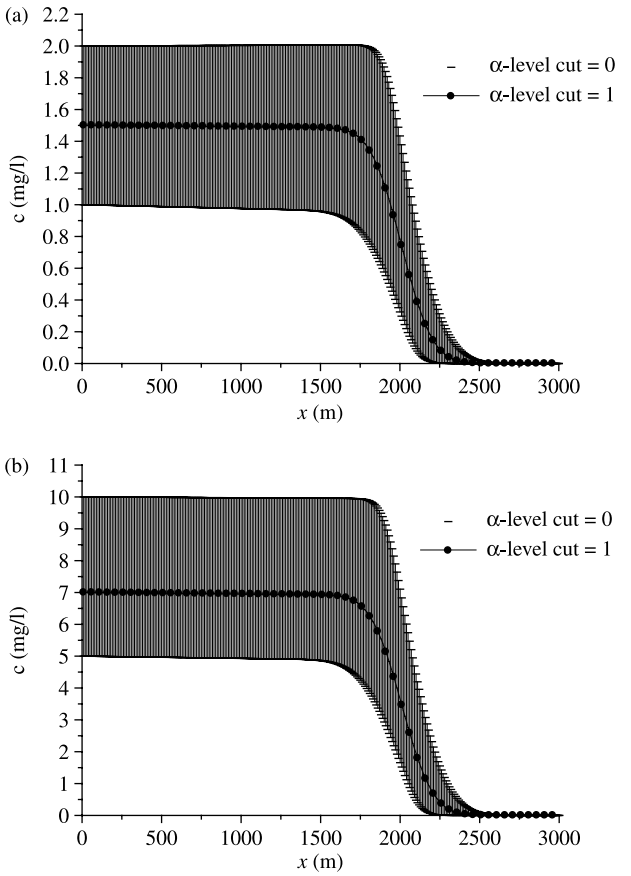

Figure 13 | Minimum and maximum bounds of concentration (α-level cut = 0) and the most confident value (α-level cut = 1) along the stream. (a) Chlorophyll-a (CL); (b) biochemical oxygen demand (BOD).

model are quoted in Table 3. Most of these values are similar to those used by Mpimpas *et al.* (2001) for the study of water pollution in a coastal area. It should be pointed out that, in the one-dimensional advection–dispersion equation, the dispersion coefficient incorporates effects of cross-sectional averaging of a non-uniform flow. This often leads to values for the dispersion coefficient in the 1-D equation that are not necessarily physically realistic compared to turbulent diffusion values. Turbulent diffusion in flowing water is generally of the order of $1\,m^2/s$ and the remainder is additional dispersion by shear because of the non-uniform velocity distribution in a cross-sectional area. For example, although the fuzzy dispersion coefficient used herein is defined by the triplet (1, 3, 5), much higher values may result if a stream of substantially larger width is considered.

The concentrations of the input loads of pollutants are expressed in mg/l, except for the coliforms, which are expressed as numbers per 100 ml, while the pollution sources were placed at the upstream end of the river. For each parameter the input loads are expressed as fuzzy numbers. The number of nodes was 1,000, the spacing between adjacent nodes was 10 m and the number of particles introduced initially was 1,000. The time step was 2 s, the flow velocity was considered constant at all nodes equal to 0.50 m/s and the computation was continued until a time equal to 5,000 s was reached. The triplets of the fuzzy input loads of the pollutants used in the model are quoted in Table 4. The concentrations of some of the quality parameters (CL, BOD, DOD and NH) at the 150th node of the domain ($x = 1{,}500$ m) expressed in fuzzy form are presented in Figure 12. The results correspond to a time equal to 5,000 s from the inception of the computation. The resulting fuzzy numbers are not triangular, although the deviation from triangular, with the exception of Figure 12(c), may be considered as small. Figure 13 depicts the minimum and maximum bounds of concentrations (α-level cut = 0) of CL and BOD and the most confident values (α-level cut = 1) along the river, at $t = 5{,}000$ s. The results show how the concentration imprecision is very sensitive to the uncertainty width (degree of fuzziness) of the various coefficients.

As stated previously the lower (minimum) and upper (maximum) bounds of the values of concentrations cannot be obtained by using simply the extreme values of the various coefficients without the use of interval operations. For the model used in this study, which contains the 30 fuzzy parameters listed in Table 3, the enormous number of 2^{30} computations is required in order to obtain all possible values of concentration for each quality parameter, ignoring the fuzziness of input loads. If the fuzziness of input loads is taken into account, the number of computations increases to 2^{40}. From the whole range of possible values, the minimum and maximum values of concentrations of the pollutants at each node can be determined. On the other hand, with the use of the present fuzzy model, just one computation is sufficient to determine the extreme values of concentration at each node (α-level cut = 0), and five computations for the determination of the extreme concentration values at all five α-level cuts considered in the present study.

CONCLUSIONS

One of the problems in water pollution modelling is the vagueness in the values of pollutant sources and biochemical coefficients, arising either from natural randomness in time and space or from indirect measurements and a limited number of samples. With fuzzy modelling we can represent imprecise data and produce imprecise output in the form of fuzzy numbers, with minimal input data requirements and without the need of a large number of computations.

The fuzzy approach combined with the finite-difference method was employed for the solution of the one-dimensional advection–dispersion equation of both a conservative and a non-conservative substance for two different types of input pollutant loads: a point source at the origin of coordinates and a Gaussian-hill initial load. The results derived from the numerical computation expressing the dispersion and decay coefficient as fuzzy parameters agree closely with those of the analytical solution, confirming the accuracy of the numerical technique.

The finite-difference method combined with fuzzy analysis was also used for the study of pollution in a stream, considering ten water quality parameters. The solution yields the most confident value of each quality variable, and is very efficient for the determination of the extreme values of these quantities, a task for which a vast number of computations would be required.

REFERENCES

Alefeld, G. 1983 *Introduction to Interval Computations*. Academic Press, New York.

Ang, A. H. & Tang, W. H. 1984 *Probability Concepts in Engineering Planning and Design*, vol. 2. Wiley, New York.

Brown, L. C. & Barnwell, T. 1987 *The Enhanced Stream and Water Quality Models QUAL2 and QUAL2-UNCAS. Documentation and User Model.* US Environmental Protection Agency, Environmental Research Laboratory, Athens, GA.

Dou, C., Woldt, W., Bogardi, I. & Dahab, M. 1995 Steady state groundwater flow simulation with imprecise parameters. *Water Resour. Res.* **31**, 2709–2719.

Ducstein L. & Plate E. (eds) (1987). *Engineering Reliability and Risk in Water Resources* Nijhoff, Dordrecht.

Ganoulis, J. (ed.) 1991 *Water Resources Engineering Risk Assessment, NATO ASI Series*, vol. G29. Springer Verlag, Berlin.

Ganoulis, J. 1994 *Engineering Risk Analysis of Water Pollution Probabilities and Fuzzy Sets*. VCH, Weinheim.

Ganoulis, J., Anagnostopoulos, P. & Mpimpas, H. 2003 Fuzzy numerical simulation of water quality. In: *Proc. 30th IAHR Congress, Theme B, August 2003, Thessaloniki, Greece*. IAHR, Spain. pp. 165–174.

Ganoulis, J., Mpimpas, H. & Anagnostopoulos, P. 1995 Coastal water assessment using fuzzy arithmetic. In: *Proc. IEEE Workshop on Nonlinear Signal Image Processing, 20–22 June, Porto Carras, Greece,* vol. II. pp. 1015–1018.

Guymon, G. L. & Yen, C. C. 1990 An efficient deterministic-probabilistic approach to modelling regional groundwater flow–2. Application to Owens Valley, California. *Water Resour. Res.* **26**, 1569–1581.

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

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

Klir, G. & Folger, T. 1988 *Fuzzy Sets, Uncertainty and Information*. Prentice-Hall, Englewood Cliffs, NJ.

Kruse, R., Gebhart, J. & Klawonn, F. 1994 *Foundations of Fuzzy Systems*. Wiley, New York.

Moore, R. E. 1979 *Methods and Applications of Interval Analysis.* SIAM Studies in Applied Mathematics, Philadelphia.

Mpimpas, H. 1998 *Numerical Investigation of the Pollutant Dispersion with Fuzzy Logic* (in Greek). PhD thesis, Department of Civil Engineering, University of Thessaloniki, Thessaloniki, Greece.

Mpimpas, H., Anagnostopoulos, P. & Ganoulis, J. 2001 Modelling of water pollution in the Thermaikos Gulf with fuzzy parameters. *Ecol. Modell.* **142**, 91–104.

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

Nickel, K. (ed.) 1986 *Interval Mathematics 1985* Springer Verlag, Berlin.

Perk, V. 1996 *Uncertainty Issues in Modelling the Influence of Bed Sediments on Water Composition.* Netherlands Geographical Studies No. 200, University of Utrecht, Amsterdam.

Plate, E. 1991 Probabilistic modelling of water quality in rivers. In: *Proc. Water Resources Engineering Risk Assessment*, NATO ASI Series. vol. G29 (ed. J. Ganoulis), pp. 137–166. Springer Verlag, Berlin.

Rall, L. B. 1986 Improved interval bounds for range of functions. In: *Interval Mathematics 1985* (ed. K. Nickel), pp. 143–155. Springer Verlag, Berlin.

Shafike, N. G. 1994 *Groundwater Flow Simulations and Management under Imprecise Parameters.* PhD thesis, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ.

Silvert, W. 1997 Ecological impact classification with fuzzy sets. *Ecol. Modell.* **96**, 1–10.

Zadeh, L. A. 1965 Fuzzy sets. *Inf. Control* **8**, 338–353.

Zimmermann, H. J. 1991 *Fuzzy Set Theory and Its Applications*, 2nd edition. Kluwer, Boston.

First received 23 August 2007; accepted in revised form 17 February 2008

APPENDIX A. FUZZY NUMBERS AND FUZZY OPERATIONS

Fuzzy sets are those in which membership cannot be defined on a yes/no basis, because their boundaries are vague. The central concept of the fuzzy set theory is the membership function, μ, which represents the degree to which an element belongs to a set, and is expressed by a number between 0 and 1 (0 represents the minimum and 1 the maximum confidence for an element to belong to the set). A fuzzy number is a fuzzy set of real numbers whose membership function reaches unity and is convex (Zimmerman 1991; Kruse *et al.* 1994). A fuzzy number $\tilde{A}$, (where the overscore " $\sim$ " means fuzzy), is defined in mathematical terms as: $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) : x \in R, \mu_{\tilde{A}}(x) \in [0,1]\}$, where $\mu_{\tilde{A}}(x)$ is the membership function of x in $\tilde{A}$. The simplest type of a fuzzy number is the triangular. A triangular fuzzy number (TFN) can be defined by a triplet (x_1, x_2, x_3), as shown in Figure 1. The membership function for a TFN is defined as

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & x < x_1 \\ \frac{x - x_1}{x_2 - x_1} & x_1 \le x \le x_2 \\ \frac{x_3 - x}{x_3 - x_2} & x_2 \le x \le x_3 \\ 0 & x > x_3 \end{cases}$$

Let us consider two triangular fuzzy numbers $\tilde{A}$ and $\tilde{B}$ defined by the triplets $\tilde{A} = (x_1, x_2, x_3)$ and $\tilde{B} = (y_1, y_2, y_3)$. According to Kaufmann & Gupta (1985), their addition and subtraction are defined as

$$\text{Addition} : \tilde{A} + \tilde{B} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

$$\text{Subtraction} : \tilde{A} - \tilde{B} = (x_1 - y_3, x_2 - y_2, x_3 - y_1)$$

For multiplication and division operations, triplets cannot be used. However, the computation can be performed using the confidence interval $\overline{A_\alpha}$ at each α-level (Kaufmann & Gupta 1985): $\forall \alpha \in [0, 1]$: $\overline{A_\alpha} = [\alpha_1^{(\alpha)}, \alpha_2^{(\alpha)}] = [(x_2 - x_1)\ \alpha + x_1, -(x_3 - x_2)\ \alpha + x_3]$ (where the overscore "$^-$" means interval).

APPENDIX B. ARITHMETIC OPERATIONS IN INTERVAL NUMBERS

An interval number $\overline{A}$ is defined as (Alefeld 1983; Nickel 1986)

$$\overline{A} = [\alpha_1, \alpha_2] = \{t | \alpha_1 \le t \le \alpha_2\}$$

where α_1 is the left bound and α_2 is the right bound of the interval $\overline{A}$. The operations on the intervals $\overline{A} = [\alpha_1, \alpha_2]$ and $\overline{B} = [\beta_1, \beta_2]$ are defined as

$$\overline{A} + \overline{B} = [\alpha_1 + \beta_1, \alpha_2 + \beta_2]$$

$$\overline{A} - \overline{B} = [\alpha_1 - \beta_2, \alpha_2 - \beta_1]$$

$$\overline{A} * \overline{B} = [\min(\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2\beta_2), \max(\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2\beta_1, \alpha_2\beta_2)]$$

$$\overline{A} / \overline{B} = [\min(\alpha_1/\beta_1, \alpha_1/\beta_2, \alpha_2/\beta_1, \alpha_2/\beta_2), \max(\alpha_1/\beta_1, \alpha_1/\beta_2, \alpha_2/\beta_1, \alpha_2/\beta_2)]$$

$$-\overline{A} = -[\alpha_1, \alpha_2] = [-\alpha_2, -\alpha_1].$$

Considering the three intervals $\overline{A}$, $\overline{B}$ and $\overline{C}$, the following properties can be defined:

$$\text{Associativity} : (\overline{A} + \overline{B}) + \overline{C} = \overline{A} + (\overline{B} + \overline{C}),$$

$$(\overline{A}\ \ \overline{B})\overline{C} = \overline{A}(\overline{B}\ \ \overline{C})$$

$$\text{Commutativity} : \overline{A} + \overline{B} = \overline{B} + \overline{A}, \quad \overline{A}\ \ \overline{B} = \overline{B}\ \ \overline{A}$$

$$\text{Subdistributivity} : \overline{A}(\overline{B} + \overline{C}) \subseteq \overline{A}\ \ \overline{B} + \overline{A}\ \ \overline{C}.$$