Consideration of parameters and boundary conditions uncertainties in water balance and solute transport simulation

AMIR HASSAN ZADEH, RENE BLANKENBURG & PETER-WOLFGANG GRAEBER

Technische Universitaet Dresden, Institute of Waste Management and Contaminated Site Treatment, Pratzschwitzer Str. 15, 01796 Pirna, Germany

rene.blankenburg@mailbox.tu-dresden.de

Abstract In this paper, a fuzzy-set theory-based approach is presented to incorporate the verbal expert knowledge with the uncertainties of the input data in a numerical model. It allows for the consideration of the fluctuation range of measurements, input parameters (e.g. soil and solute transport parameters) and boundary conditions. To extend numerical models to operate with fuzzy input parameters, interval arithmetic is used. Due to the characteristics of unsaturated zone processes, a nonlinear optimization problem occurs, where with given ranges of parameter sets min/max values must be found by means of gradient and objective function calculations. Furthermore, a second method, called fuzzy analysis library, is tested. This method provides a statistical approach based on an evolutionary algorithm so that no gradient calculations are performed. Both methods are applied to water flow and solute transport processes in the soil zone. The results of both methods are compared to each other.

Key words water balance; solute transport; fuzzy set theory; uncertainty analysis; unsaturated zone; fuzzy analysis library

INTRODUCTION

The basis of typical environmental information systems is usually given by means of numerical models. Their application requires a high amount of expert knowledge which is often verbal information. Therefore, this expertise cannot be combined directly with numerical models.

The input data of such models usually comes with uncertainties which result from imprecise measurements or heterogeneities of the quantities to be measured. To combine the verbal expert knowledge with the uncertainties of the input data in a numerical model, a methodology based on fuzzy set theory is described in this paper. It allows for the consideration of the fluctuation range of (laboratory) measurements, input parameters (e.g. soil and solute transport parameters) and boundary conditions. In particular, it is important if the required parameters can not be sufficiently estimated or the density measurement is inadequate. Similarly, if model predictions are calculated, both the impact and temporal change of boundary conditions, can be taken into account.

Fuzzy logic (Kruse *et al.*, 1993) is an analysis methodology developed to incorporate uncertainty into decision making and control models. Fuzzy logic allows inclusion of imperfect information, whatever the cause. There are key benefits to applying fuzzy tools. Fuzzy tools provide a simplified platform where the development and analysis of models require reduced development time compared with other approaches. As a result, fuzzy tools are easy to implement and modify. Nevertheless, despite their "user-friendly" outlet, fuzzy tools perform as well as or better than other soft approaches to decision making under uncertainties.

During the past two decades fuzzy logic has been applied to a multitude of environmental engineering challenges. Bardossy *et al.* (1988, 1995) were among the first to incorporate imperfect information from expert judgement and subjective information in groundwater flow simulation. They carried out fuzzy rule-based modelling for water movement in the unsaturated soil matrix. Other attempts at applying fuzzy logic in groundwater simulation were developed by Schulz & Huwe (1997). They presented the incorporation of fuzzy soil hydraulic properties and boundary condition into closed-form solutions of the Darcy-Buckingham equation to analyse the impact on 1D steady-state water flow and on the estimation of maximum evapotranspiration rates. Also, an extension of the fuzzy flow problem with the incorporation of a finite difference solution into a nonlinear optimization procedure was investigated (Schulz *et al.*, 1999). Verma *et al.* (2009)

presented a method based on fuzzy set theory to express imprecision of input data, in terms of fuzzy number, to quantify the uncertainty in prediction. Huang *et al.* (2010) developed a fuzzy-based simulation method (FBSM) through coupling distributed hydrological model with fuzzy vertex analysis for water resources management.

This paper addresses fuzzy-based simulation for 2D movement of water flow and solutes in variably-saturated porous media with the incorporation of a finite element method into two different methods (i) nonlinear optimization tool “NLPQLP” (Schittkowski, 2006) and (ii) fuzzy analysis library “FALIB” (Möller *et al.*, 2007). The proposed fuzzy methods are implemented by modifying the computer code CHAIN_2D (Simunek & van Genuchten, 1994). Both approaches are applied to steady-state and transient water flow and solute transport processes in the soil zone.

An application of the described approaches considers the simulation of solute transport using the classic advection-dispersion equation. It can be used for the numerical implementation of leachate forecasts according to the German Soil Protection Law (BBodSchG). The point of interest is here the assessment of the amount (concentration and yield) of contaminants infiltrated into the groundwater at the transfer zone between the unsaturated and saturated zone. As a result of the fuzzy input parameters, the breakthrough curves for concentration and yield are not just represented by a single graph. They are now described as envelopes. The results of the described methods were comparable, despite their different approach.

The remaining part of this paper is organized in the following sections: (1) provides an explanation of the fundamental aspects of fuzzy logic; (2) incorporation of imprecision parameters and boundary condition into 2D movement of water flow and multiple solutes in variably-saturated porous media is presented; (3) fuzzy based simulation with the program of NLPQLP and FALIB; (4) performing the first application, i.e. fuzzy transient flow simulation within a dike body; (5) deals with imprecision input data and boundary conditions to leachate forecasts; and (6) conclusions.

Basics of fuzzy logic

Fuzzy set theory was first created by Professor Lotfi Zadeh in the 1960s to mathematically represent uncertainty and vagueness. This theory reflects human reasoning in a non-probabilistic sense to generate decisions. At its most basic form fuzzy logic considers variables to be included in a set based on their degrees of membership, rather than absolute membership.

In the classical set theory the membership of elements in relation to a set is assessed in binary terms according to a tight condition. An element either belongs or does not belong to the set; the boundary of the set is tight. As a further development of classical set theory, fuzzy set theory permits the gradual assessment of the membership of elements in relation to a set; this is described with the aid of membership function. A fuzzy set is defined as follows:

If X represents a fundamental set and x are the elements of this fundamental set, to be assessed according to an (lexical or informal) uncertain proposition and assigned to a subset A of X, the set:

$$\tilde{A} = \left\{ \left(x, \mu_A(x) \right) \middle| x \in X \right\}$$

is referred to as the uncertain set of fuzzy set on X. $\mu_A(x)$ is the membership function of the fuzzy set $\tilde{A}$. If

$$\mu_A(x) \geq 0 \quad \forall x \in X$$

$$\sup_{x \in X} \left(\mu_A(x) \right) = 1$$

holds, the membership function is referred to as normalized. A fuzzy subset A is called a fuzzy number, if A is a normal, convex fuzzy set of the set of real numbers. Calculations with fuzzy numbers allow the incorporation of uncertainty on parameters, properties, geometry, initial conditions, etc.

As an example, Fig. 1 shows purely linguistically the following expert statement “The hydraulic conductivity of this soil is always greater than 70 cm/d and no greater than 200 cm/d. Almost certainly the hydraulic conductivity can be estimated to 130 cm/d.

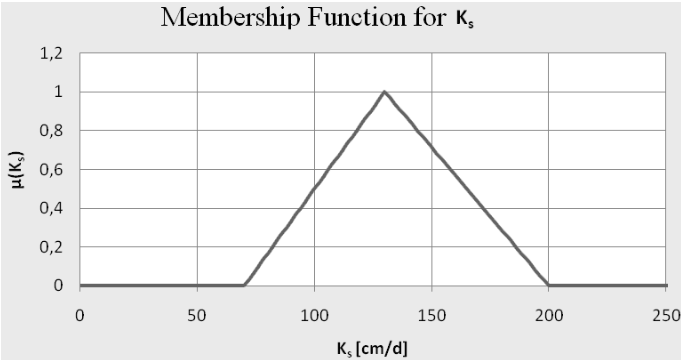

Fig. 1 Triangular membership function for the unsaturated hydraulic conductivity.

The α -level cut (A_α) of a fuzzy subset A is the set of those elements, which have at least a membership value greater than or equal to α:

$$A_\alpha = \{x \in X, \mu_A(x) \geq \alpha\}$$

The supplement of a fuzzy number (supp) is the set of real numbers:

$$\text{supp}(A) = \{x \in X \,|\, \mu_A(x) \geq 0\}$$

This horizontal representation is very close to human thinking where uncertainty and imprecision is usually expressed by associating different "intervals of confidence" with different "levels of presumptions" and accepting a range of parameter values as wide as possible with a decreasing level of presumption (Schulz & Huwe, 1999).

In order to be able to define mathematical operations with imprecision quantities and/or fuzzy numbers, it is necessary to extend the Cartesian product to imprecision quantities.

If $X_1,....,X_n$ and Y are sets, and f is a mapping of the Cartesian product $X_1 \times ... \times X_n$ to Y, then f can be extended to operate on the Cartesian product $A_1 \times ... \times A_n$ of fuzzy subsets of $X_1,....,X_n$ with membership functions $\mu_{A_1}(x),...,\mu_{A_n}(x)$. The image of $A_1 \times ... \times A_n$ in Y is the fuzzy subset B with the membership function:

$$\mu_B(y) = \begin{cases} \sup\{\min\left(\mu_{A_1}(x_1),...,\mu_{A_n}(x_n)\right), y = f(x_1,...,x_n)\} \\ 0 \text{ if no}(x_1,...,x_n) \in X_1 \times ... \times X_n \text{ such that } f(x_1,...,x_n) = y \end{cases}$$

By using α-level cut, operations on fuzzy sets can be reduced to operations within the interval arithmetic (Kruse *et al.*, 1993)

If $A_1,...,A_n$ are fuzzy numbers of A^*, $R^n = R \times ... \times R$ is the Cartesian product and $\varphi : R^n \to R$ is a continuous mapping,

$$[\hat{\varphi}(A_1,...,A_n)]_\alpha = \hat{\varphi}(A_{1_\alpha},...,A_{n_\alpha}) = \{y \,|\, y = \varphi(x_1,...,x_n); x_1 \in A_{1_\alpha} \wedge ... \wedge x_n \in A_{n_\alpha}\}$$

A_{i_α} denotes the α-cut of fuzzy number A_i.

Mathematically it means that operations on imprecision numbers for the individual α-cuts can be attributed to appropriate operations within interval arithmetic. In other words, to create the membership function of imprecision numbers for each α-cut, minimum and maximum possible function values must be found. This requires complex functions or operators with the application of a nonlinear numerical optimization procedure.

PROBLEM DEFINITION AND MODEL DEVELOPMENT

Water balance and solute transport simulation

As a simulation tool, the CHAIN_2D computer program (Simunek *et al.*, 1994), available as a public domain code, was selected. It is a 2D model for the simulation of water flow and contaminant transport in variably saturated porous media (i.e. both, saturated and unsaturated zone). The water flow is represented by the Richards equation for saturated and unsaturated flow. It contains a sink term for water uptake by plant roots. The flow region can be composed of non-uniform soils to incorporate the effects of anisotropy. The contaminant transport is modelled by the advection-dispersion equation which is modified for simulating nonlinear non-equilibrium reactions. The boundary conditions for the contaminant transport could be constant concentration or constant flux. The flow and transport equations are solved numerically using the Galerkin Finite Element method.

The governing equation for 2D flow in variably saturated media can be written as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x_i}\left[k \left(K_{ij}^{A} \frac{\partial h}{\partial x_i} + K_{iz}^{A} \right) \right] - s$$

with θ as the volumetric water content, x_i are the spatial coordinates (x, y, z), h is the capillary pressure head, t is the time, K_{ij}^{A} and K_{iz}^{A} are the components of the dimensionless anisotropy tensor K and K is the function of the unsaturated hydraulic conductivity. The variable s represents the sink or source term. The water content in the medium depends on the capillary pressure head in the pores and can be described by the water retention curve.

The governing equation for 2D solute transport processes in the unsaturated soil zone is described as:

$$\frac{\partial \theta_c}{\partial t} + \frac{\partial \rho_s}{\partial t} = \frac{\partial}{\partial x_i}\left(\theta D_{ij} \frac{\partial c}{\partial x_j} \right) - \frac{\partial q_i c}{\partial x_i} + \mu_w \theta c + \mu_s \rho s + \gamma_w \theta + \gamma_s \rho - S c_s$$

where c and s are solute concentrations in the liquid and solid phases, respectively; q_i is the ith component of the volumetric flux density, μ_w and μ_s are first order rate constants for solutes in the liquid and solid phases, respectively; γ_w and γ_s are zero order rate constants for the liquid and solid phases, respectively; ρ is the soil bulk density, S is the sink term in the water flow equation, c is the concentration of the sink term, and D_{ij} is the dispersion coefficient tensor for the liquid phase.

Fuzzy soil hydraulic parameters

The soil hydraulic characteristics in the CHAIN_2D code can be described by nine parameters. If we assume that these input parameters are imprecise and represented by fuzzy numbers, a triangular or trapezoidal membership function can be defined for each parameter. Here, the following input format was selected:

Thr	ths	tha	thm	**Alfa**	**n**	Ks	Kk	thk
.01	0.45	.0	0.45	**.01**	**1.6**	26.495	26.495	0.45
.01	0.45	.0	0.45	**.09**	**2.6**	26.495	26.495	0.45
.01	0.45	.0	0.45	**.04**	**2.3**	26.495	26.495	0.45
.01	0.45	.0	0.45	**.06**	**2.3**	26.495	26.495	0.45

The first line represents the left bound of the α_0-level and the second line the right bound of the α_0-level. The appropriate values for the α_1-level represent the third and fourth line. In the above example, the imprecision parameters are “Alfa” and n (van Genuchten parameters), whereby “Alfa” is characterized by trapezoidal membership function and n is given by triangular form (the values of the α_1-Levels are the same). The interval boundaries of the α-level cuts of the fuzzy inputs create the upper and lower constraints for the optimization algorithm. The imprecision parameters are picked out from an input file.

Fuzzy solutes transport parameters

The input data for transport parameters of solutes in Chain_2D is similar to the definition of the fuzzy soil parameters.

Fuzzy boundary conditions

The CHAIN_2D program implements different types of flow boundary conditions. These conditions are divided in boundary conditions of first and second type. The boundary condition first type represents the pressure head boundary condition, while the boundary condition second specifies fluxes over the model boundary. The Chain_2D assigns a signed integer boundary code for each type. This value must be set positive for boundary conditions of first type and negative or zero for boundary conditions of second type. Here, the following input format was selected:

BC-Code	α_0-left	α_0-right	α_1-left	α_1-right
1	8.0	8.0	–3.0	2.0
–1	–0.2	0.2	–0.05	0.1

Since it is recommended to use only third type boundary conditions for solute transport, no input information is given here.

Fuzzy modelling with Nonlinear Optimization Procedure NLPQLP

This section provides an algorithm for controlling the NLPQLP procedure (Schittkowski, 2006) and the simulation program CHAIN_2D (Fig. 2). First, all fuzzy parameters are read from three separate input files and also membership functions for each parameter and boundary condition are obtained. Then a first estimation for given objective functions is computed (e.g. pressure head, concentration). On the next step, based on these functions values, fuzzy gradients are calculated. After embedding the fuzzy finite element scheme into the program, the NLPQLP procedure is called. Based on sensitivity analysis of the uncertain input parameters by means of gradient and objective function calculations, a number of simulation runs have to be performed until the algorithm finds the minimum value. Since the computation of gradients and objective functions are independent for each parameter, it can be done easily in parallel.

The parameter IFAIL shows the reason for terminating a solution process. Initially, IFAIL must be set to zero. On return, IFAIL could contain the values shown in Table 1:

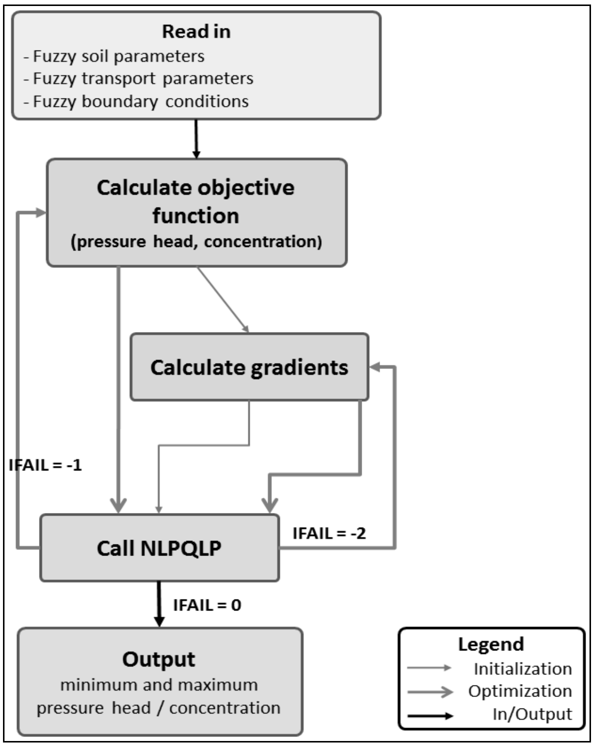

Fig. 2 The main control program for incorporation fuzzified CHAIN_2D into NLPQLP.

Table 1 Description of the return codes of the NLPQLP procedure.

IFAIL	Meaning for further action
–1	Compute objective function and constraint values for all variables and call NLPQLP again.
–2	Compute new gradient values and call NLPQLP again.
0	A minimum objective function is found; Optimality conditions satisfied; the optimization is successfully terminated; the results (value of the objective function and/or parameter combination) can be issued.
>0	An error in NLPQLP occurred (e.g. too many iterations, overflow, under-running, no solution found). In this case it is to be specified by the user.

Fuzzy modelling with fuzzy analysis library

As a second fuzzy modelling tool, Fuzzy Analysis Library (FALIB) was combined with CHAIN_2D program (mainly as a black box model). It has been adapted from the department of civil engineering (Möller *et al.*, 2007). Fuzzy analysis is a multi-purpose framework for predicting system responses with respect to uncertainty of input parameters in a numerically efficient way. It can be combined with any arbitrary analysis software as a linear/nonlinear deterministic solution.

This library uses a mathematical procedure, Fuzzy Randomness, which represents a holistic extension in relation to conventional procedures for the treatment of uncertainties. Using Fuzzy Randomness, all types of uncertainty i.e. stochastic, informal and lexical uncertainties can be described.

The software solution Fuzzy Analysis Library enables the mapping of fuzzy input parameters onto fuzzy result parameters. As a mapping model, any arbitrary analysis program, e.g. FE-Programs, programs for dynamic analysis or safety assessment can be applied. Fuzzy Analysis is performed with the aid of new numerical algorithms for uncertainty processing. The core procedure is called α-level optimization and operates according to a modified evolution strategy and is particularly suitable for nonlinear problems. For the fuzzy input parameters, a sufficiently high number of α-levels are chosen, and the fuzzy results are obtained α-level by α-level via fuzzy analysis.

Fuzzy Analysis Library is available as a FORTRAN Library and can be called and controlled by a simple FORTRAN program.

Example problem

In this section, both above described approaches are applied to a solute transport process with steady-state water flow in the soil zone.

Example: leachate forecast

Leachate forecasts are claimed to evaluate the hazard to the groundwater caused by contamination in the subsurface. The source is described by a concentration either in the leachate or in the soil *vs* time. The hazardous material is subjected to retardation and generally also to degradation or decay during its transport as solute through the unsaturated zone to the saturated or groundwater zone. Because of the complex nature of soils and their transport properties it is recommended to use computer models for the leachate forecast. But quite a lot of parameters are needed to describe the properties of the source, of the transport through a heterogeneous unsaturated zone and also of the climatic boundary conditions and the fluctuating groundwater table. The assessment of these parameters is expensive and prone to errors; the uncertainty of the values is very high.

Figures 3 and 4 show the assessment of the contaminant concentration infiltrated into the groundwater at the transfer zone between the unsaturated and saturated zone (point of evaluation) using NLPQLP and FALIB, respectively.

DISCUSSION AND CONCLUSIONS

In this paper, two successful incorporations of the fuzzy based simulation approached (NLPQLP and FALIB) to the CHAIN_2D computer program are presented to express imprecision and

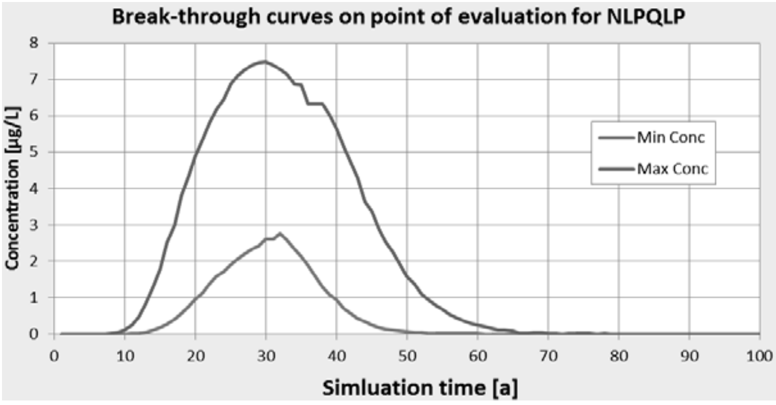

Fig. 3 Break-through curves for the leachate forecast using NLPQLP.

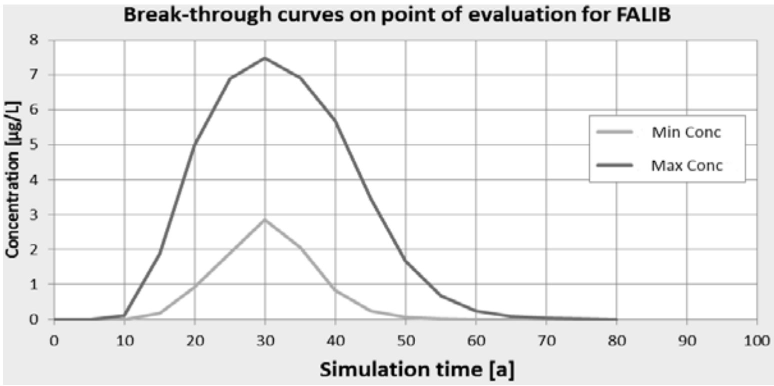

Fig. 4 Break-through curves for the leachate forecast using FALIB.

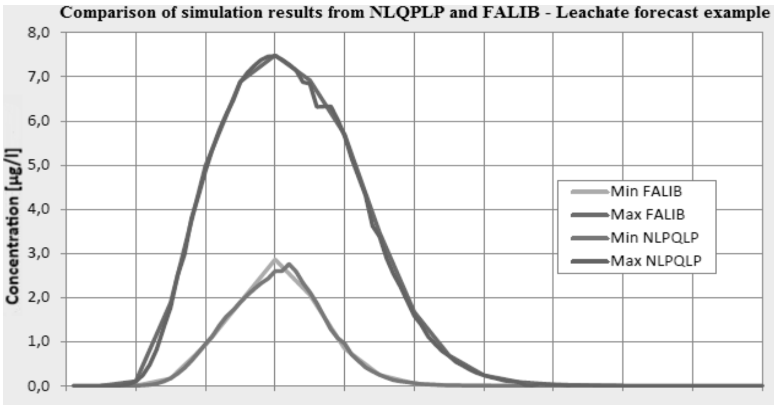

Fig. 5 Comparison of simulation results from NLQPLP and FALIB – leachate forecast test case.

vagueness of soil and solute transport parameters and boundary conditions in a non-probabilistic sense. The programming organization for both codes benefit from parallel executions. Figure 5 shows a comparison of the approaches for the above test case with the same input. It can be seen that quite a good match is achieved and the differences are tolerable.

However, the number of simulations to be carried out for both methods is not directly comparable, since they can be very variable during any time step. FALIB performs a number of run between 500 and 1000, while the number of runs for NLPQLP is between 2 and 3000.

Since both methods use fuzzy theory and are also very well parallelizable, especially for unsteady processes, they can easily be programmed, codes remain simple and require little computer time. One advantage comes from the fact that the result will be two curves (envelopes), which directly reflects uncertainties of the input parameters on the simulation results.

REFERENCES

Bardossy, A. & Duckstein, L. (1995) *Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological and Engineering System.* CRC Press Inc., New York.

Bardossy, A., Bogardi, I. & Kelly, W. E. (1988) Imprecise (Fuzzy) information in geo-statistics. *Math. Geol*. 20(4): 287–311.

Dou, C., Woldt, W., Bogardi, I. & Dahab, M. (1997) Numerical solute transport simulation using fuzzy sets approach. *J. Con. Hyd.* 27(1–2), 107–126.

Huang, Y., Chen, X., Li, Y. P., Huang, G. H. & Liu, T. (2010) A fuzzy-based simulation method for modeling hydrological processes under uncertainty. *Hydrol. Processes* 24, 3718–3732.

Kruse, R., Gebhardt, J. & Klawonn, F. (1993) *Fuzzy-Systeme*. Teubner, Stuttgart.

Möller, B. & Reuter, U. (2007) *Uncertainty Forecasting in Engineering.* Springer Verlag, Berlin.

Schittkowski, K. (2006) NLPQLP: A Fortran implementation of a sequential quadratic programming algorithm with non-monotone line search – user guide. University of Bayreuth, Bayreuth.

Schulz, K. & Huwe, B. (1997) Water flow modeling in the unsaturated zone with imprecise parameters using a fuzzy approach. *J. Hydrol.* 201, 211–229.

Schulz, K. & Huwe, B. (1999) Uncertainty and sensitivity analysis of water transport modeling in a layered soil profile using fuzzy set theory. *J. Hydroinformatics* 1.2, 127–138.

Šimunek, J., Vogel, T. & van Genuchten, M. Th. (1994) The CHAIN_2D code for simulating two-dimensional movement of water, heat, and multiple solutes in variably-saturated porous media. Agricultural Research Service, US Salinity Laboratory. Riverside, California: Research Report. Version 1.1.

Verma, P., Singh, P., George, K. V. & Singh H. V. (2009) Uncertainty analysis of transport of water and pesticide in an unsaturated layered soil profile using fuzzy set theory. *Applied Mathematical Modelling* 33, 770–782.