Model Abstraction Techniques Related to Parameter Estimation and Uncertainty

Yakov Pachepsky, Rien van Genuchten, Ralph Cady,
and Thomas J. Nicholson

in

Proceedings of the

International workshop on

Uncertainty, Sensitivity, and Parameter Estimation for Multimedia Environmental Modeling

Held August 19-21, 2003, at the U.S. Nuclear Regulatory Commission Headquarters
11545 Rockville Pike, Rockville, Maryland, USA

Pachepsky, Y., R. van Genuchten, R. Cady, and T. J. Nicholson. 2004. Model abstraction techniques related to parameter estimation and uncertainty. *In*: T. J. Nicholson, J. E. Babendreier, P. D. Meyer, S. Mohanty, B. B. Hicks, and G. H. Leavesley (eds.), Proc. Int. Workshop "*Uncertainty, Sensitivity, and Parameter Estimation for Multimedia Environmental Modeling*", August 19-21, 2003, Rockville, MD, NUREG/CP-0187, pp. 107-109, U.S Nuclear Regulatory Commission, Rockville, MD.

Model Abstraction Techniques Related to Parameter Estimation and Uncertainty

Yakov Pachepsky[1], Rien van Genuchten[2], Ralph Cady[3], and Thomas J. Nicholson[3]

[1]Environmental Microbial Safety Laboratory, USDA-ARS, ypachepsky@anri.barc.usda.gov;
[2]U.S. Salinity Laboratory, USDA-ARS, rwang@ussl.ars.usda.gov;
[3]U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research,
TJN@nrc.gov and REC2@nrc.gov

Model abstraction is a 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 (Frantz, 2002). The need for model abstraction has been recognised in simulations of complex engineering and military systems that show that increased level of detail does not necessarily imply increased accuracy of simulation results, but usually increases computational complexity and may make simulation results more difficult to interpret. Similar observations have been made for simulations of subsurface flow and transport problems. Model abstractions that lead to reduced computational overhead and complexity can enable risk assessments to be run and analyzed with much quicker turnaround, with the potential for allowing further analyses of problem sensitivity and uncertainty. In addition, because of the highly heterogeneous nature of the subsurface, the issues of data collection and parameter estimation are as essential as computational complexity. While increased levels of detail in the data currently do not necessarily imply increased accuracy of the simulations, it usually does imply increased data collection density. Finally, model abstraction is important in enhancing communication. Simplifications that result from appropriate model abstractions may make the description of the problem more easily relayed to and understandable by others, including decision-makers and the public. It is often imperative to explicitly acknowledge the abstraction strategy used and its inherent biases, so that the modeling process is transparent and tractable.

Model abstraction explicitly deals with uncertainties in model structure. Model abstraction techniques and examples of their application in subsurface flow and transport include (a) using pre-defined hierarchies of models, (b) simplifying process descriptions based on the specific range of input parameters, i.e., reducing dimensionality, (c) parameter elimination based on simulation results, i.e., sensitivity analysis, (d) combining system states whose distinctions are irrelevant to the simulation output, i.e., combining individual stream tubes in a stochastic transport model, or upscaling based on aggregation, (e) dividing a model into loosely connected components, executing each component separately, and searching for constraints that execution of one component can impose on other components, i.e., running a flow model independently of the transport model, (f) combining states involving similar sequences and distinctions among the individual sequences that are irrelevant to the final outcome, i.e., abstracting the iterative plume construction to the transport of particle ensembles undergoing non-Brownian motion, (g) replacing continuous variables by class variables, i.e., using regression trees to develop pedotransfer functions used for hydraulic parameters estimations, or genetic algorithms in model calibration optimization, (h) temporal aggregation, i.e., replacing several closely-spaced events with a single event, (i) aggregating entities in a natural hierarchical structures, i.e., replacing a heterogeneous soil profile with an equivalent homogenous profile, (j) function aggregation to provide a coarser list of states or output information from existing entities, i.e., representing the water regime of a soil layer by means of either infiltration or evaporation, while neglecting redistribution, (k) using probabilistic inputs to develop lumped models, i.e., statistical averaging of flow and transport behavior for temporal and spatial upscaling, (l) using look-up tables to simplify the input-output transformation within a model or model component by means of a decrease in computational effort, (m) rule-based solutions of model equations, i.e., using cellular automata in flow and transport problems, (n) metamodeling with neural networks, i.e., neural network approximations of a range of output scenarios for a particular remediation site, (o) spatial correlation-based metamodeling, i.e., using spatial correlations in flow and transport data assimilation, and (p) wavelet-based metamodeling.

Applications of model abstraction require criteria to select a simpler model, justify validity, and quantify questions being addressed. The criteria have yet to be developed based on quantified uncertainty and cost-benefit analyses. For purposes of vadose zone water flow and solute transport modeling, simplicity may be related to the number of processes being considered explicitly in the simulations, details of the discretization, runtime, number of measurements for parameter estimation, and correlations among parameters. Validity must be related to variability in data and to the uncertainty in the simulation results. Questions being addressed relate to specific outputs defined in terms of probability thresholds or physical thresholds for pre-defined locations in space in time.

During the first phase of this project we developed prospective directions for testing the model abstraction process using high-density data sets for water flow in typical environments. We concluded that, for model abstraction in flow and transport model *development*, the prospective direction should be on *model structure* modifications, whereas the prospective direction for model abstraction in flow and transport model *parameterization* should be on *model behavior* modification. Similarly, the prospective direction for model abstraction in flow and transport *simulations* should be on *model form modification*. Field data sets for a humid environment (Fig. 1) and for an arid environment (Fig. 2) were selected based on their completeness and complexity to explore specific issues, e.g., complex three-dimensional processes rendered as two and one-dimensional processes, or replacing directly measured soil hydraulic properties by pedotransfer function estimates. Future work with field data sets will compare the efficiency of model analysis techniques and provide a basis for developing rule-based strategies for model abstraction in the area of subsurface water and solute transport.

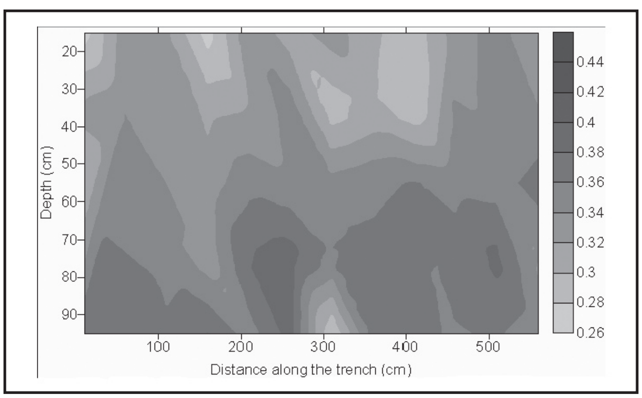

Figure 1. A snapshot of soil water contents monitored for 1 year (along with pressure heads and solute concentrations) along a trench in a loamy soil at the Bekkevoort site, Belgium.

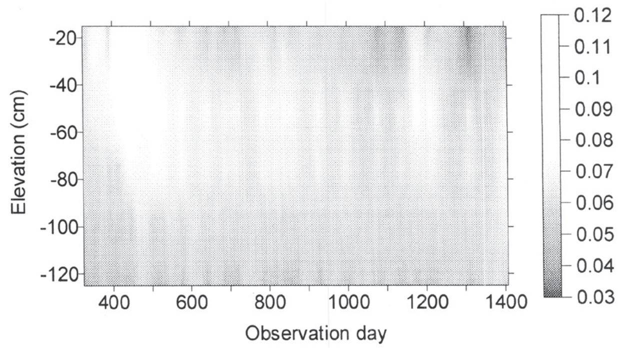

Figure 2. Image map of soil water contents monitored for 3 years (along with soil temperatures) at the USGS Amargosa Research Site, Nevada.

References

Diederik, J., J. Simunek, A.Timmerman and J.Feyen, 2002, "Calibration of Richards' and convection–dispersion equations to field-scale water flow and solute transport under rainfall conditions," *Journal of Hydrology*, 259: 15–31. (Belgium field study)

Frantz, Frederick K., "A Taxonomy of Model Abstraction Techniques," Computer Sciences Corporation, One MONY Plaza, Mail Drop 37-2, Syracuse, NY, June 2003. (Available at the U.S. Air Force Research Laboratory's Web site: http://www.rl.af.mil/tech/papers/ModSim/ModAb.html)

Pachepsky, Yakov, Martinus Th. van Genuchten, Ralph Cady, and Thomas J. Nicholson, "Letter Report: Task 1—Identification and Review of Model-Abstraction Techniques," U.S. Department of Agriculture, Agricultural Research Service, Beltsville, Maryland, February 27, 2003.

USGS, Amargosa Desert Research Site Web site: http://nv.usgs.gov/adrs/, 2003.