THE MULTIPLICITY OF MODELS OF FLOW AND TRANSPORT IN UNSATURATED ZONE – A CURSE OR A BLESSING?

Y. Pachepsky[1], A. Guber[1,2], M. T. Van Genuchten[3], J. Šimunek[2], D. Jacques[4], T. J. Nicholson[5], R. E. Cady[5]

[1] EMSL, USDA-ARS, Bldg. 173-BARC-WEST, 10300 Baltimore Ave., Beltsville, MD 20705, USA, Yakov.Pachepsky@ars.usda.gov
[2] Department of Environmental Sciences, University of California Riverside, Riverside, CA 92521, USA, Andrey.Guber@ars.usda.gov, jiri.simunek@ucr.edu
[3] Salinity Laboratory, USDA-ARS, 450 West Big Springs Rd., Riverside CA 92507, USA, rvang@ussl.ars.usda.gov
[4] SCK·CEN, Boeretang 200, 2400 Mol, Belgium, djacques@sckcen.be
[5] ONRR, U.S. Nuclear Regulatory Commission, Two White Flint North, 11545 Rockville Pike, Rockville, MD 20852, USA, tjn@nrc.gov, rec2@nrc.gov

ABSTRACT. Many conceptually different models have been developed to simulate flow and transport in vadose zone. For practical purposes, parameters in these models are often estimated from readily available data using pedotransfer functions. Many pedotransfer functions have been developed, and it is not known which one is the most suitable for a specific site. This adds to the model multiplicity. Selection of the best model is always uncertain at best. The objective of this work is to make an argument for using several models rather than looking for the best model.

One method of using several conceptually different models has been recently termed 'model abstraction'. Model abstraction is the methodology for reducing the complexity of a simulation model while maintaining the validity of the simulation results with respect to the question that the simulation is being used to address. Simpler models can be used along with the more complex to improve the reliability and reduce uncertainty of simulations, to make the modeling and its results more explicable and transparent to the end users, and to enable more efficient use of available resources in data collection and computations. An example of the systematic model abstraction will be presented that has been developed for the case of the infiltration in variably saturated soil under natural rainfall conditions when the complex original model gave inexplicable simulation results. Abstracting base model led to more transparent process description with a simpler model and more robust parameter set without the loss in accuracy of simulating soil water fluxes as the key output.

Another way to take advantage of the model multiplicity is to combine results obtained from several models using weights reflecting models' performance. This approach called multimodeling is appropriate for the case of using several pedotransfer functions to simulate flow and transport in unsaturated zone. The similarity in results from different models may require the dimensionality reductions using, for example, the singular value decomposition. An example of multimodeling for the case of the infiltration in variably saturated soil under natural rainfall conditions will be presented to show that that monitoring of the soil water regime in combination with multimodel simulations can be a viable approach to simulating field water flow in the vadose zone.

Overall, the model multiplicity presents an opportunity to combine expert and factual knowledge on unsaturated flow and transport in different natural settings. A systematic use of several models can advance understanding monitoring data and improve the use of monitoring resources.

1.- Introduction

The complexity of flow and transport pathways at the specific site usually may be easily perceived, but it is often difficult to represent it in mathematical equations of the model without making strong simplifying assumptions (Beven, 2002). This implies that several different models could be consistent with the available observations. Such multiplicity of models reflects the multiplicity of possible conceptual approaches to representation of complex subsurface processes in mathematical form tractable within limitations of existing computer and measurement technologies (Neuman et al., 2003).

The massive effort in developing criteria to select the best model has not lead to a univocal solution. Rather, it has shown that type of model and data, as well as the intended use of the modeling results affect the best model selection. All error-based methods condition the evaluation and comparison of models on the available data. Using the reasonability of forecasts to evaluate models, e. g. with the GLUE methodology (Beven and Binkley, 1992), does not exclude the subjective element in selecting cutoffs and in defining the reasonability. Including measures of model complexity based on the number of model parameters is hardly applicable to nonlinear models of the variably saturated flow. The difficulties in selecting the best from several non-linear models have been encountered in various modeling fields. Eventually, modelers came to realize that simultaneous use of several models can be a beneficial alternative to the quest for the moving target of the best model.

The objective of this paper is to provoke the discussion of the value of concurrent model use. Two systematic approaches to the concurrent use of several models of flow in vadose zone are presented. One approach is the model abstraction which consists in a systematic simplification of a complex model and generating a series of simpler models. Another approach is the multimodeling which consists in assigning weights to the simulation results from different models and using the weighted averages of these results.

2.- Model abstraction in vadose zone modeling

Model abstraction (MA) is defined as *methodology for reducing the complexity of a simulation model while maintaining the validity of the simulation results with respect to the question that the simulation is being used to address.* The need in MA addresses the concerns that the use of overly complex simulation models may cause an excessive burden of data collection and computations as well as difficulties in interpreting simulation results and conveying the simulation approach to both technical and lay audiences. The presumed risk of leaving some important process or feature out often leads the model users to employing fairly complex flow and transport models. However, many of the detailed features, events and processes represented in these complex models may have limited influence at a specific site.

The feasibility of MA has been demonstrated in many research and engineering fields that give ample examples of models having strikingly different complexity and yet the same accuracy, for example, in regional ground water assessments (Kelson et al., 2002), chemical engineering (Diwekar, 1994), marine ecology (Stillman et al., 2000), runoff generation (Jakeman and Hornberger, 1993), population dynamics (Stephens et al., 2002), demographic projections (Smith, 1997), battlefield simulations (Sisti and Farr, 2005), etc. In contaminant hydrology, MA is possible because the complexity of flow and transport pathways at a specific site is easy to perceive but difficult to represent in mathematical equations of the model without making strong simplifying assumptions. Different sets of plausible assumptions lead to different models that are consistent with the available observations. The multiplicity of models reflects the multiplicity of possible conceptual approaches to representation of complex subsurface processes in mathematical form tractable within limitations of existing computer and measurement technologies (Neuman et al., 2003).

2.1. Model Abstraction Techniques

The methodology of MA has been developing for more than 30 years in various research and engineering fields (Meisel and Collins, 1973; Zeigler, 1976; Innis and Rextad, 1983; Fishwick, 1995; Frantz, 1995; Caughlin and Sisti, 1997; Davis and Bigelow, 2003; Van Ness and Scheffer, 2005). Most of the MA techniques are specific to the type of mathematical models used in a specific field. The synopsis and annotated examples of MA in modern flow and transport modeling are presented in (Pachepsky et al., 2007). These MA techniques are summarized in Fig. 1. Two main targets of abstraction are (1) the model structure, i.e. the formal description of specific processes and their interactions that affect flow and transport variables, and (2) the parameter determination, i.e. the estimation of constants and functions serving as coefficients in model equations.

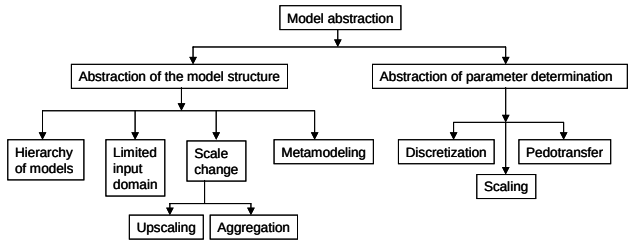

Fig. 1. Categories of model abstraction techniques relevant to flow and transport modeling in subsurface hydrology.

Structure of hydrologic models is changed with MA via (a) using pre-defined hierarchies of models, (b) delimiting input domain, (c) scale change done by either upscaling or aggregation, and (d) metamodeling. A predefined hierarchy contains a series of progressively more simple conceptual and corresponding mathematical representations of porous media. The class of MA techniques based on the delimiting input domain utilises the fact that some features, events, or processes may be not relevant for a given set of scenarios or for a given set of model outputs. Scale change provides transitions between four operational scales – core, profile/pedon, field, and watershed scales that are of interest in contaminant hydrology. MA with scale change alters model equations, variables and parameters with two classes of methods: upscaling and aggregation. Upscaling MA methods use the fine-scale model and the fine scale media properties to derive the coarse-scale model equations and to relate the coarse-scale and fine-scale transport parameters. On the contrary, MA with aggregation does not assume any relationship between model parameters and equations at the fine and at the coarse scales. Parameters of the coarse-scale model are lumped, and are subject to calibration with field data. Aggregation is also being done without the change in the model equations by combining several soil horizons or geologic strata. Metamodeling seeks to simulate the input-output relationships of the complex model with a simple statistical relationship or, recently, with more complex data mining tools such as artificial neural networks.

MA applied to parameterization of hydrologic models does not provide a substitution for the model calibration, but rather seeks to obtain reasonable estimates of parameter values and their variability. Such estimates are useful for setting initial values of parameters for the model calibration, using non-calibrated model in pilot studies and field campaign designs,

assigning values of parameters that are shown to be not sensitive, etc. Abstraction of parameter determination affects parameter estimation via (a) discretization, (b) scaling, and (c) pedotransfer functions. The discretization abstraction techniques replace the continuously varying spatial fields of parameters with piece-wise constant spatial distributions. The scaling abstraction techniques are useful when models are coarsened, i.e. the grid cell size is substantially increased but no changes in the model structure are made. The scale-dependence in model parameters is usually encountered in such cases, and scaling relationships are needed to convert the original parameter values and measurement data to the parameters values of the coarsened model. The pedotransfer techniques convert the readily available data to the hydraulic and transport parameters of the unsaturated flow and transport. Those parameters are notoriously difficult to measure, and a substantial effort has been made to estimate these parameters from the data available from soil survey or borehole logs. The empirical functions used for such estimating are often called pedotransfer functions. Large databases have been assembled to encompass variety of soil properties in developing pedotransfer functions in different parts of the world. New powerful heuristic tools, such as artificial neural networks and regression trees, appeared to be useful in the development of pedotransfer functions. Still, the accuracy of pedotransfer functions outside of their development dataset remains essentially unknown. This is addressed in the second part of this paper.

2.2. Model Abstraction Implementation

The MA implementation has to conform to the requirements of objectiveness, systematic implementation, comprehensiveness, and efficiency (Neuman et al., 2003). The MA process starts with an existing *base* model that can be calibrated and used in simulations. The *key output* of the model is defined that provides the necessary and sufficient information to decide on issues of interest. The base model may need abstraction for one or more of the following reasons (a) the base model includes a complex description of processes that cannot be observed well and yet need to be calibrated; the calibrated values of parameters of those processes are very uncertain, (b) the base model propagates uncertainty in the initial distributions, parameters, and forcing in a manner that creates an unacceptable uncertainty of the key output, (c) the base model produces inexplicable results in terms of the key output, (d) the base model requires an unacceptable amount of resources for computations, data preprocessing, or data post-processing, e.g. the base model is not suitable to be used as a part of a real-time modeling system that requires short computer runtimes, (e) the base model lacks transparency to be explicable and believable to the users of the key output.

The MA process includes the following steps (1) justify the need for the MA, (2) review the context of the modeling problem, (3) select applicable model techniques, (4) determine the MA directions that may give substantial gain, (5) simplify the base model in each direction. Statistical criteria based on guidelines by Hill (1998) can be used to justify the need in MA in case it is related to the uncertainty in calibrated parameter values or in the model key output. The context of the modeling problem has to be reviewed to realize what details and features of the problem are omitted or de-emphasized when the abstraction is performed, and thus to warrant the comprehensiveness and objectiveness of the MA process. MA can lead to simplifications via (a) the number of processes being considered explicitly, (b) process descriptions, (c) coarsening spatial and temporal support, (d) the number of measurements for the reliable parameter estimation, (e) reduced computational burden, (f) data pre-processing and post-processing. Detailed description of the suggested MA process can be found in Pachepsky et al. (2007) where classes of MA techniques are specified that lead to each type of the simplification. Each abstracted model has to be parameterized and confirmed in the uncertainty context.

2.3. Model Abstraction Case Study

The MA case study was design to illustrate the MA process and techniques. The main objective of the test case was to demonstrate how the MA can be applied to understanding and prediction of soil water fluxes at a relatively humid site where transport may be affected by the presence of soil macropores and related preferential flow phenomena. The test site was located in Bekkevoort, Belgium (Jacques, 2000). The 1.5-m deep test trench was dug in a loamy soil and was instrumented to measure soil water content and soil matric potential in 60 locations at five depths hourly, soil temperature in 6 locations at five depths hourly, and soil water fluxes in three locations at two depths once in one to four days. Vegetation was removed and a thin layer of fine gravel covered the soil. Soil water monitoring continued for 384 days. Soil was sampled at 90 locations at three depths to measure soil water retention in laboratory.

The base model was the Richards equation of water flow in variably saturated porous media. The model was calibrated using almost 200,000 measurements of soil water content. The key output of the model was the total soil water flux at two measurement depths and at the bottom of the soil profile over three wetting-drying periods.

Observed and simulated soil water contents are shown in Fig. 2. The topsoil exhibited more weather related variations in the water content than the subsoil. Still, the relatively dry periods between days 120 and 160 and between days 190 and 210 are reflected in lower soil water contents at depths down to 75 cm. Calibrations with the Richards flow model for the five-layer profile resulted in relatively high accuracy (Fig. 2).

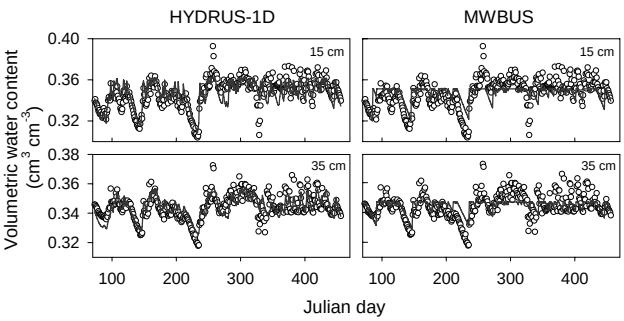

Fig. 2. Observed (symbols) and simulated (using the calibrated model; lines) daily average water contents. Observed values are averages across the transect for each depth and for each day.

The need in MA was justified by the fact that the calibrated base model predicted substantial runoff whereas no runoff was observed at the site during the monitoring period. The simulated infiltration fluxes were much smaller than measured (Fig. 3). The base model produced inexplicable results in terms of the key output (Jacques et al, 2002).

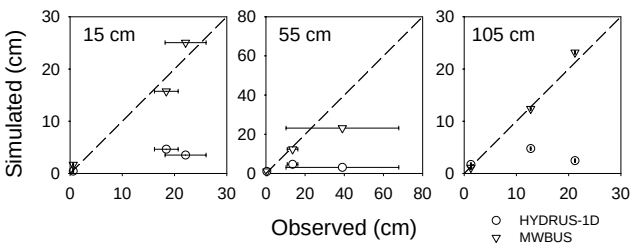

Fig. 3. Simulated and observed cumulative fluxes over three wetting-drying periods at depths of 15 cm, 55 cm, and 105 cm. Error bars show standard deviations. Both HYDRUS-1D and MWBUS have been calibrated.

The models abstraction was designed as shown in Fig. 4. Four classes of applicable model abstraction techniques were selected. The base model 0 was abstracted. Model structure abstraction was achieved using the hierarchy of model (models 1,5,7, 8 in Fig. 4), aggregation (models 2, 6, 7, and 8), and metamodeling (model 4). The hierarchy of models included the Richards equation and a simple soil water budget model. Aggregation was done by replacing the layered soil profile with the homogeneous profile. Metamodeling was done with the backpropagation artificial neural network that used daily and weekly precipitation as the input to estimate monthly cumulative soil water fluxes. Nineteen pedotransfer functions (see below) were used in model abstraction of parameter determination. HYDRUS1-D and MWBUS software packages were used to calibrate the abstracted models and to run Monte Carlo simulations to evaluate the uncertainty in calibrated model outputs.

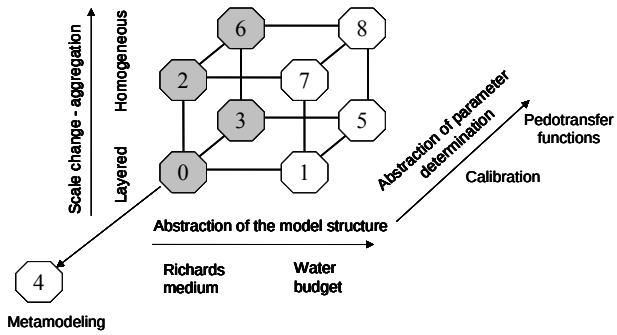

Fig. 4. Design of the model abstraction application. The Richards media was simulated with the HYDRUS-1D software (Simunek et al., 1999); the water budget model was MWBUS (Pachepsky et al., 2007). The model numbering is explained in the text.

The abstraction with the hierarchy of models was useful. The simple soil water budget model was less accurate in predictions of soil water content compared with the Richards model (Fig. 2). However, unlike the more complex Richards model, this simple soil water budget model correctly predicted the absence of runoff and measured cumulative soil water fluxes (Fig. 3). The prediction of runoff was an artifact of the Richards model calibration in absence of measured boundary fluxes. This abstracted model appeared to be instrumental in both explaining behavior of the complex model and in predicting the key output – soil water fluxes.

The abstraction with aggregation was not useful in this case. The Richards model was less accurate with respect to soil water contents and continued to generate large simulated runoff when a homogeneous soil layer was introduced. The neural network metamodel was extremely accurate in estimating cumulative soil water fluxes. It was also five orders of magnitude faster than the numerical algorithm coded in HYDRUS-1D. Results of the abstraction of parameter determination are described in the next section.

3.- Multimodeling in simulations of flow in vadose zone

Modeling of flow in variably saturated soils requires water retention and hydraulic conductivity parameters that are impractical to measure for large-scale projects. As opposed to saturated flow, the nonlinearity of the unsaturated flow constitutive (hydraulic) properties gravely complicates calibration of the variably saturated flow models against field monitoring data. As an alternative, pedotransfer functions are routinely used to relate the hydraulic parameters to readily available data on soil properties that can found on soil maps or in soil survey reports.

Since pedotransfer functions are empirical regression-type relationships, their accuracy outside of its development region is essentially unknown. A wealth of pedotransfer information has recently accumulated in nearly all parts of the world. Unfortunately, no good method currently exists to decide which pedotransfer function model should be used for a specific site or application (Pachepsky and Rawls, 2004). Climate predictions faced similar uncertainties in model

selection in the 1980s. To deal with these uncertainties, multimodel prediction has emerged as a popular technique in climate prediction (Barnston et al., 2003; Palmer et al., 2000; 2004; Shukla et al., 2000). The objective of multimodel prediction is to reduce modeling errors by combining forecasts of various independent models (Doblas-Reyes et al. 2000; Doblas-Reyes et al., 2005; Hagedorn et al., 2005).

The efficiency of multimodel predictions was demonstrated for streamflow forecasts by Regonda et al. (2006). Multimodel prediction methods are now also being used in groundwater modeling. For example, Ye et al. (2004) suggested using weighted results of several spatial variability models for unsaturated fractured tuff to run flow simulations for situations where standard information criteria provide an ambiguous ranking of the models such that it does not justify selecting one of them and discarding all others. Methods of weighing predictions obtained from different groundwater flow models have been discussed in detail by Poeter and Anderson (2005). Recently the feasibility of applying multimodel simulations to water flow in the vadose zone was demonstrated (Guber et al., 2006).

Since its introduction by Bates and Granger (1969), multimodel prediction has been subject to much debate that can be summarized into two questions: (a) is a multimodel prediction better than the single best forecast, and (b) what is the best approach to weigh predictions obtained with the different models. In this work, the multimodel simulation approach has been applied to the analysis of a large database of information on unsaturated soil water dynamics along a 6-m transect (seee section 2.3). We were particularly interested in the effects of using different methods of weighting the multimodel simulations of the field water regime as obtained with different pedotransfer functions.

3.1. Methods for Multimodel Simulations

We consider the case where several PTF models exist to estimate parameters in the constitutive relationships (soil hydraulic properties) needed for application of the governing Richards equation to particular variably saturated flow problems. The multimodel simulation may then mean either (a) combining output of several PTFs for the hydraulic parameters into a single parameter set and then running the flow model with this parameter set, or (b) running the flow model with outputs of individual PTFs for the hydraulic parameters and then combining the obtained outputs of the flow model. In this study we consider the multimodel simulations in the sense of option (b), which is commonly used in meteorological predictions.

Combining simulation results from N models is carried out as:

$$S = \overline{Y} + \sum_{i=1}^{N} w_i \left(F_i - \overline{F}_i \right) \qquad (1)$$

where S is the multimodel simulation result, $\overline{Y}$ is the mean of the observed values, "i" is the model number (i=1,2,…, N), N is the total number of PTF models, F_i is the simulation result from model "i", $\overline{F}_i$ is the mean of the simulation results obtained with model "i", and w_i are the weights on the simulation results from individual models. The weights are found by using observations and simulations over a training period. The weighing methods are evaluted in terms of the accuracy of the multimodel (i.e., errors in reproducing the training or hindcast datasets), and the reliability of the multimodel (i.e., errors in reproducing the test datasets). Equation (1) shows that over the training period the multimodel prediction equation relates deviations from the average of observed values (S-$\overline{Y}$) with deviations from the average of the simulated values (F-$\overline{F}_i$).

Existing methods to combine predictions according to (1) differ in the way in which the weights w_i are obtained. Such methods have been reviewed by Clemen (1989), Burnham and Anderson (2002), Armstrong (2001), and Jolliffe and Stephenson (2003), among others. We applied and compared the following six methods.

1. Using only the model that has been the best with the training dataset.
2. Using the arithmetic average of results from all models.
3. Superensemble forecasting (Krishnamurti et al., 2000). If the measured values of the variable to be simulated are Y, then the weights w_i are determined by treating Eq. (1) as a multiple linear regression equation with $S - \overline{Y}$ as the dependent variable and $F_i - \overline{F}_i$ as the independent variable.
4. Regression with singular value decomposition. The main problem of using regression to search for w_i is the existence of relatively high correlations between F_i. Having correlated independent variables in regression, or multicollinearity, does not preclude using the resultant regression for predictions within the range of observation (Neter and Wasserman, 1974). However, it leads to very inaccurate regression coefficients that are not easily interpreted, and a matrix for the system of equations for calculating weights that is numerically close to singular (Kharin and Zwiers, 2002). Yun et al. (2003) showed that improvements in superensemble forecasts can be achieved by applying the singular value decomposition (SVD) technique to solve the system of equations for the coefficients w_i.
5. Bayesian model averaging (BMA). The method is based on th assumption that for any given forecast there is a "best" model; while we do not know what that model is, uncertainty about the best model can be quantified using Bayesian model averaging (Raftery et al. (2003). The weight w_i is the posterior probability of the simulation f_i being the best one, and is based on the performance of model "i" in the training period. The w_i add up to one.

6. Using information theory. The Akaike information criteria is used to derive weights for individual models (Burnham and Anderson, 2002; Poeter and Andersen, 2005)

3.2. Pedotransfer Functions and Flow Simulations

The literature was searched for pedotransfer functions to estimate the soil water retention and hydraulic conductivity properties from soils data available at the site. We only used PTFs that had been developed from relatively large (more than 200 samples) databases. To estimate soil water retention, we selected 19 PTFs developed in different regions. Equations for the PTFs were taken from the appendices of Guber et al. (2006) and Pachepsky et al. (2007)[1]. Five of the PTFs (those by Campbell and Shiozawa, 1992; Mayr and Jarvis, 1999; Rawls and Brakensiek. 1985; Saxton et al., 1986; and Williams et al., 1992) estimated parameters of the equation of Brooks and Corey (1964), while four other equations (Varallyay et al., 1982; Vereecken et al., 1989; Wösten et al., 1999) estimated parameters of the van Genuchten (1980) equation. Nine pedotransfer functions (Baumer, 1992; Canarache, 1993; Gupta and Larson, 1979; Hall et al., 1977; Petersen et al, 1968; Rajkai and Várallyay, 1992; Rawls et al., 1982; Rawls et al., 1983; Tomasella and Hodnett, 1998) estimated water contents at several fixed pressure heads. The van Genuchten parameters were evaluated by fitting the van Genuchten equation to the water retention points obtained from those PTFs. Two PTFs (Vereecken et al., 1989; Varallyay et al., 1982) used van Genuchten equation assuming m=1. We also used the Rosetta software (Schaap, 2004) to generate van Genuchten parameters from texture and bulk density.

We could find only two PTFs for the saturated hydraulic conductivity, K_{sat}, that were developed and/or tested using large databases (Rawls et al., 1998; Wösten et al., 1999). Since the multimodel prediction did not seem feasible with only two PTFs, we used the PTF developed by Rawls et al. (1998), which gives estimates of the median and corresponding 25% and 75% probability levels, to generate random values of K_{sat}.

The HYDRUS-1D software (Simunek et al., 1998) was used to run the simulations. This software gives options to run simulations either with the Brooks-Corey water retention equation or the van Genuchten equation .

3.3. Multimodeling results

The Richards equation was used to simulate flow at the field site using hydraulic parameters obtained with the pedotransfer functions. An example of results of these simulations is shown in Fig. 5 for the 15-cm depth. The scatter in the simulated water contents was found to be substantial. In general, the simulated soil water contents were lower than the observations, and also decreased somewhat faster during dry periods. Temporal variations in the soil water content were more pronounced with some PTFs than with others. However, inspection of the Fig. 5 shows that the different pedotransfer functions produced almost parallel time series that were shifted in various degrees with respect to each other. Correlation coefficients between simulated soil water contents exceeded 0.95 in more than 48% of the cases, while very few correlation coefficients were less than 0.5. Similar results were obtained for the other depths.

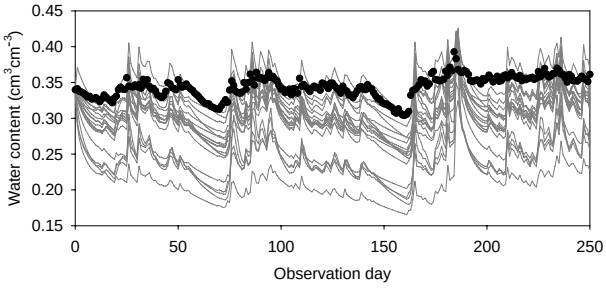

Fig. 5. Daily average water contents at 15 cm depth; lines - simulated with 19 pedotransfer functions, circles – observed.

The accuracy and reliability of different methods to build the multimodel prediction is compared in Fig. 6. The unconstrained superensemble was found to be the worst weighing method characterized by high accuracy and very low reliability (data not shown in Fig. 4 because they are outside the axis scales). Assigning equal weights (simple averaging) was the second worst method. This method also had testing RMSE values larger than the training RMSEs, although the difference was not as dramatic as with the superensemble method. The accuracy of the unconstrained superensemble was the worst of all methods except the superensemble. Bayesian model averaging fared only slightly better than simple averaging. Using only the best model was a reasonably good approach in terms of accuracy. However, the reliability of this method was almost the same as with simple averaging and Bayesian averaging. Results of the information theory method were no different from the using the best model. Using singular value decomposition in the superensemble appeared to be the best method in that both the accuracy and reliability RMSE values almost two times less than those of all other methods considered.

The results in Fig. 6 further shows that training with one month of data leads to lower reliability than when two or three months of data is used in the training. We found a weak trend of increasing accuracy with the duration of training. Fig. 4 also compares multimodel results with results obtained with the Richards model calibrated on data from the entire observation period. The error bars for the calibrated model were obtained from the statistical distributions of RMSE over

[1]The FORTRAN code to estimate water retention with pedotransfer functions used in this study is available upon request from the corresponding author.

the same training and testing periods that were used for the multimodel predictions.

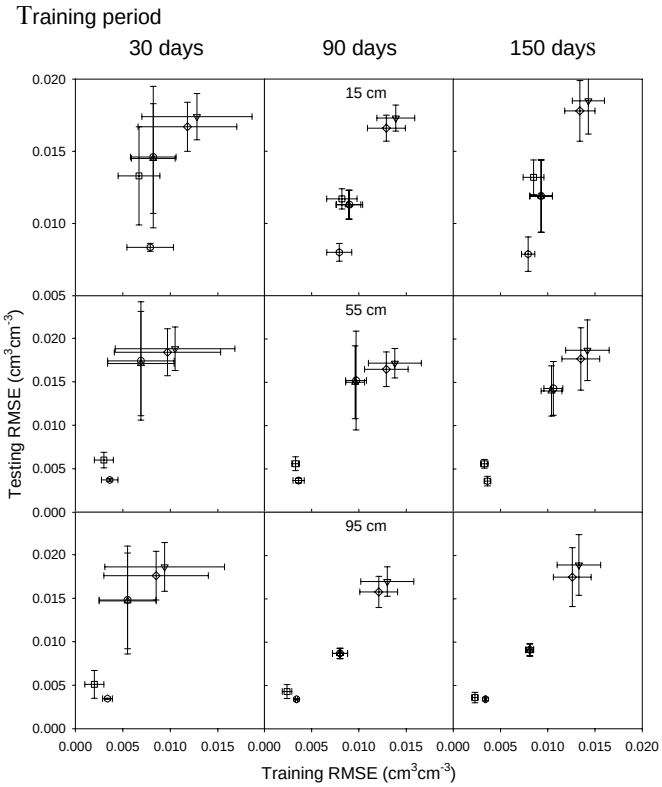

Fig. 6. Root-mean squared errors of multimodel water content simulations for the training and testing datasets; results are for the best in training model (○), the arithmetic average of all individual models (▽), the superensemble with singular value decomposition (□), Bayesian averaging (◊), the use of the information theory (△), and calibrated Richards equation (⬡).

The data in Fig. 6 show that the accuracy of the superensemble with singular value decomposition (SVD) is comparable to the accuracy and reliability of the calibrated Richards equation, even when weights are derived from one month of observations. The overall accuracy and reliability of the SVD multimodel was slightly less, or comparable to, the accuracy of the calibrated model.

4.- Concluding remarks

We observed two types of possible advantages from the use of multiple models. Model abstraction explained the strange behavior of the complex model, and provided the correct description of the system behavior and plausible parameter ranges. The similarity in accuracy between the calibrated Richards model for the layered soil and the multimodel predictions with pedotransfer functions suggest the interesting possibility of using soil moisture monitoring data to obtain a more accurate predictive soil water flow model. Monitoring soil moisture and finding weights for the multimodel prediction may well be a very viable approach to simulating field water flow in the vadose zone, rather than only calibrating the flow model which is known to be a daunting task.

The intensive use of models in subsurface contaminant hydrology has resulted in development of many MA techniques that to-date have been used mostly in research. Potential benefits of MA include improvement of understanding and communication of modeling results, more robust predictions, and better understanding of essential factors and their representation in models. This makes MA an attractive methodology for engineering modeling applications. The MA process can be set as a transparent step-by-step formalized procedure of justification of the use of a simplified model. An important feature of models abstraction is the explicit treatment of model structure uncertainty. The model structure, along with the data uncertainty, and scenario uncertainty, is known to introduce the uncertainty in modeling results. Unlike the uncertainty in input data, in model parameters, and in scenarios, the effect of the model structure uncertainty on the uncertainty in simulation results is usually impossible to quantify in statistical terms. Using MA, a series of models with feasible structures can be built and evaluated in a systematic manner. Each of the models is evaluated from results of an ensemble of simulations by its accuracy to measurement data and by its predictions with respect to scenarios that have not been observed.

The list of methods to build the multimodel in this work is far from exhaustive. Young (2002) and Regonda et al. (2006) list several other methods. Recent developments include using patterns found in experimental data to adjust the weights of individual forecasts in multimodel predictions (Zheng et al., 2004). There is no reason why a nonlinear combination of individual simulations could not be used in multimodels. Using data mining techniques, such as artificial neural network, may be beneficial.

Overall, the model multiplicity presents an opportunity to combine expert and factual knowledge on unsaturated flow and transport in different natural settings. A systematic use of several models can advance understanding monitoring data and improve the use of monitoring resources.

References

Armstrong, J.S. 2001. Combining forecasts. *In*: Principles of Forecasting: A Handbook for Researchers and Practitioners, J.S. Armstrong (ed.), Kluwer Academic, 417–439.

Barnston, A.G., S. Mason, L. Goddard, D.G. Dewitt, and S.E. Zebiak. 2003. Multimodel ensembling in seasonal climate forecasting at IRI. Bull. Amer. Meteor. Soc. 84:1783–1796.

Bates, J.M., and C.W.J. Granger. 1969. The combination of forecasts. Operations Research Quarterly, 20: 451-468.

Baumer, O.M. 1992. Predicting unsaturated hydraulic parameters. p.341-354. *In*: M.Th. van Genuchten et al. (ed.) Proc. Int. Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California, Riverside, CA.

Brooks, R.H., and A.J. Corey. 1964. Hydraulic properties of porous media. Hydrol. Paper 3. Colorado State Univ., Fort Collins.

Burnham, K.P., and D.R. Anderson. 2002. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, New York, USA. 353 pp.

Campbell, G.S., and S. Shiozawa. 1992. Prediction of hydraulic properties of soils using particle size distribution and bulk density data. p. 317-328. *In* M. Th. van Genuchten et al. (eds.) Proc. Int. Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California, Riverside, CA.

Canarache, A. 1993. Physical-technological maps - a possible product of soil survey for direct use in agriculture. Soil Technol. 6:3-16.

Caughlin, D. and Sisti, A.F. 1997. A Summary of Model Abstraction Techniques. In: Proceedings of Enabling Technology for Simulation Science I Conference, (Orlando FL., 22-24 April 97). SPIE Conf. 3083, pp 2-13.

Clemen, R. T. 1989. Combining forecasts: A review and annotated bibliography. Int. J. of Forecasting, 5, 559-583.

Davis, P. K., and Bigelow, J. H. 2003. Motivated metamodels: synthesis of cause-effect reasoning and statistical metamodeling. MR-1570. RAND, 2138, Santa Monica, CA.

Diwekar, U. M. 1994. How Simple Can It Be? - A Look at The Models for Batch Distillation. Computers & Chemical Engineering 18: S451—S457.

Doblas-Reyes, F.J., M. Déqué, and J.-Ph. Piedelievre. 2000. Model and multimodel spread and probabilistic seasonal forecasts in PROVOST. Quart. J. Royal. Meteorol. Soc. 126:2069-2088.

Doblas-Reyes, F.J., R. Hagedorn, and T.N. Palmer. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting. Part II: Calibration and combination. Tellus A, 57:234-252.

Fishwick, P. A. 1995. Simulation Model Design and Execution. Englewood Cliffs: Prentice-Hall, 1995

Frantz, K. F. 1995. A taxonomy of model abstraction techniques," In: Alexopoulos, C., Kang, K., Lilegdon, W. R., and Goldsman, D. (Eds.) Proceedings of the 1995 Winter Simulation Conference, pp.1413-142.

Graham, R. J., A.D.L.Evans, K.R. Mylne, M.S.J. Harrison, and K.B. Robertson. 2000. An assessment of seasonal predictability using atmospheric general circulation models. Quart. J. Royal Meteorol. Soc. 126: 2211–2240.

Guber, A. K., Y.A. Pachepsky, M.T. van Genuchten, W.J. Rawls, D. Jacques, J. Simunek, R.E. Cady, T.J. Nicholson. 2006. Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. Vadose Zone J., 5:234-247.

Gupta, S.C., and W.E. Larson. 1979. Estimating soil water retention characteristics from particle-size distribution, organic matter percent, and bulk density. Water Resour. Res. 15(6): 1633-1635.

Hagedorn, R., F.J. Doblas-Reyes, and T.N. Palmer. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. Tellus, 57A: 219-233.

Hall, D.G.M., M.J. Reeve, A.I. Thomasson, and V.F Wright. 1977. Water retention, porosity and density of field soils. Soil Surv. Tech. Monogr. 9. Rothamsted Experimental Station, Lawes Agricultural Trust, Harpenden, UK.

Hill, C. M. 1998. Methods and Guidelines for Effective Model Calibration. U. S. Geological Survey Water Resources Investigations Report 98-4005, Denver, CO, 90 pp.,

Innis, G., and E. Rextad. 1983. Simulation model simplification techniques. Simulation, 41:7-15.

Jacques, D. 2000. Analysis of Water Flow and Solute Transport at The Field Scale. PhD Thesis no 454. Faculteit Land bouwkundige en Toegepaste Biologische Wetenschappen, K.U. Leuven, Belgium, 255 pp.

Jacques, D., Simunek, J., Timmerman, A., and J. Feyen. 2002. Calibration of Richards and Convection–Dispersion Equations To Field-Scale Water Flow And Solute Transport Under Rainfall Conditions. J. of Hydrology, 259: 15-31.

Jakeman, A.J. and G.M. Hornberger.1993. How Much Complexity is Warranted in a Rainfall-Runoff Model? Water Resources Research, 29: 2637-2649,

Jolliffe, I.N., and D.B. Stephenson, 2003. Forecast Verification: A Practitioner's Guide in Atmospheric Science. Wiley and Sons, 240 pp.

Kelson, V.A., Hunt, R.J., and H.M. Haitjema. 2002. Improving a Regional Model Using Reduced Complexity and Parameter Estimation. Ground Water, 40: 132-143.

Kharin, V., and F.W. Zwiers. 2002. Climate predictions with multimodel ensembles. J. Climate 15: 793-799.

Krishnamurti, T.N, C.M. Kishitawal, Z. Zhang, T. Larow, D. Bachiochi, and E. Williford. 2000. Multimodel ensemble forecasts for weather and seasonal climate. J. Climate 13:4196–4216.

Mayr, T., and N.J. Jarvis. 1999. Pedotransfer functions to estimate soil water retention parameters for a modified Brooks-Corey type model. Geoderma 91:1-9.

Meisel, W. S. and D. C. Collins.1973. "Repromodeling: An approach to efficient model utilization and interpretation. IEEE Transactions on Systems, Man, and Cybernetics SMC-3: 349-358,

Neter, J., and W. Wasserman. 1974. Applied linear statistical models: Regression, analysis of variance, and experimental designs. Richard D. Irwin, Homewood, IL.

Neuman, S. P., Wierenga, P. J., and Nicholson, T. J. 2003. A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites. NUREG/CR 6805. U. S. Nuclear Regulatory Commission. Washington, D.C. 20555-0001.

Pachepsky, Y. A., Guber, A. K., Van Genuchten, M. T., Nicholson, T. J., Cady, R. E., Simunek, J., Schaap, M. C. 2007. Model Abstraction Techniques for Soil Water Flow and Transport. NUREG/CR-6884. U.S. Nuclear Regulatory Commission, Washington, D.C. (available at http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr6884/)

Pachepsky, Y.A., and W.J. Rawls. 2004. Status of pedotransfer functions. Pp. vii-xviii. *In*: Pachepsky, Y. A., and W. J. Rawls (eds.), Development of Pedotransfer Functions in Soil Hydrology. Elsevier, Amsterdam.

Palmer, T.N, A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Díaz, F.J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J-F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A.P. Morse, B. Orfila, P. Rogel, J-M. Terres, and M.C. Thomson. 2004. Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). Bull Am. Meteor. Soc 85:853–872.

Palmer, T.N., C. Brankovic, and D.S. Richardson. 2000. A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. Quart. J. Royal Meteor. Soc. 126:2013–2035.

Petersen, G.W., R.L. Cunningham, and R.P. Matelski. 1968. Moisture characteristics of Pensylvania soils. 1. Moisture retention as related to texture. Soil Sci. Soc. Am. Proc. 32: 271-275.

Poeter, E., and D. Anderson. 2005. Multimodel Ranking and Infe-rence in Ground Water Modeling. Ground Water 43:597-605.

Raftery , A.E., R. Balabdaoui , T. Gneiting, and M. Polakowski. 2003. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. Technical Report no. 440, Department of Statistics, University of Washington.

Rajkai, K., and Gy. Várallyay. 1992. Estimating soil water retention from simpler properties by regression techniques. p.417-426. *In*: M. Th. van Genuchten et al.(eds.). Proc. Int. Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils University of California Riverside, Riverside, CA.

Rawls, W.J., and D.L. Brakensiek. 1985. Prediction of soil water properties for hydrologic modeling. p.293-299. *In*: E. B. Jones and T. J. Ward (eds.), Proc. Symp. Watershed Management in the Eighties. April 30-May 1, 1985, Denver, CO., Am. Soc. Civil Eng., New York, NY.

Rawls, W.J., D. Giménez, and R. Grossman. 1998. Use of soil texture, bulk density and slope of the water retention curve to predict saturated hydraulic conductivity. Trans. ASAE 41: 983-988.

Rawls, W.J., D.L. Brakensiek, and B. Soni. 1983. Agricultural management effects on soil water processes, Part I. Soil water retention and Green-Ampt parameters. Trans. ASAE 26:1747 -1752.

Rawls, W.J., D.L. Brakensiek, and K.E. Saxton.1982. Estimation of soil water properties. Trans. ASAE 25:1316-1320.

Regonda, S.K., B. Rajagopalan, M. Clark, and E. Zagona. 2006. A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin. Water Resources Research, Vol. 42, W09404, doi:10.1029/2005WR004653.

Saxton, K.E., W.J. Rawls, J.S. Romberger, and R.I. Papendick. 1986. Estimating generalized soil-water characteristics from texture. Soil Sci. Soc. Am. J. 50:1031-1036.

Schaap, M.G. 2004. Graphic user interfaces for pedotransfer functions. p. 349-356. *In*: Pachepsky, Y. A., and W. J. Rawls (eds.), Development of Pedotransfer functions in Soil Hydrology. Elsevier, Amsterdam.

Shukla J, Anderson J, Baumhefner D, Brankovic C, Chang Y, Kalnay E, Marx L, Palmer T, Paolino D, Ploshay J, Schubert S, Straus DM, Suarez M, Tribbia J. 2000. Dynamical seasonal prediction. Bull Amer. Meteor. Soc 81:2593–2606

Simunek, J., M. Sejna, and M.Th. van Genuchten. 1998. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 2.0, IGWMC - TPS - 70, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 202pp.

Sisty, A.F. and S.D. Farr. 2005. Model Abstraction Techniques: An Intuitive Overview. Air Force Res. Laboratory/IFSB, Rome, NY(http://www.rl.af.mil/tech/papers/ModSim/ ModAb_Intuitive.html).

Smith, S.K. 1997. Further Thoughts on Simplicity and Complexity in Population Projection Models. International Journal of Forecasting, 13: 557 – 565.

Stephens, P. A., Frey-Roos, F. Arnold, W., and Sutherland, W. J. 2002. Model Complexity and Population Predictions. The Alpine Marmot As A Case Study. Journal of Animal Ecology, 71: 343–361.

Stillman R. A. , McGrorty, S., Goss-Custard, J. D. and A. D. West. 2000. Predicting Mussel Population Density and Age Structure: The Relationship Between Model Complexity and Predictive Power. Marine Ecology Progress Series, 208: 131–145.

Tomasella, J., and M.G. Hodnett. 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. Soil Sci. 163:190-202.

Van Genuchten, M. Th. 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44: 892-898.

Van Ness, E. H., Scheffer, M. 2005. A strategy to improve the contribution of complex simulation models to ecological theory. Ecological Modelling, 185: 153-164..

Varallyay G., K. Rajkai, Ya.A. Pachepsky, and R.A. Shcherbakov. 1982. Mathematical description of soil water retention curve. Pochvovedenie, (4):77-89 (in Russian).

Vereecken, H., J. Maes, J. Feyen, and P. Darius. 1989. Estimating the soil moisture retention characteristics from texture, bulk density and carbon content. Soil. Sci. 148:389-403.

Williams, J., P. Ross, and K. Bristow. 1992. Prediction of the Campbell water retention function from texture, structure, and organic matter. p. 427-442. *In*: M. Th. van Genuchten et al. (eds.) Proc. Int. Workshop on Indirect methods for Estimating the Hydraulic Properties of Unsaturated Soils. University of California Riverside, Riverside, CA.

Wösten, J.H.M., A. Lilly, A. Nemes, and C. Le Bas. 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90:169-185.

Ye, M., S.P. Neuman, and P.D. Meyer. 2004. Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. Water Resour. Res., 40, W05113, doi:10.1029/2003WR002557.

Young, G., 2002. Combining forecasts for superior prediction. Preprints, 16th Conf. on Probability and Statistics in the Atmospheric Sciences, Orlando, FL, Amer. Meteor. Soc., 107– 111.

Yun, W. T., L. Stefanova, and T.N. Krishnamurti. 2003. Improvement of the Multimodel Superensemble Technique for Seasonal Forecasts. J. of Climate 16: 3834-3840.

Zeigler, B. 1976. Theory of modeling and simulation. New York, New York, Wiley and Sons.

Zheng, X., M. Sugi, and C.S. Frederiksen. 2004. Interannual variability and predictability in an ensemble of climate simulations with the MRI-JMA AGCM. J. Meteor. Soc. Japan 82:1–18.