NUREG/CR-7221

Integrating Model Abstraction into Subsurface Monitoring Strategies

Office of Nuclear Regulatory Research

AVAILABILITY OF REFERENCE MATERIALS
IN NRC PUBLICATIONS

SR-CR 08/2016

Integrating Model Abstraction into Subsurface Monitoring Strategies

Prepared by:
Y. Pachepsky[1], A. Guber[1], A. Yakirevich[2],
F. Pan[1], T. Gish[1], M. Kouznetsov[2],
M. Van Genuchten[3], R. Cady[4], T. Nicholson[4]

[1]United States Department of Agriculture
Agricultural Research Service
Environmental Microbial and Food Safety Laboratory
Hydrology and Remote Sensing Laboratory
Beltsville, MD 20705

[2]Department of Environmental Hydrology & Microbiology
Zuckerberg Institute for Water Research
Blaustein Institutes for Desert Research
Ben-Gurion University of the Negev
Sede Boqer Campus, 84990, Israel

[3]Department of Mechanical Engineering, COPPE/LTTC
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ CEP 21945-970, Brazil

[4]Division of Risk Analysis
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

T.J. Nicholson, NRC Project Manager

NRC Job Code N6730

Office of Nuclear Regulatory Research

ABSTRACT

NUREG/CR-7221 reports on integrating modeling abstraction techniques into subsurface monitoring strategies. This research is the culmination of many field and modeling studies conducted by the USDA/Agricultural Research Service (ARS) at their Beltsville Area Research Center. The research design was to identify and examine near-surface water flow pathways by monitoring performance indicators within the unsaturated zone and local water-table system. The peak tracer concentration and the time to peak concentration at several monitoring locations served as the performance indicators. The objective was to apply model abstraction techniques in designing monitoring networks such as those used at nuclear waste and decommissioned facilities. The level of spatial and temporal detail in characterizing soil properties (e.g. water contents and hydraulic conductivities) is based upon the model abstraction considerations. Simplifications may omit significant processes and conditions that control the water and contaminant migration. The ARS field studies provided detailed databases for modeling water and chemical tracer movement in 2- and 3- dimensions to facilitate understanding of what processes and properties could be simplified (abstracted). Model abstractions included using pedotransfer functions for hydraulic conductivity, soil profile homogenization, and unsaturated zone omission. This latter abstraction proved to be the most accurate in generating a monitoring network that reflected the calibrated model. A comprehensive sensitivity analysis was performed to identify possible directions of model simplification in the model abstraction process. The integration of model abstraction into monitoring strategies based upon the ARS field and modeling findings was documented and reviewed by international soil scientists. The studies were jointly funded by NRC and USDA/ARS.

FOREWORD

This technical report was prepared by the USDA's Agricultural Research Service (ARS), under a follow-on to Interagency Agreement 05-005 (RES-05-005, JCN N6730) with the Office of Nuclear Regulatory Research. The objective was to characterize how groundwater model abstractions (simplifications) within conceptual site models impact subsurface monitoring strategies. These monitoring strategies are used to confirm the performance of nuclear waste and decommissioned facilities. Numerical models which assess environmental impacts rely on both the conceptual site model simplifications (abstractions) and the monitoring data. The monitoring networks should incorporate the significant processes and conditions of the model abstraction, and be compatible with the performance indicators of the numerical models.

The findings in this report represent the culmination of efforts presented in NUREG/CR-6884 and NUREG/CR-7026. These reports were previously developed under Interagency Agreement 02-008 (RES-02-008), "Model Abstraction Techniques for Soil Water Flow and Transport." The research presented in this document builds upon the previous work by characterizing the accuracy of ground-water monitoring networks derived from abstracted subsurface models.

NRC licensees and NRC licensing staff routinely develop and assess numerical models of subsurface transport based upon conceptual site models. Simplifications in these models are due to uncertainties in identifying the significant flow and transport processes, parameter estimations and soil structure as reflected in the model framework. Examples of these simplifications include: soil profile homogenization, unsaturated zone omission, and pedotransfer function usage to estimate hydraulic conductivities. This report directly addresses how such simplifications impact ground-water monitoring networks. NRC staff can use the information presented here to review licensees' monitoring and modeling programs, and assess the use of more compact, efficient ground-water models.

This report is not a substitute for NRC regulations, and compliance is not required. Consequently, the approaches and methods described in this report are provided for information only, and publication of this report does not necessarily constitute NRC approval or agreement with the information contained herein. Similarly, the use of product or trade names in this report is intended for identification purposes only and does not constitute endorsement by either the NRC or USDA/ARS.

Michael F. Weber
Director of Nuclear Regulatory Research

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

This study was designed and performed to investigate the opportunities and benefits of integrating model abstraction techniques into subsurface monitoring strategies. The study focused on future applications of modeling in contingency planning and management of potential and actual contaminant release sites within the scope of US NRC operations. The main objective was to develop methods for incorporating model abstraction techniques into the design of subsurface hydrologic monitoring and performance assessment programs. This study is based on a systematic model abstraction methodology that was developed and tested in previous work on flow and transport in soils and shallow groundwater systems.

A comprehensive review of groundwater monitoring network (GMN) design techniques was conducted first. The review demonstrated the recent increased use of physically-based pollutant fate and transport models in monitoring network designs. Since GMN designs are based on a conceptual model of the presumed subsurface flow and transport conditions, and since abstraction of the model structure leads to a range of conceptual models and their mathematical counterparts, an opportunity arises to design GMNs that decrease uncertainty in the predictions of individual abstracted models, or in weighted predictions from several models combined via model averaging. Another opportunity is to use model abstraction via parameter estimation to augment existing monitoring networks for discriminating between different conceptual and mathematical abstractions. These opportunities are pursued in this work and applied to models that are obtained via model abstraction for full three-dimensional (3D) flow and transport in variably saturated flow domains.

The USDA-ARS OPE3 experimental field site near Beltsville, MD, has been studied extensively for more than 10 years using geophysical, biophysical, remote sensing, and soil and groundwater monitoring methods. Available data were analyzed using systematic procedures within a broad vadose zone modeling context developed in our previous work. A major focus was on subsurface structural units and features that could drastically change the fate and transport of contaminants in the vadose zone, as well as on projected trajectories of the contaminant plume in groundwater.

Solute transport in soils and shallow groundwater at the site could potentially be affected by such features as the presence of a restrictive fine-textured layer that is not fully continuous laterally; the complex topography of the restrictive layer favoring preferential flow and transport along preferred pathways along its surface relief; the presence of natural capillary barriers, possible funnel flow in a coarse-textural layer between the more fine-textured layers, and localized high-conductivity parts of the soil pore space.

The abstraction methodologies were applied to data from the originally designed field site showing lateral transport of a surface-applied conservative tracer pulse, with transport being controlled by regular irrigation pulses and natural precipitation. The vadose zone was monitored for soil water contents and pressure heads, while groundwater levels and tracer concentrations in groundwater at three depths were also recorded. Fine-scale ground-penetration radar and time-lapsed electrical resistivity surveys, along with the borehole logs, contributed to the development of the original conceptual model in which subsurface structures exerted strong controls on flow and transport. A cluster analysis of borehole data showed the presence of nine soil materials in the studied subsurface domain.

A comprehensive sensitivity analysis was performed to identify possible directions of model simplification in the model abstraction process. Two sensitivity indexes, a first-order sensitivity coefficient and a derivative-based measure of sensitivity (section 5.1), were used to reflect the effect of parameter variations on variations in simulated values. These indexes were first computed as daily values, with their time series providing insight into the spatio-temporal sensitivity of solute transport in the variably saturated heterogeneous domain. The sensitivity of the base HYDRUS-3D model to the transport parameters and soil materials depended upon both location and time. The HYDRUS-3D code was found to be more sensitive to the transport parameters than modified TOUGH2. Time series of first-order sensitivity coefficient were aggregated to obtain an average and standard deviation of daily value over the observation period for each of eight soil materials. Ranking soil materials according to the aggregated sensitivity indexes allowed identification of those soil materials most appropriate for model simplification by material elimination.

The proposed approach to abstract a model by decreasing the number of soil materials consisted of replacing the eliminated soil material with one of the closest materials according to the cluster analysis used for the material definition. The simulations were run with the three closest soil materials as possible replacement for the eliminated material. Different replacement materials were selected when the HYDRUS-3D or modified TOUGH2 codes were used, possibly because of differences in averaging hydraulic conductivity procedures.

Four examples of model abstraction were developed using the manually calibrated HYDRUS-3D code as a base model. The examples included using pedotransfer functions for the hydraulic conductivity, profile aggregation, ignoring the unsaturated zone, and combining the use of pedotransfer functions with scaling. The efficiency of a particular abstraction option was evaluated by comparing simulation results of the calibrated and abstracted models using three performance indicators, namely simulated peak concentration, time to reach the peak concentration, and total mass flow through observation nodes and transects at different distances from the release points based on simulations of 10 year-long tracer flow and transport.

The simulation domain that mimicked an NRC site was set within the boundaries of the GPR survey, and included the tracer application plot and 56 groundwater observation wells. Forty eight scenarios were simulated, each one being different in terms of the location and depth of the tracer release, groundwater dynamics, and precipitation. Pedotransfer estimates of soil water retention and the saturated hydraulic conductivity were generated using the Rosetta software and the tabulated summary of the Rawls databases, respectively. Estimates were found for each of 20,000 simulation nodes.

Abstraction of parameter estimation using empirical pedotransfer functions was found not to be efficient in our study due to the considerable sensitivity of HYDRUS-3D modeling results on the saturated hydraulic conductivity. Pedotransfer predictions of the saturated hydraulic conductivity were relatively poor as compared to calibrated values. Therefore, pedotransfer-based simulations produced results that were quite different from the calibrated model results. Time to peak concentration and total mass were better discriminators between the calibrated and abstracted models than the peak concentration itself.

The scale mismatch between measurements reflected in the pedotransfer database and the computational grid discretization was a plausible explanation of the low efficiency of the pedotransfer abstraction process. Pedotransfer abstraction for this reason was complemented with scaling abstraction. The increase in the saturated hydraulic conductivity with the size of grid cells was simulated assuming a power-law dependence of hydraulic conductivity on the ratio of

characteristic lengths (section 5.2.8). The accuracy of simulations with the concurrent use of pedotransfer and scaling abstraction was very close to that of the calibrated model, thus confirming the efficiency of systematic model abstraction.

Profile aggregation, i.e. assuming a single soil material for all depths was another model abstraction technique implemented in this study. A motivation for this abstraction technique is possible dominance of one soil material in the simulation domain. Profile aggregation had a major effect on tracer transport in our study. Introducing a homogeneous profile resulted in drastic changes in the spatial pattern of the performance indicators. The model that used profile aggregation performed the poorest.

Ignoring the unsaturated zone implies ignoring the retardation of water and chemical fluxes due to the nonlinear decrease in hydraulic conductivity with a decrease of water content in the absence of preferential flow and transport in the vadose zone. However, ignoring the unsaturated zone did not lead to noticeable changes of the performance indicators. A systematic difference between abstracted and calibrated modeling results was found for the times to peak concentration, but not for the maximum concentration and the total mass transported. The abstraction scheme which ignored the unsaturated zone was acceptable if the key performance indicators were the peak concentration and the total contaminant flux which passed through a transect or observation well. This method of model simplification was found to be acceptable because of a relatively shallow vadose zone at our site.

It should be used with caution at sites with deep phreatic surfaces where characteristic times of vertical transport in unsaturated zone and lateral transport in groundwater are comparable. The procedure of this work needs to be re-applied to make sure that the travel time in the unsaturated zone can be ignored which may be the case in semi-arid zone.

Mathematical models developed with model structure abstractions should have different underlying conceptual hypotheses. Since a monitoring network is always designed based on a conceptual model, designs may need to be different for models that are abstracted differently. A monitoring network design could either jointly use the designs resulting from several conceptual models, or develop a particular design to discriminate between models with the goal of selecting the more plausible one.

Two approaches were developed and implemented for situations in which monitoring network augmentation is envisaged for flow and transport in a variably saturated three-dimensional subsurface domain.

A sensitivity-based statistical inference method has been proposed in steady-state groundwater studies to rank candidate monitoring locations by their potential to improve model predictions and the accuracy of parameter estimation. The method was expanded in this study for applications in variably-saturated 3D flow and transport domains by redefining parameter sensitivities for non-stationary conditions. The modified statistical inference method was tested first using simulated data of the tracer experiment as obtained with the calibrated model so as to rank the existing observation wells in terms of their importance for model calibration. The method provided stable results indicating that (a) one (deeper) observation depth rather than two would be sufficient in the calibration, and (b) that the relative importance of wells is defined by either their presence in subdomains where fast changes in flow and transport occur within small distances, or their presence in relatively large subdomains where flow is slow compared to other parts of the OPE3 simulated field site. The sensitivity based statistical inference method was further used to augment the observation network with the aim of improving the HYDRUS-3D calibration.

Observations at the lower of the two depths appeared to be more important, and followed the previously noted trend of importance of observations in subdomains where fast changes in flow and transport occur within small distances.

Performance indicators were used to compare monitoring locations selected for the base model and for models abstracted using pedotransfer, profile aggregation, and ignoring the vadose zone by placing the tracer into groundwater at the very beginning of the simulations. Locations selected for the pedotransfer abstraction were close to those for base model if the peak concentration was used as the performance indicator. Soil profile aggregation was the only abstraction technique that generated a monitoring network dissimilar to the network obtained using the calibrated HYDRUS model. This occurred primarily due to existence of preferential lateral flow paths in the heterogeneous simulation domain, which were entirely discarded in the homogeneous profile abstraction process. Soil profile aggregation abstraction in general may well be suitable for small-scale applications or when vertical soil stratification is not important hydrologically. However, for large-scale projects natural succession of soil genetic horizons producing different hydrological regimes at different depths cannot be ignored when water flow and chemical transport are modeled.

Abstraction by ignoring the unsaturated zone appeared to be the most accurate in predicting the tracer concentrations and fluxes, and consequently generated essentially the same monitoring network as the calibrated model. This occurred because the unsaturated zone was very thin in our study, while vertical transport dominated at the release points. These hydrological conditions created a specific flow regime in which the tracer was transported preferentially in the vertical direction before approaching the groundwater. The results of this abstraction can differ for a deeper unsaturated layer, especially when the soil profile contains horizontal lenses with contrasting hydraulic properties. The presence of such structural factors, and their effects, need to be evaluated at the conceptual model development stage. Model abstraction then can help to quantify their possible role at the mathematical modeling stage of a project.

Model discrimination can be one of several objectives of groundwater monitoring design. A simplistic search for a location where predictions using two models differ the most can give unsatisfactory results since both data and model predictions are uncertain. The novel method developed in this work method determines new monitoring locations where the total information gain is maximized. The latter is computed based on estimates of the uncertainty in modeling results and uncertainty in observations. The discrimination was sought for the original model and the model obtained with the soil profile aggregation abstraction.

 Pedotransfer functions were used to develop the ensemble of models for estimating the uncertainty in modeling results obtained with the numerical 3D flow and transport model. Peak tracer breakthrough concentrations were used to define the information gains. The determination of the new locations to augment existing ones was conducted on a 2-D grid. The information gain peaked in small area, and additional observation locations were very well spatially defined.

Overall, integrating model abstraction and monitoring strategies appears to be a logical step given that both the mathematical models and the groundwater monitoring network designs are based on conceptual site models of subsurface structural and geochemical conditions. Developments in the fields of monitoring data assimilation and geophysical data fusion will further help to interfacing of the two technologies. Since the use of mathematical fate and transport models in groundwater monitoring designs is expanding, one should expect further intertwining of modeling and monitoring based on common conceptual modeling.

ACKNOWLEDGMENTS

ABBREVIATIONS AND ACRONYMS

1D, 2D, 3D	One-dimensional, Two-dimensional, Three-dimensional
ANN	Artificial neural networks
BTC	Breakthrough curve
C	Clay
CV	Coefficient of variation
DGPS	Differential global positioning system
ET	Evapotranspiration
FAO	Food and Agriculture Organization of the United Nations
FTM	Model of the subsurface flow and transport
EM	Electromagnetic induction
FD	Finite difference
GLUE-BMA	Combined Generalized Likelihood Uncertainty Estimation (GLUE) and Bayesian Model Averaging (BMA) methods
GMN	Groundwater monitoring network
GMS	Groundwater Modeling System
GPR	Ground penetrating radar
GWD	Groundwater depth
L	Loam
LS	Loamy sand
LSA	Local sensitivity analysis
HYDRUS	Software package for simulating the movement of water, heat and multiple solutes in variably-saturated media
MA	Model abstraction
MATLAB	A programming environment for algorithm development, data analysis, visualization, and numerical computation
MCP	Multi-sensor capacitance probe
MODFLOW	Three-dimensional finite-difference ground-water flow model
MT3D	Modular 3-D Solute Transport model
MWBUS	Model of Water Budget of Unsaturated Soil
NRC	U. S. Nuclear Regulatory Commission
OPR	Observation-prediction statistic
PTF	Pedotransfer function
RADAN	GSSI's state-of-the-art post-processing software
REV	Representative elementary volume
ROSETTA	Windows based program that implements artificial neural network to soil hydraulic parameters
S	Sand
SA	Sensitivity analysis
SCL	Sandy clay loam
SDSi	Standard deviation of the sensitivity index
SiCL	Silt clay loam
SiL	Silt loam
SL	Sandy loam
TDR	Time Domain Reflectometry
TOUGH	Numerical model for simulating the coupled transport of water, vapor, non-condensable gas, and heat in porous and fractured media

Roman symbols

$\underline{\mathbf{A}}$	Sensitivity matrix
b	Exponent
C	Solute concentration
c	speed of light
$Cpeak$	Peak concentrations
D	Index of agreement
$\mathbf{D}$	Hydrodynamic dispersion tensors
D_{ij}	Components of the hydromechanical dispersion tensor
Di	Derivative-based sensitivity index
E	Efficiency index
E	Soil dielectric constant
ET	Daily average evapotranspiration
$\overline{ET}$	Evapotranspiration over simulation period
GWD	Groundwater depth
H	Matric pressure head
H	Hydraulic head
h_{cf}	Pressure head at the top of capillary fringe
I	Kullback-Leiber divergence
J	Total information gain
Kv	Total information gain with prior probabilities of models being correct
K_s	Saturated hydraulic conductivity
$K_{S,R}$	Saturated Hydraulic conductivity estimated from Rawls' pedotransfer function
$\mathbf{K}$	Hydraulic conductivity tensor
l	Pore connectivity parameter
L_g	Size of the soil cores used to determine $K_{S,R}$
L_s	Characteristic length of the numerical grid used to solve flow and transport problem.
M	A parameter of van Genuchten's water retention model
N	A parameter of van Genuchten's water retention model
O	Measured values
$P_r^{(n)}$	Posterior probability
P	Predicted values
P	Probability
$\mathbf{Q}$	Darcy flux vector
q_r	Darcy flux vector components
Q	Water flux soil hydraulic parameters
QC	Chemical mass flux
R	Retardation factor
r^2	Coefficient of determination
RMSE	Root Mean Square Error
S	Sink/source term
S_e	Relative saturation
Si	Sensitivity measure
$S_{Y'}$	Prediction standard deviation

Slope X	Phreatic surface slope in X direction
Slope Y	Phreatic surface slope in Y direction
T	Time
Tpeak	Time of peak concentrations
V	Variance
v_i	Components of flow velocity
V	Groundwater velocity
V_m	Volume of geological media
v	GPR wave velocity
$\underline{\mathbf{W}}$	Matrix of weights of observations
X_i	Generic factor
$\overline{X}_i$	Nominal value of factor X
x,y,z	Coordinates
Y	Model output
$\overline{Y}_i$	Nominal value of factor Y
Z	Depth

Greek Symbols

α	A parameter of van Genuchten's water retention model
a_L	Longitudinal dispersivity
a_T	Transverse dispersivity
δ	Dirac delta function
Δ	Increment
ε	Accuracy parameter
$\mathcal{E}$	Expected value
σ_r^2	Variance
σ_Y^2	Variance of the distribution of observations about the expected value
λ	First order decay rate
Λ	Finite difference operator
μ	Specific water capacity
Θ	Volumetric water content
θ_r	Residual water content
θ_s	Water content at saturation
τ	GPR signal travel time

1 INTRODUCTION

1.1 Background

U.S. NRC staff reviews of performance assessments of nuclear facilities (e.g., decommissioning of the facilities, management of low-level and high-level radioactive waste disposal sites) frequently involve assessing models for subsurface water flow and solute transport in the vicinity of a nuclear facility (U.S. NRC, 1993, 1994). These models seek to represent complex and highly transient subsurface systems. Representations of those complex systems in existing models range from very simple to extremely sophisticated formulations (Neuman et al., 2003; Reilly and Harbaugh, 2004; Hill and Tiedeman, 2007; Valocchi, 2012). One reason for this is the inherent complexity of the subsurface, which needs to be dramatically simplified to be expressed in mathematical terms. Different sets of simplifying assumptions can be used that lead to different models (Neuman et al., 2003). And even when the types and relative positions of subsurface structural units are selected, knowledge about the locations of the boundaries of these units, and about the flow and transport properties of the soils and fractured rocks making up those units is uncertain. This creates substantial uncertainty with regard to the assignment of parameter values in the models.

The existence of several candidate models for the same observation site or for the same phenomenon is common in the environmental sciences and their applications, as well as in other disciplines. Making use of several models has been shown to be more beneficial than looking for the "single best" model. The use of simpler modes to complement a complex model is a promising approach (Bigelow and Davies, 2003; Van Ness and Scheffer, 2005). Simple models have advantages regarding the collection of data, the computations involved, the interpretation of the simulation results, and conveying the simulation approach to both technical and lay audiences. On the other hand, the explicit description of mechanisms in more complex models may cause their better performance outside of conditions for which the simple models were initially developed or tested.

Currently, different research fields have been adopting various approaches for selecting or deriving simple models for use along with more complex models (e.g., Casteletti, 2011; Lawrie and Hearne, 2007; Raick et al., 2006). One of these methods consists of a systematic derivation of simpler models from the original complex model. This method has been termed "model abstraction". The systematic model simplification in subsurface hydrology was first suggested by Neuman and colleagues (2003) in their work performed for the U.S. Nuclear Regulatory Commission. Examples of useful simplifications of ground water models were later presented by Hill (2006). For soils, the model abstraction methodology, including those techniques and examples used in simulating water flow, was developed in a recent interagency study by Pachepsky, et al. (2006).

One essential application of subsurface modeling is the development of monitoring programs and strategies. Monitoring of subsurface processes has been expanded significantly in the past several decades to address soil, groundwater, and surface water contamination. A wide variety of monitoring design techniques has been proposed (Minsker, 2003; Quevauviller et al., 2009). Yet, the central conundrum of subsurface monitoring design--the struggle between subsurface complexity and available monitoring resources still presents a challenge at decommissioning and new reactor sites--as well as at other potential point source contamination sites. Ideally, subsurface environments at potential or actual contaminated sites should be described using state-of-the-art conceptual models in conjunction with extensive monitoring data collected at the

site. While potentially very realistic, these models generally introduce large data requirements and often focus on parameters that have only minimal influence on the specific subsurface processes of concern to anticipated contaminant transport. Limited costs and resources, however, tend to make this approach an unlikely option. Instead, there is a tendency toward modeling complex and highly transient subsurface flow and transport processes with greatly simplified models that overlook fine-scale heterogeneities that may well control contaminant flow and transport at coarser scales (Neuman, et al., 2003).

Model abstraction techniques can be utilized to reduce the complexity of a natural system to its essential components and processes through a series of conceptualizations, selection of processes, and identification of parameters and boundary conditions. The result of successful model abstraction is to simplify the representation of a complex, natural groundwater system to a manageable level while still capturing an acceptable level of detail and realism (Pachepsky et al., 2006). The central hypothesis of this study is that models of intermediate complexity as developed using model abstraction can be useful for the design of feasible monitoring networks and schedules at actual and potential contaminated sites. The selection of a specific model should depend on the purpose of monitoring, such as obtaining calibration data, demonstration of compliance with a regulatory limit, or for obtaining information to remediate a plume.

1.2   Objectives

The objective of this research was to develop methods for incorporating model abstraction techniques into the design of subsurface hydrologic monitoring and performance assessment programs. The enhanced abstraction strategies and models needed to be tested using a site-specific database on flow and transport of tracers from pre-defined sources.

1.3   Approach

The research effort utilized a unique, long-term monitoring data set from the ongoing environmental monitoring program at the ARS OPE3 site, along with datasets collected at that same site during subsequent tracer experiments (Pachepsky et al., 2011). Both the vadose zone and groundwater were sampled continuously, which should present an advantage given the importance of the vadose zone in evaluating groundwater recharge and contaminant transport to groundwater (Hunt et al., 2008; Harter and Hopmans, 2004.). As appropriate for the shallow groundwater site, the base conceptual model to be simplified was selected to be fully three-dimensional. The highly nonlinear nature of flow in the unsaturated zone makes the modeling results dependent upon mathematical approximations needed for numerical solutions using a realistic temporal and spatial discretization. To evaluate this issue, two available codes (HYDRUS-3D and TOUGH 2) were used concurrently in this study.

The following requirements were considered to be important:
    1. The model abstraction techniques should be selected to support the performance assessment modeling and monitoring at the field/watershed scale;
    2. The research should recognize a variety of data aggregation methods used in performance assessment; data abstraction should be considered along with model abstraction;
    3. The effect and efficiency of model abstraction must be characterized in statistical terms; and
    4. Monitoring should improve both the performance of the abstracted models and discrimination between the abstracted models.

The following research questions were addressed to meet the above requirements:

1. Do the analyses and evaluations produce a sufficient technical basis to designate the appropriateness of the abstractions as represented by the different models, depending upon the "complexity" of the results (all data, filtered data or derivative data) used for the comparisons?
2. When does a simplification create significant error and/or uncertainty, and what monitoring data support that determination?
3. For the various abstracted models using filtered data or derivatives of the data, what test methods could be used to predict where (or when) to sample for a) an optimal reduction of uncertainty in the model; and b) discriminating indicators (or criteria) between models of different abstraction?
4. How can monitoring assist in determining which abstracted model and its assumptions are appropriate for estimating transport given the release scenario and site-specific subsurface flow conditions (e.g., presence of preferential pathways, perching units, fracture connectivity)?
5. Given the goals stated in the voluntary EPRI/NEI Industry Ground-Water Protection Initiative (Kim, 2009), what monitoring densities (i.e., areal, depth and time determinants) for that magnitude of release at the OPE3 field site would be needed to assure a high (or absolute) probability that the release would be detected?
6. How much uncertainty can be reduced by monitoring the unsaturated zone close to the release location, as opposed to monitoring at the water table, and which performance indicators could best be used and monitored in the unsaturated zone?
7. How could model abstraction techniques be coupled with site attributes and conditions to select the proper monitoring strategy to provide confidence in the selection of an appropriate remediation technique or techniques?

Monitoring network design issues includes (a) creating an initial monitoring network and (b) augmenting an existing network. The modeling-based initial monitoring design at the OPE3 site was described earlier (Pachepsky et al., 2011). Here the ARS study team concentrated on augmentation of an existing monitoring network, and two reasons for this approach are considered:
- improving model performance by augmenting the network to include optimal locations to decrease uncertainty in the predictions, and
- improving model discrimination by augmenting the network to include the best locations for distinguishing between two models.

2   MONITORING AND MODEL ABSTRACTION
IN SUBSURFACE HYDROLOGY

This section provides an overview of the state-of-the art of model abstraction and model-based monitoring design, and outlines the interactions between these two components of environmental assessments and predictions. Modeling and monitoring have substantial potential to complement each other to improve studies of any contamination problem. This is because the mathematical model and the monitoring network both require a conceptual model of the contaminated site, with both at the same time providing information about the validity of the invoked conceptual model. Monitoring and modeling both should also recognize the non-uniqueness of possible conceptual models, and thus should foresee their possible application to several models, rather than singling out only one model, even if thought to be the best.

Model abstraction is a methodology for reducing the complexity of a simulation model while maintaining the validity of the simulation. Model abstraction, if conducted in a systematic manner, generates a variety of both conceptual and mathematical models. Any one of the conceptual models could serve as a basis for the monitoring design. The monitoring design itself can be useful by allowing one to distinguish between the mathematical models and their conceptual counterparts, and by eliminating models that generate unacceptable results. On the other hand, an ensemble of simplified models can serve as a source of data about uncertainty in model predictions, which then can be incorporated implicitly or explicitly in the monitoring design process.

2.1 Model Abstraction in Subsurface Hydrology

2.1.1 Background

Model abstraction techniques stem from a need to improve the reliability of simulations, and to reduce their uncertainty, to make modeling and its results more explicable and transparent to end users, and to enable more efficient use of available resources in data collection and computations.

An important feature of model abstraction is the explicit treatment of model structure uncertainty. Model structure, along with data uncertainty and scenario uncertainty, is known to introduce uncertainty in the modeling results. Unlike uncertainty in input data, model parameters and scenarios, the effect of model structure uncertainty on uncertainty in the simulation results is usually impossible to quantify in statistical terms. However, using model abstraction, 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 ability to match observed data, and by its predictions with respect to scenarios that have not been observed.

Earlier collaborative work of ARS and NRC staff resulted in the definition of several categories of model abstraction techniques relevant to subsurface flow and transport modeling (Pachepsky et al., 2006). These categories included abstractions of model structure and model parameter determination, and were based on a comprehensive review of model simplification techniques developed for subsurface flow and transport (Fig. 2-1). A brief review of available model abstraction techniques is given below for clarity and reference purposes; more detailed descriptions are in the above cited NUREG report.
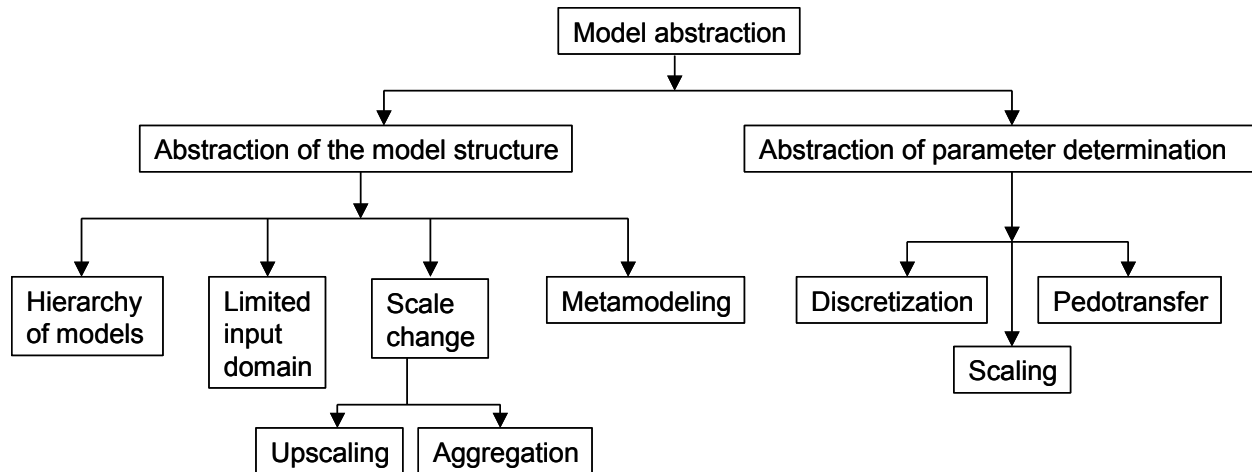
Figure 2-1   Categories of model abstraction techniques relevant to flow and transport modeling in subsurface hydrology.

2.1.2 Abstraction of Model Structure

Hierarchies of models. An example of the possible hierarchy of models available to simulate water flow in variably-saturated porous subsurface media is presented in Fig. 2-2. A pre-defined hierarchy of models has been suggested previously for flow and transport in structured media (Altman et al., 1996). Fig. 2-2 shows a schematic representation of increasingly complex models that may be used to simulate preferential flow and transport in macroporous soils or unsaturated fractured rock. The simplest water budget model describes the accumulation of water in the soil matrix and its discharge when the water content exceeds the field water capacity, or during evaporation periods (Fig. 2-2a). The classical approach of simulating flow/transport processes in a vadose zone devoid of macropores or fractures is to use the Richards equation for variably-saturated water flow and the advection-dispersion equation for solute transport (Fig. 2-2b). The simplest situation (Fig. 2-2c) for a fractured medium arises when the Richards and advection-dispersion equations are still used in an equivalent continuum approach, but with composite (double-hump type) hydraulic conductivity (permeability) curves of the type shown by Durner (1994) and Mohanty et al. (1997), rather than the classical smooth curve for the relative permeability shown in Fig. 2b.

More involved dual-porosity type models (Fig. 2-2d) result when the medium is partitioned into fracture and matrix pore regions, with water and/or solutes allowed to exchange between the two liquid regions (Ventrella et al., 2000; Šimůnek et al., 2003). Different formulations of this type are possible. For example, one could permit transient variably-saturated flow in the fractures only, while allowing water to exchange between the fracture and matrix domains. The latter situation leads to both advective and diffusive exchange of solutes between the fractures and the matrix, but still without vertical flow in the matrix (e.g., Zurmühl and Durner, 1996; Zurmühl, 1998). Dual-permeability models (Fig. 2-2e) arise when water flow occurs in both the fracture and the matrix domains. Examples are models by Pruess (1991), Gerke and van Genuchten (1993) and Jarvis (1999). These models use different formulations for the exchange of water between the fracture and matrix regions. In some models, (e.g., Wilson et al., 1998) more than two domains are considered, each one having its own hydraulic properties. The modeling approach can be refined further by considering transient flow and/or transport in well-defined discrete fractures without (Fig. 2-2f) or with (Fig. 2-2g) interactions between the

fractures and matrix. The latter approach is based on the assumption that the flow and transport equations of the fracture network can be solved in a fully coupled fashion with the corresponding equations for the matrix (e.g., Therrien and Sudicky, 1996).

Limited input domain. Model abstraction techniques based on delimiting the input domain rely on the notion that some feature or process may not be relevant for a given class of scenarios or for a given set of model outputs. A reduction in the spatial dimensionality is one application of this technique. Two-dimensional representations of the subsurface are sometimes redundant and 1D representation may suffice as shown by Wang et al. (2003). In another example, Guswa and Freyberg (2002) explored the possibility of using a 1D model to characterize solute spreading in a medium containing low-permeability lenses; they found that the 1D macroscopic advection-dispersion equation closely matched results of the 2D model when the equivalent conductivity of the domain was less than the geometric mean conductivity. This example shows that one should expect a change in parameter values when the dimensionality is reduced.

Scale change. Scale is a complex concept having multiple connotations. The notion of <u>support</u> is important to characterize and relate different scales in subsurface hydrology. Support is given by the length, area, or volume of a sample or element for which a single value of a porous medium
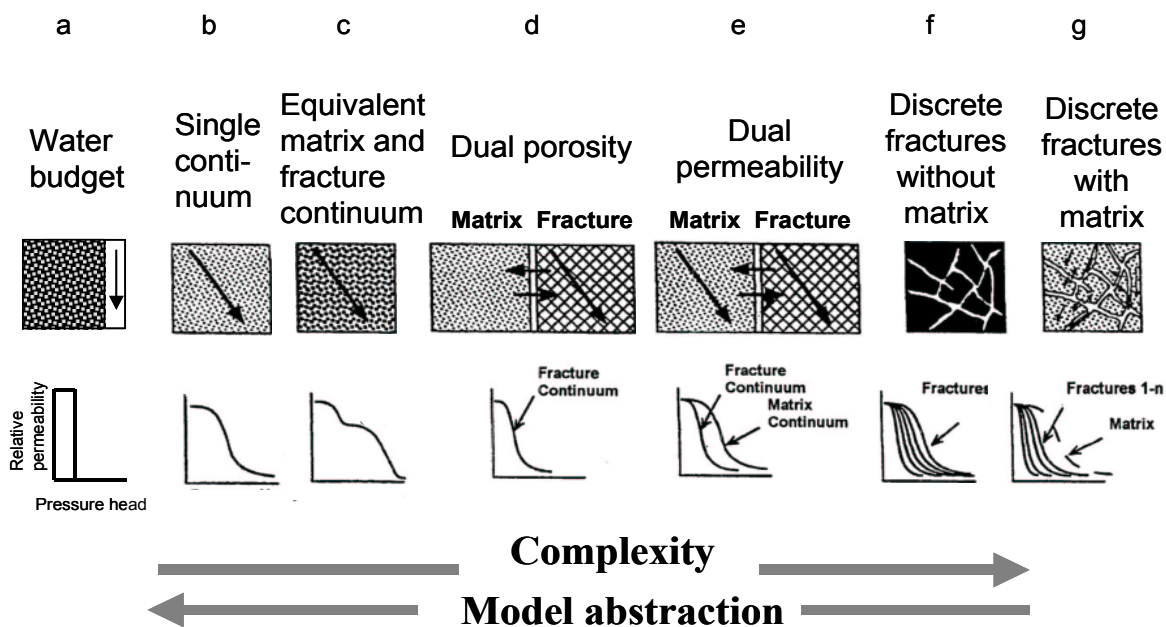
Figure 2-2   Hierarchy of models to simulate water flow and solute transport in structured soils or in unsaturated fractured rock (after Altman et al., 1996).

property is defined and no variations in this or other properties are taken into account. The size of an individual sample, and the size of a discrete spatial element in the flow model, are typical examples of support. The terms "<u>resolution,</u>" "<u>pixel size,</u>" "<u>grid size,</u>" and "<u>voxel size</u>" are often used to define the resolution in terms of length, area, or volume. An area or volume that is sampled with given support determines the <u>extent</u> of measurements. Distances between sampling locations define the <u>spacing</u>, which is also a function of scale. Blöschl and Sivapalan (1995) suggested characterizing scale not with a single value of support, extent, or spacing, but

with a position of a particular hydrological dataset in the 3D space of coordinates 'support-spacing-extent'. Therefore, change(s) in one or more of the numerical values of support, spacing, or extent leads to a change in scale.

In practice, vadose zone investigations generally define the support of core scale, soil profile or "pedon" scale, field scale, and watershed scale operationally, i.e. based on the measurement setup and available equipment. Increasing the linear size by about two orders of magnitude generally corresponds to a transition from one of these scales to the next one, as shown schematically in Fig. 2-3. Changes in the spatial scale are usually reflected by changes in the temporal scale as shown in this figure.

Scale change with upscaling. This category of model abstraction recognizes the need to alter the model structure when the spatial scale changes. Model equations, variables and parameters change as the scale changes. A key premise of upscaling is the possibility to derive parameters of a coarser-scale model from parameters of a finer-scale model.
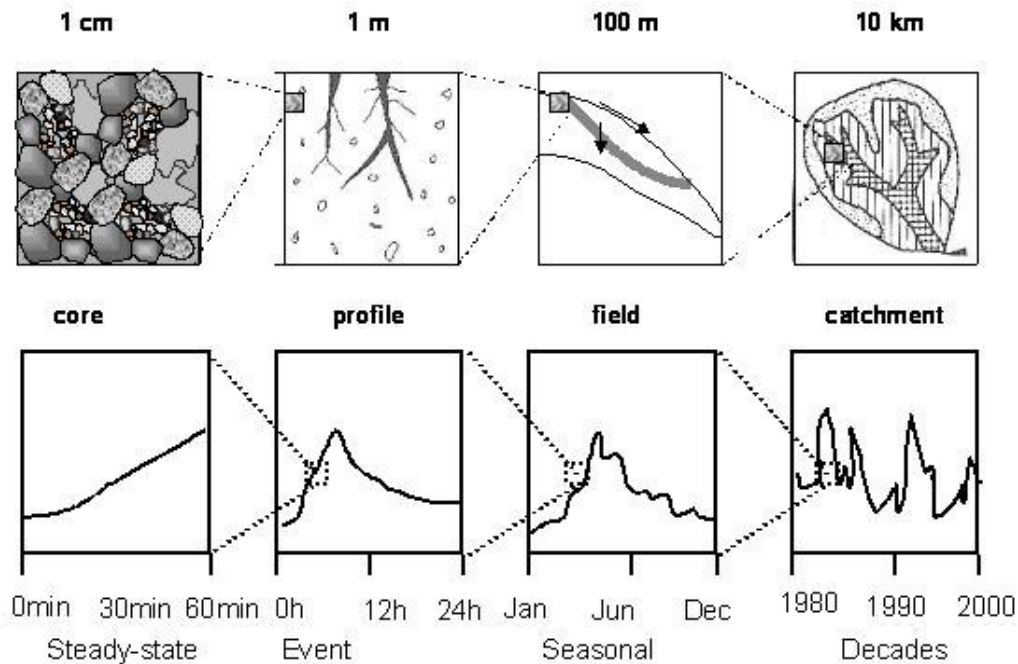
Figure 2-3   Spatial and temporal operational scales in hydrology (after Blöschl and Sivapalan, 1995).

To be effective, all upscaling techniques must use the correct statistical representations of relatively rare but essential features at the finer scale since those features often govern relevant

media properties or processes at the coarser scale. For example, if the hydraulic conductivity is distributed lognormally, then five percent of a fine-scale representative elementary volume

(REV) will conduct 95% of all flow, which implies that relatively rare features at the smaller scale may control flow at the coarser scale. Macropores (i.e. relatively large pores) provide another example of the importance of relatively rare fine-scale features. As compared to smaller matrix pores, macropores are rare in soils and are easy to miss during sampling. Yet, continuous macropores largely control rates of water flow and solute transport in a soil profile. Similarly, lateral preferential flow pathways in the subsurface are fine-scale features at the field or watershed scale, and relatively difficult to locate. Yet, lateral preferential flow pathways often control lateral flow and solute transport at the larger scale.

Scale change with aggregation. Aggregation also leads to a change in the governing equations and model parameters. However, unlike with upscaling, no relationship is assumed between the model parameters at the fine and coarse scales. Parameters of the coarser-scale model are deemed to be lumped, which suggests that field data are needed to calibrate the coarse scale models. An example is the use of a water budget soil flow model (Fig. 2-2a) at the field scale, while the Richards equation is used at the soil profile scale (Fig. 2-2b). Soil water retention is then parameterized in terms of field capacity at the coarse scale, whereas soil water retention curves are used at the fine scale. The value of water content at field capacity cannot be attributed to a single value of the soil water potential, and therefore no relationship exists between the coarse- and fine-scale soil water retention parameters.

Aggregation can be also implemented without a change in the model equations by combining several soil horizons or geologic strata. For flow and transport in the vadose zone, one common application is to replace a heterogeneous soil profile with an equivalent homogenous profile while retaining the Richards equation as the flow model (Zhu and Mohanty, 2002). The flow and transport parameters are still lumped in that case. Attempts to determine the effective hydraulic properties of the equivalent homogeneous soil profile from the layer properties have shown that the effective properties depend not only on the layer properties as such, but also on the type of predominant water regime (infiltration or evaporation).

Many examples of distributed watershed modeling with varying degrees of sub-watershed aggregation have recently been developed in surface hydrology. In general, excessive aggregation worsens model performance. For example, Boyle et al. (2001) showed that aggregation from eight to three sub-watersheds did not worsen model performance, whereas aggregating all sub-watersheds into only one watershed did.

Metamodeling. Metamodeling is a group of abstraction methods that uses results of multiple simulation runs to extract information helpful for simplifying a complex model. Metamodeling literally means modeling a model. Also known as the repro-modeling (Meisel and Collins, 1973) or response surface modeling, metamodeling creates a relatively simple empirical model intended to mimic the performance of a large complex model in order to reproduce the object model's input-output relationships (Davis and Bigelow, 2003). A common way to develop a metamodel is to generate "data" from a number of large-model runs and then to use statistical methods to relate model input to model output without attempting to understand the model's internal working.

Machine learning methods have become a popular means to build metamodels. Artificial neural networks (ANNs) are the most widespread among these since they are powerful approximators and, as such, are popularly used to relate multiple input variables to outputs from the complex model (Haykin, 2008). Examples of the application of artificial neural networks to mimic MODFLOW output for a range of scenarios for particular remediation sites were published as early as 10 years ago (Kron and Rosberg, 1998; Masket et al., 2000). The use of ANNs requires development of a large number of training datasets covering the range of possible scenarios of forcing variables. Generating such datasets requires extensive computing efforts, but after the ANN is trained, computations with the derive ANN are several orders of magnitude faster than simulations with the original model. An example involving regional wastewater planning in which, for reasons of computational efficiency, an artificial neural network was employed is given by Wang and Jamieson (2002). The ANN replicated the process-based model in multiple evaluations of the model output during optimization aimed at determining both the best sites and individual discharge standards. Other machine learning methods such as support vector machines, classification and regression trees, and genetic programming are increasingly being appreciated also in subsurface hydrology.

Machine learning methods can be used to simulate not only the model output of interest, but also the results of any part of the computations performed during a model run. For example, Hassan and Hamed (2001) demonstrated the use of ANNs to predict particle trajectories in a particle-tracking algorithm simulating plume migration in heterogeneous media. Their metamodeling approach substantially improved the computational efficiency of the entire algorithm.

2.1.3 Abstraction of Parameter Determination

Discretization. Using a set of discrete parameter values, instead of continuous functions, can sometimes lead to substantial simplifications. This simple but very important abstraction approach is commonly used in flow and transport modeling by employing soil and geological maps to define soil or sediment structural units that have the same hydraulic properties, even though the real-world hydraulic properties may change more gradually from one unit to another. Zonation of the hydraulic conductivity was found to be efficient when geophysical and hydrogeological data were combined (Chen et al., 2001, 2004; Straface et al., 2011). Automated zonation is now also feasible (Tsai et al, 2003), which has the advantage of not needing a specific equation to simulate the dependence of parameters on the spatial coordinates.

 Abstraction by temporal discretization is often used to create soil surface boundary conditions or discharge values. The temporal support of precipitation and evaporation or discharge is then coarsened to time intervals of months or years to speed up computations or to avoid generating synthetic weather patterns. Unfortunately, this type of abstraction can potentially also introduce substantial errors. For example, McLaren et al. (2000) studied the transport of ^{35}Cl in fractured tuff at Yucca Mountain and showed that changing the infiltration regime from a constant value of 5 mm y^{-1} to a pulsing regime with the same annual average value greatly enhanced the transport of the radioactive isotope. The effect of infiltration pulsing was model dependent; low rates of mass exchange between preferential flow zones and the matrix further enhanced transport to make the predictions more consistent with observations (Bandurraga and Bodvarsson, 1999). Relationships between flow properties and water contents are highly nonlinear in the vadose zone. Temporal coarsening for this reason may cause simulated plumes to move much slower and to become much more compact from what they would have been otherwise.

Pedotransfer functions. Parameters of models for flow in variably-saturated porous media are nonlinear functions of the pressure head or water content. The parameters in these hydraulic

functions are notoriously difficult to measure. Accurate measurement of water and solute fluxes in unsaturated soils remains a major research issue; no routine methods have been devised to-date. Therefore, defining parameters using inverse methods is often problematic. Substantial efforts have been made to estimate the hydraulic parameters from data available from soil surveys or borehole logs. Regression equations used for this purpose are often called pedotransfer functions (PTFs). These regressions relate soil water retention and hydraulic conductivity to more readily available soil parameters, such as contents of textural fractions, bulk density and organic carbon content. Extensive work has been carried to develop PTFs. The proceedings of two international conferences (van Genuchten et al., 1992, 1997), a book on this topic (Pachepsky and Rawls, 2004), and a recent review (Vereecken et al., 2010) provide a panoramic view of these fast-developing studies.

The performance of pedotransfer functions can be evaluated in terms of their accuracy, reliability, and utility. In broad terms, the accuracy of a PTF can be defined as the correspondence between measured and estimated data for the data set *from which a PTF has been developed*. The reliability of a PTF may be viewed as the correspondence between measured and estimated data for data set(s) *other than the one used to develop the PTF.* Finally, the utility of a PTF in model applications can be assessed in terms of the correspondence between certain *measured and simulated environmental variables*.

PTFs developed from regional databases have been shown to be more reliable in regions with similar soil and landscape histories. For example, PTFs developed from the country-wide US database by Rawls et al. (1982) appear to be more robust than PTFs developed from regional databases. However the accuracy of any PTF outside of its development dataset is essentially unknown, and using multiple PTFs instead of relying on a single PTF may well be preferred. Ye et al. (2004) suggested averaging spatial variability models for unsaturated fractured tuff in situations where standard information criteria provide an ambiguous ranking of the models such that it was not justified to select one model and discarding all others. Guber et al. (2006, 2008) likewise used an ensemble of pedotransfer functions to simulate water flow in variable saturated soils.

Recent developments in pedotransfer technologies have focused on the use of spatially dense physical information related to soil cover. Using topographic information is based on the hypothesis that some relationship may exist between the soil hydraulic properties and topographic variables (Pachepsky et al., 2001) since (a) the basic properties of soil are known to be related to landscape position, and (b) soil hydraulic properties are related to basic soil properties. Soil water retention was found to exhibit a strong dependence upon terrain attributes in the work by Rawls and Pachepsky (2002) who used U.S. soil survey data from the National Resources Conservation Service (NRCS). Terrain attributes were also used by Romano and Palladino (2002) to improve PTFs.

The use of hydrogeophysical, remote sensing, and crop yield data is based on a similar premise as using topography. The basic soils data should affect both the sensor readings and the soil hydraulic properties. Therefore, some relationship between the geophysical data and the hydraulic properties should be expected. A recent book on hydrogeophysics (Rubin and Hubbard, 2005) summarizes the rapid progress in this field. Cross-borehole resistivity measurements are now also showing promise in improving estimates of flow and transport properties of the vadose zone (Looms et al., 2008; Deiana et al, 2008). One recent example involved the use of airborne gamma radiometric sensing to estimate the clay content of surface soils and using a simple pedotransfer function to convert this information into a spatial representation of soil water retention parameters (Smettem et al., 2004). In another example, Timlin et al. (2003) related biophysical information from yield maps to the soil field capacity.

The use of dense auxiliary data in PTFs reflects an attempt to trade data quality for data quantity. Since dense coverage can be treated as an image, image analysis techniques may be used for segmentation and classification, as well as for delineating the structural units of soils. Data assimilation techniques may be suited also for combining soil survey and sensor information (McLaughlin, 1995). PTFs developed with auxiliary data are probably highly site-specific and, therefore, useful mostly only for sites for which they were developed. Nevertheless, the availability of sensor data can make such PTFs a viable component of supplying parameters of models for water flow in the vadose zone.

Scaling. Abstraction of scaling presumes the dependence of flow and transport parameters on the shape and dimensions of the domain over which these parameters are averaged. The saturated hydraulic conductivity and the dispersivity are most often viewed as scale dependent parameters, although parameters of geochemical kinetics are also shown to be subject to scale-dependency (Malmstrom et al., 2000; White and Brantley, 2003; Li et al., 2006). Evidence is accumulating that parameters of biological activity models show scale-dependencies, too.

Examples of observed scale dependencies in the saturated hydraulic conductivity are shown in Fig. 2-4. Ks values tend first to increase as the cross-sectional area or volume of the reference area increases and then presumably stabilize. The increase in Ks is caused by a change in the type of structural heterogeneities that are encountered as the scale increases. Large heterogeneities that are rarely viewed at the finer scales dominate flow as the scale increases. The dispersivity shows a similar behavior. Different measurement techniques (e.g., slug, single well, and multiple well pumping tests) represent different scales, and integration of their results presumes application of scaling as discussed by Neuman (1991).

The structure of surface soils in some cases can create the decrease in the value of hydraulic conductivity with the increase of the thickness of soil layers that is involved in flow (Lauren et al., 1988). There reason for that is the number of well-conducting macropores connecting top and bottom boundary of soil sample decreased as the length of soil column increased.

The scaling abstraction of parameter determination provides compatibility of the scale of measurement with the scale of model resolution. Neuman et al. (2003) suggested that this can be done by either rescaling the data to fit the scale of model resolution (which often entails averaging or upscaling over computational grid cells) or adapting model resolution to fit the scale of measurement (which often entails adapting the size of grid cells to the size of the data support).

2.1.4 Systematic Model Abstraction

Earlier collaborative work between ARS and NRC staff resulted in the development of a systematic and objective approach for model abstraction (MA) relevant to subsurface flow and transport modeling (Neuman et al., 2003; Pachepsky et al., 2006). The approach included (a) justifying the need for model abstraction, (b) reviewing the context of the modeling problem, (c) selecting applicable MA techniques, and (d) determining alternative MA directions and simplifying the base model in each direction.

The ARS study team emphasized that model abstraction is always site-specific. Pachepsky et al. (2011) provides a detailed overview of the model abstraction process. In this section the ARS study team provides a brief description.

2.1.4.1 Model Abstraction Steps

The MA process starts with an existing *base* model that can be calibrated and used in the simulations. *Key output* of the model is defined in this first step. The output provides the necessary and sufficient information required to make decisions about a certain issue of interest.

The model abstraction process includes the following steps:
> 1. Justify the need for the envisioned model abstraction
> 2. Review the context of the modeling problem
> 3. Design the model abstraction using selected techniques
> 4. Perform the model abstraction.

Model abstraction justification. Any model simplification requires calibration of the abstracted simpler model and its confirmation with multiple model runs. This is a separate modeling project that demands resources, and may be used later to justify the need for model abstraction. The base model may need abstraction for one or more of the following reasons:

- The base model includes a complex description of processes that cannot be observed well, but still needs to be calibrated; the calibrated parameter values of those processes may be very uncertain.
- The base model propagates uncertainty in the initial distributions, the parameters, and the invoked boundary conditions (forcings) in a manner that creates unacceptable uncertainty in the key output.
- The base model produces inexplicable results in terms of the key output.
- The base model requires an unacceptable amount of resources for the computations, data preprocessing, or data post-processing (e.g., the base model is not suited to be part of an operational modeling system that requires real-time data processing).
- The base model lacks transparency to make the model and its results explicable and believable to users of the key output.
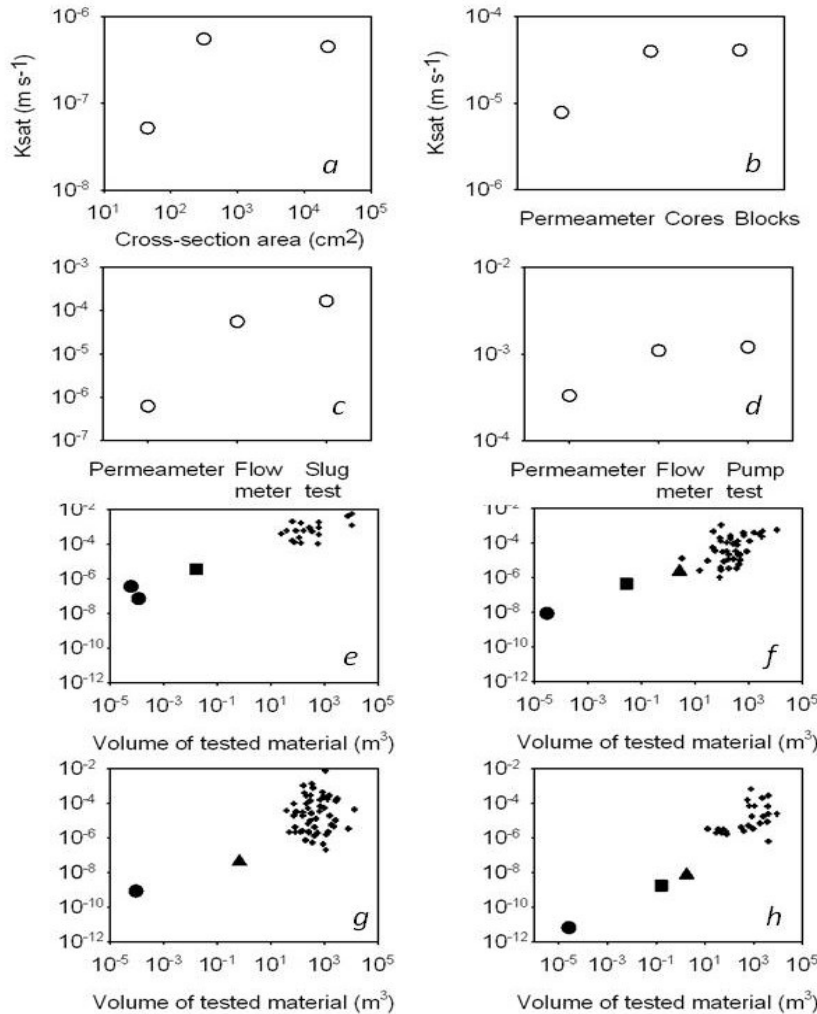
Figure 2-4   Relationships between the scale of measurement and the saturated hydraulic conductivity of soils and sediments; *a* – saprolite, North Carolina; measurements on cores, columns, and at the drain field scale (Vepraskas and Williams, 1995); *b* – Kokomo clay soil, Ohio; measurements with a permeameter, on cores, and on soil blocks (Zobeck et al., 1985); *c* – aquifer in a stream alluvium, Columbus, Ohio; measurements with a permeameter, flow meter, and using a slug test (Zlotnik et al., 2000); *d* – aquifer in glacial outwash, Cape Cod, MA (Zlotnik et al., 2000); *e, f, g, h* – aquifers in Wisconsin (Schulze-Makuch et al., 1999), *e* –glacial outwash sediments, *f* – carbonate aquifer, *g* – high magnesium limestone with shale, *h* – finely crystalline dolomite, ● - permeameter tests, ■ - piezometer tests, ▲ - packer tests, ✦ - single well pumping and specific capacity tests; oversized symbols show averages over a large number of replications.

The lack of transparency of the base model is often a result of abstraction motivated by the perception of potential users or critics of the simulation results. Incomprehensible simulation results with the base model or the unacceptable resource demand of the base model are often reasons that preclude timely completion of a modeling project.

The need for abstraction may stem also from uncertainty in the calibrated parameter values because of limited observation ability or uncertainty in the simulation results. The decision to carry out model abstraction may be based on the statistics of the parameter estimates and the key output, as discussed by Pachepsky et al. (2006).

Context of the modeling problem. The context of a modeling project has to be reviewed to assure the objectiveness and comprehensiveness of the selected model abstraction process. One must realize which details and features of the problem are to be omitted or de-emphasized when abstraction is performed. Neuman et al. (2003) list the following issues that constitute the context:

1. What is (are) the question(s) that the base and abstracted models are to address? The answer should consider (a) the existing or potential problem in which modeling is one of the solution instruments, (b) existing or potential causes of the problem, (c) issues requiring resolution, and (d) criteria to be used for deciding the level of resolution. The key output has to be provided with the spatial and temporal scale at which model abstraction is evaluated. Acceptable criteria for accuracy and uncertainty of the model output must to be established by the end users. In some cases mandatory regulations exist on performance measures that articulate the statistics to use in a particular case. If such regulations do not exist, then the statistics must be selected and defined; they should describe simple and clear ideas about the correspondence between the data and the simulations, such as how variability in the model errors compares with variability in the data, if the model residuals have trends, or if systematic relative or absolute errors exist in the predictions.
2. What type of data is available to calibrate the base and abstracted models and to test them with respect to the key output? An essential prerequisite is to have a database that is as broad as possible. The database must include data from public and private sources, cover both quantitative and qualitative (expert) information, and encompass both site-specific and generic information. A list of base model inputs and outputs would provide a convenient template for the necessary parts of the database. It is imperative to have statistics of all available model input and measurable model output data, including (a) the initial distributions of water contents and/or pressure heads, and the concentrations of solutes of interest, (b) surface and subsurface soil properties including horizon or layer thicknesses, porosities, bulk densities, sorption parameters, and in some cases redox conditions, (c) the forcings that provide the boundary conditions and the source/sink terms, and (d) the model parameters typical for the site. The latter can be inferred from an ensemble of pedotransfer functions, provided the necessary soil survey and/or borehole log data are available to serve as pedotransfer inputs. Statistic measures include the type of statistical distribution used, median values and variability estimates, information about observed outliers, and correlations between parameter values.
3. To make sure that the abstracted models are sound, additional information may have to be collected to ensure that the abstracted models include descriptions of all essential processes of flow and transport at a given site. This information may be of lower quality compared with the necessary part of the database. For example, again, one must be sure to not ignore certain small-scale internal heterogeneities that may have a dominant effect on flow and/or transport at the scale of interest.

2.1.4.2 Model Abstraction Design and Example

The design of a model abstraction project should include selection of the model abstraction techniques, definition of uncertainties that have to be simulated using Monte Carlo or other methods, and selection of the software to be used (Pachepsky et al., 2006). Decisions about which model abstraction technique to use depend on the reasons for the model abstraction. In general, model abstraction can lead to simplifications via:

- the number of processes being considered explicitly,
- adopting simpler process descriptions,
- coarsening the spatial or temporal support
- the number of measurements to be used for reliable parameter estimation,
- reduced computational burden, and
- simplified pre-processing and post-processing data.

An example of model abstraction design and implementation was developed in a study focusing on infiltration in a variably-saturated soil subject to natural rainfall (Pachepsky et al., 2006). The base model for that study was previously developed by Jacques et al. (2002) who simulated water flow in a layered soil using a single-continuum pore space model (media b in Fig. 2-2). Parameters were estimated by calibration using measured water content and pressure head time series. Key output was the soil water flux at three depths. The base model provided an excellent fit to the soil water contents but failed to simulate measured soil water fluxes.

The results initially could not be explained well (Jacques et al., 2002). Fig. 2-5 shows the selected model abstraction techniques, the uncertainty treatment, and the software selected (Pachepsky et al., 2006). Four abstraction techniques were used. Two techniques simplified the base model via (1) a simpler process description, i.e., by changing the porous media model from a single-continuum (Fig. 2-2b) to a water budget model (Fig. 2-2a), and (2) by aggregating soil layers into a single layer with effective hydraulic properties. Two other techniques simplified parameter estimation by using only laboratory data, or an ensemble of pedotransfer functions. Fifty Monte-Carlo simulations were performed, with each abstracted model using data on variability of the calibrated, measured or estimated hydraulic properties. The HYDRUS 2D and MWBUS software packages were used to simulate water flow in the single-continuum medium and water budget model, respectively.

The abstracted model correctly simulated fluxes and showed why inexplicable results had been obtained with the more complex model. Simultaneous calibration of a large number of parameters in the base model caused unrealistic flow simulations that included substantial runoff generation, whereas no runoff was observed during the experiments. The model abstraction process via pedotransfer functions at the same site showed that an ensemble of pedotransfer functions produced a satisfactory representation of field uncertainty in the soil hydraulic properties (Guber et al., 2006).

2.1.5 Model Abstraction vs. Arbitrary Selection of a "Simple" Model

Models used in performance assessments of waste disposal sites generally ignore small-scale heterogeneities in space and time. Preferential pathways are typically ignored when a single-continuum pore space representation is used. Observed strongly asymmetrical distributions of the pore water velocity are then often represented by an average velocity. Similarly, strongly asymmetrical temporal distributions of actual rainfall or discharge are represented by rates

averaged across long periods of time. Thus, the often-used models are actually abstractions of more realistic models that explicitly could account for spatial and temporal heterogeneities.

One argument for using a simplified model based on averaging is that locating and quantifying small-scale spatial and temporal heterogeneities has substantial conceptual and resource limitations. However, the magnitude of errors caused by ignoring the small-scale heterogeneities is never known *a priori*. Recent research shows that heterogeneities are common, and that even slight variations in soil texture may lead to flow barriers and associated lateral flow. These lateral flow paths may lead to solute mass losses, lowed vertical fluxes, and higher longitudinal dispersivities (Looms et al., 2008). This shows that it is imperative to evaluate possible prediction errors when the "simple" model is used. Deriving a simple model from a more complex base model provides opportunities for such an evaluation.
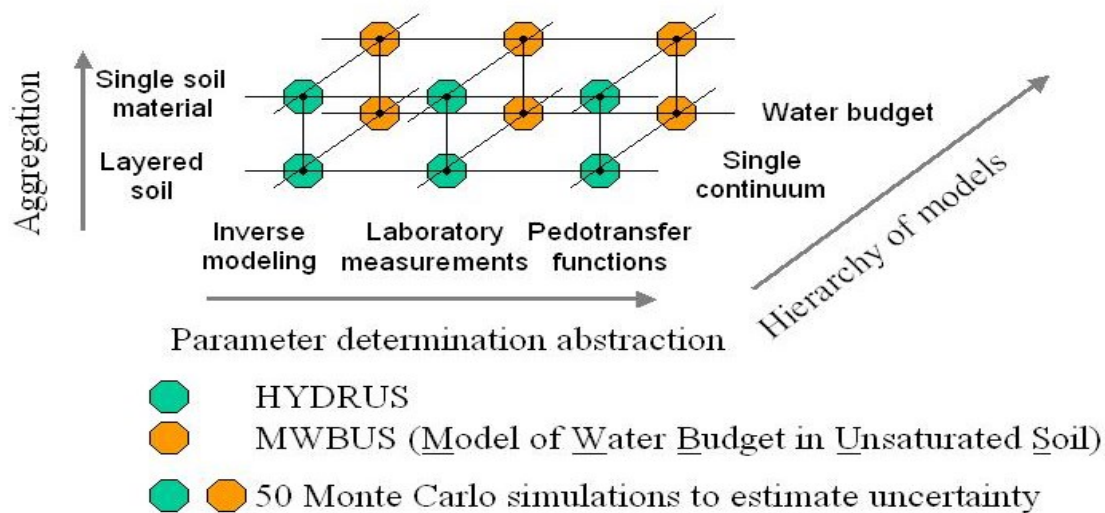
Figure 2-5   Design of model abstraction via model structure and parameter determination

2.2 Approaches to Modeling-Based Monitoring Optimization

2.2.1 Background

Significant efforts have been invested in the development of techniques to design groundwater monitoring networks (GMN), particularly for groundwater quality monitoring. A range of state-of-the art reviews and guidance documents have been published about GMNs during the last decade (Bloomfield, 2000; Hassan, 2003; Minsker, 2003; U.S. EPA, 2005; U.S. DOE, 2004; Kollat et al., 2011). It is generally agreed that there is no unique "best" way to conduct a long-term monitoring optimization project. Multiple guidance documents, tools, and standardized methods and approaches which utilize qualitative, temporal, and/or spatial-statistical methods have been applied successfully to a range of sites. The most significant advantage conferred by any optimization approach is the fact that they are used to apply consistent, well-documente procedures, which incorporate formal decision tools, to the process of evaluating and optimizing monitoring programs (U.S. EPA, 2005).

2.2.2 Components of a Groundwater Monitoring Network Design

The design of modern groundwater monitoring networks requires consideration of a range of factors that include having adequate spatial and temporal coverage of sampling sites, balancing potentially competing objectives within a monitoring program, considering the complex nature of geologic, hydrologic, and other environmental factors, accounting for the significant uncertainty about many parameters used in the design process, and considering the range of applicability of various methods in network design including their relative strengths and weaknesses (Bloomfield, 2000).

Implementation of a GMN always requires clearly stated objectives. Loaiciga et al. (1992) proposed a classification of objectives that forms the basis of Fig. 2-6. Specifically, the objective of ambient monitoring is to establish an understanding of characteristic regional groundwater trends with time. Detection monitoring is designed to identify the presence of targeted parameters, such as contaminants, as soon as they exceed background or established levels. Compliance monitoring is established to examine whether or not a set of specified groundwater-monitoring requirements, usually for chemical constituents, is met, for example near waste disposal facilities. Research monitoring on the other hand is a characteristically detailed spatial and temporal groundwater sampling effort designed to meet specific research goals. Objectives of the monitoring network define the monitoring scale.

Phase-dependent classification of monitoring objectives includes four general categories (U.S. EPA 2004): identification of changes in ambient conditions; detection of flow and transport of the environmental constituent of interest, demonstration of compliance with regulatory requirements; and demonstration of the effectiveness of a particular response activity or action.

The dynamic nature of a groundwater monitoring network is an important factor in the design of an effective groundwater monitoring network. Network design may be an iterative process in which initial sampling programs are often revised or updated as a result of previously collected data. Thus network augmentation or reduction is a characteristic feature of dynamic GMNs. In addition, the objectives of the monitoring network may also change with time.
A dynamic GMN design hence includes an iterative validation-monitoring-refinement cycle (Hassan, 2003).

GMNs must be developed according to a conceptual model of the site that includes geological and hydrogeological boundaries; the physical structure of the aquifer; information about recharge, discharge, groundwater flow pattern and mechanisms, rock-groundwater interactions, and the effects of unsaturated zone processes. Central to conceptual site model development is the collection, fusion and integration of appropriate geological data. Neuman et al. (2003) noted that it is often possible to postulate hydrogeologic conceptual models or hypotheses for a site on the basis of a broad range of publicly available geologic and geographic information about its surroundings.

Additional conceptualizations can be implemented on the basis of generic data about similar regions and the properties of similar materials elsewhere. Still, it is important to realize that each site is unique and very likely to reveal additional properties and features when characterized in some detail locally. Local characterization is hence essential for postulating acceptably robust conceptual hydrogeologic models for a particular site. The broader the available database, the

more robust the conceptualization likely becomes. Key data categories include site and regional physiography, topography, climate, meteorology, soils, vegetation, land use, geomorphology, geology, geophysics, surface and subsurface hydrology, inorganic and organic hydrochemistry, radiochemistry, natural and anthropogenic isotopes, and remotely sensed data. Of special relevance to hydrogeologic model development are regional and local site data that allow one to define the distribution of hydrostratigraphic units on a variety of scales; their geologic structure; rock and soil types; their textural, physical, flow and transport properties; fluid types; the states of fluid saturation, pressure, temperature and density; chemical constituents and isotopes; major contaminants in soil, rock and groundwater; and their sources.

Common constraints on groundwater monitoring networks include budgetary limitations, including availability and stability of funding, opportunities for developing dynamic GMNs, access to appropriate sites, environmental impact from the monitoring setups, availability of staff with appropriate skills, and technical limitations such as remote data collection.

2.2.3 Methods of Groundwater Monitoring Network Design

Existing GMN design methods differ in terms of the objectives of the design, the invoked physical model, the methods and objectives of optimization, the use of statistical representations of data, and the dimensionality of the spatial domain. Various classifications of GMN design methods have been proposed. One of the earlier classifications of GMN design methods (Hassan, 2003) is shown in Fig. 2-7. The two main categories and their branches were identified earlier by Loaiciga et al. (1992). However, the different methodologies developed under each of the general themes of Loaiciga et al. (1992) have been added to each methodology associated with either a network design theme or a network augmentation theme (Hassan, 2003). The author noted that that many of the developed approaches belong to more than one of the categories in Fig. 2-6. For example, a probability-based approach may be cast within an optimization framework, a variance-reduction approach may be based on kriging or co-kriging, and a simulation approach may be linked to an optimization technique. Kollat et al. (2011) recently listed such categories of methodological approaches that include expert-based hydrogeologic judgments, deterministic physics based simulations, geostatistical interpolation approaches, Monte Carlo physics-based simulations, and statistical filtering/data assimilation.

One of the primary distinctions among GMN design methods is whether the method uses a physics-based mathematical model of subsurface flow and transport (FTM), or employs geostatistics and statistical inference based on the past observations. Each of the two approaches has advantages and disadvantages. Grabow et al. (2000) noted that geostatistical methods as generally applied are essentially interpolation techniques and are most useful for siting wells to better define a plume once it is located and its boundaries are identified. However, these methods are not as good at extrapolating as they are at interpolating, and thus may be of limited help in predicting plume movement. On the other hand, flow and transport model-based (FTM-based) methods provide essential forecast information on plume movement, and thus can be instrumental in siting wells to characterize future plume development. The forecast skill of such models, however, depends on the conceptual model of the subsurface and can be seriously compromised if the conceptual model is flawed.

Recent developments in geostatistical methods in GMN design have progressed well beyond the optimal interpolation approach. An advanced, purely geostatistical/statistical toolbox, MAROS, was introduced with much success in 2003 (Aziz, 2003, Ling et al., 2004). Additional work in the same period was done on redundancy reduction using geostatistics (Nunes et al., 2004; Wu et al., 2005). Since then, geostatistics-based methods began to include FTM components for various

reasons. For example, Rivest et al. (2012) noted that groundwater contaminant plumes often display a curvilinear anisotropy, which conventional kriging and geostatistical simulation approaches generally fail to reproduce. They proposed to use a FTM to relate the spatial concentration patterns in natural coordinates directly with simulated flow patterns. Shlomi and Michalak (2007) proposed two new geostatistical tools based on using spatial structures of simulated flow and transport extracted from a FTM for spatial interpolation and uncertainty estimation. The new tools yielded results that are superior to those obtained by kriging, with a better reproduction of the true plume shape and lower uncertainty. Kollat et al. (2011) further demonstrated advantages of using data assimilation in GMN design. Data assimilation assumes knowledge of uncertainties in both the measurement and the modeling results. To estimate uncertainties in the modeling results, Kollat et al. (2011) used the Ensemble Kalman filter, with uncertainties in the modeling results being estimated from the ensemble FTM. The trend of merging geostatistical methods and FTMs seems to be very productive and is expected to continue.
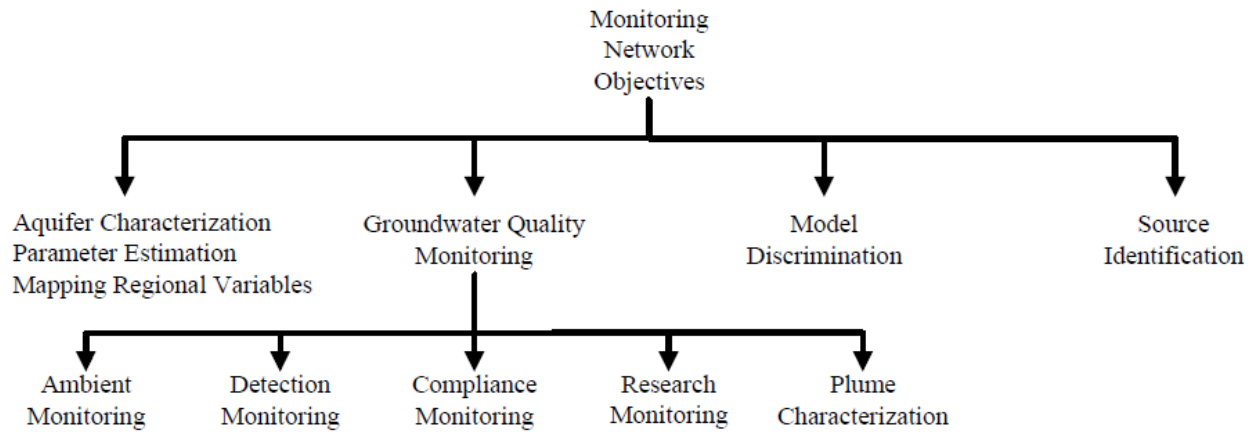
Figure 2-6   Classification of groundwater monitoring network objectives (after Loaiciga et al, 1992).
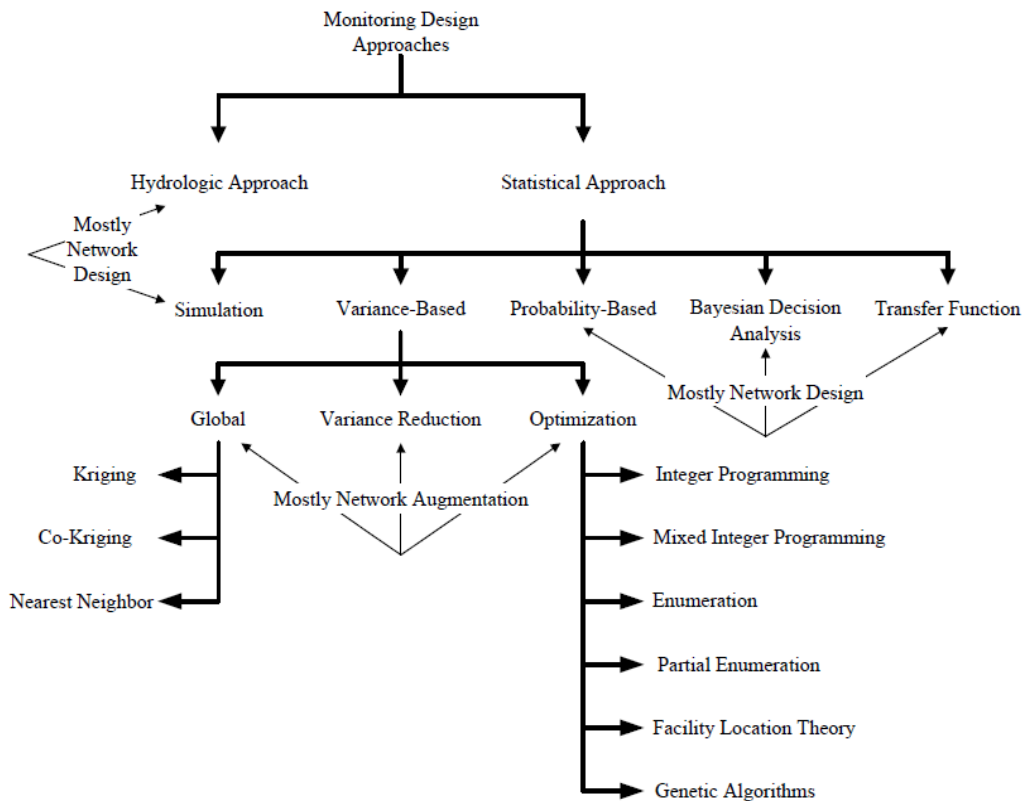
Figure 2-7   Classification of groundwater monitoring design approaches (after Hassan, 2003).

2.2.4 Groundwater Monitoring Network Design Based on Flow and Transport Models

Developments in using flow and transport models (FTMs) in the design of groundwater modeling networks (GMNs) during the last 20 years are summarized in Table 2.1. The table shows that FTMs have been used most often within a Monte Carlo simulations framework. Random selections of flow and transport parameters, locations of the contaminant sources (leaks), and boundary conditions are then used to develop probability distribution functions (PDFs) of contaminant concentrations at predefined locations. In several cases it was sufficient to obtain statistical moments, rather than full PDFs. Less often FTMs have been used to determine sensitivities of simulated concentrations to the parameter values. Deterministic representations of plumes with a single set of subsurface media parameters were used most often in past, but not lately.

FTMs have been implemented for both static and dynamic network designs. Proposed dynamic FTM-based designs mostly considered only augmentations of an existing network. For example, the assumption is generally made that wells in operation before network augmentation will continue to function after the expansion (Hudak and Loaiciga, 1992; Lee and Kitanidis, 1996; Dhar and Datta, 2007; Tonkin et al., 2007; Kim and Lee, 2007; Chadalavada and Datta, 2008; Datta et al., 2009). Some authors did not specify a trigger for the network expansion, while others assumed that the management stages were defined such that eventually the well network would be augmented (e.g. Grabow et al., 2000). The common assumption is that the period of additional well installation is much shorter that the period between network augmentation events. Potential locations for the new wells typically were set at the nodes of a predefined grid (e.g. Cieniawski et al., 1995). Use of a dynamic design emphasizes efficient use of collected information, such as allowing the past behavior of the plume to serve as a predictive model for placing future wells (Grabow et al., 2000). The term "sequential design" has been utilized seemingly interchangeably with the term "network augmentation." Sequential GMN designs appear to be much more cost-effective than static one-time setups such as used in the work of Kim and Lee (2007).

Two-dimensional flow and transport domains have been used most often in the past. Three-dimensional layered domains with wells located in different layers were investigated by Hudak and Loaiciga (1992), and more recently by Kollat et al. (2011). The computational load using 3D models in the past called for simplification of the solute transport simulations by using particle tracking methods (Storck et al., 1997). Evidence suggests that in some situations two-dimensional simulations are poor approximations of natural three-dimensional systems (Hassan et al., 1998). Based on a case study review, Montas et al. (2000) concluded that 2D models can be very useful tools to study and predict contaminant behavior in natural aquifers, and thus can be used to design monitoring well networks to detect and characterize plume spreading and evolution in regional aquifers.

The objective of FTM-based GNM designs have been related mostly with the detection, delineation and characterization of contaminant plumes. The probability of detection associated with any GMN design can be expected to increase with increased density of the network and/or increased frequency of sampling. A sensitivity analysis by Meyer et al. (1994) showed that the predicted performance of a given number of wells can decrease significantly as the heterogeneity of the porous medium increases. Poor estimates of the hydraulic conductivity were further shown to result in too optimistic estimates of network performance.

Model discrimination was addressed in early studies by Usunoff et al. (1992) and Nordquist and Voss (1996). The former authors noted that conceptual model uncertainties often are the main

source of prediction uncertainties. Orders of magnitude differences in peak concentrations and arrival times were found when performing long-term predictions with the two models used in their study. They emphasized that simulations of the proposed experiment with all models are necessary but not sufficient since data obtained from the experiment could be fitted with several of the models. Nordquist and Voss (1996) tested the hypothesis that measurements in regions where alternative models produce the most divergent predictions will give the best information possible for deciding which of the candidate models is the most appropriate.

Until recently few if any studies have addresses the effects of uncertainty in observed data on GMN designs using FTMs. Datta et al. (2009) may have been the first by using explicitly treated measurement errors as normally distributed additive terms. They indicated that source identification errors depend upon assumptions in the measurement error characteristics.

A variety of optimization algorithms have been used to solve the extremely complex problem of well placement optimization under multiple constraints. Substantial progress has been achieved with the advent of evolutionary algorithms (Minsker, 2003). Comparison of results obtained with different algorithms showed potentially substantial differences between different multi-objective optimization approaches (Cieniawski et al, 1995; Kollat and Reed, 2006). The ARS study team note that monitoring designs for FTM parameter estimation are not considered in this report; reviews and summaries of progress can be found elsewhere (e.g., Tsai et al., 2003; Hill and Tiedeman, 2007).

The need to monitor both the vadose zone and groundwater has been emphasized repeatedly in the literature. Hudak and Loaiciga (1999) provided a compelling argument that comprehensive subsurface monitoring programs should include contaminant detectors in both the vadose and saturated zones. This work demonstrated that vadose zone detectors can provide early warnings of an impending groundwater contamination problem, and also yield information relevant to placing groundwater monitoring wells. Ward (2006) reported that, at the Hanford site, fine-scale geologic heterogeneities, including grain fabric and lamination, were observed to have a strong effect on the large-scale behavior of contaminant plumes, primarily through increased lateral spreading resulting from anisotropy. To our knowledge, no studies on optimization of FTM-based coupled monitoring of vadose zone and groundwater have been reported to date. Although conceptually the purposes and constraints for monitoring flow and transport in a coupled groundwater and vadose zone system may be similar in some aspects to monitoring flow and transport in groundwater only, the process rates in groundwater can be quite different from those in the vadose zone. Opportunities for sampling of dissolved contaminants hence are also different.

Research on GMN optimization have only recently addressed uncertainties in the conceptual models that form the basis of the invoked FTM. These uncertainties are now increasingly acknowledged, and both model discrimination and model combination are viewed as potential directions to address these uncertainties (see e.g. Refsgaard et al., 2012).

Overall, the steady progress in physical model-based optimization of groundwater monitoring has resulted in the development of a variety of powerful and robust optimization methods. Multi-objective criteria are now increasingly being used successfully in many studies. Multiple sets of model parameters have been used similarly in Monte Carlo simulations and sensitivity analyses. Furthermore, geostatistics-based GMN design methods are now starting to benefit FTM simulations. In particular, FTM-based GMN designs can be improved by incorporating information about uncertainties in both data and the model predictions.

Table 2-1 Selected studies proposing methods for model-based groundwater monitoring network design

No.	Source	Static /Dynamic[§]	Objective	Optimization criteria	Optimization algorithm	Model use[¶]; model dimensions
1	Usunoff et al., 1992	S	Model discrimination	If distances between simulation vector results obtained for all model pairings are large enough, then the proposed experiment can be used most effectively to select one of the models (provided 'reality' is closely reproduced by that model).	None	Simulations for pairs of conceptual models, with each model calibrated first with measured data and then with measured data and simulation results from the other model; 1D.
2	Hudak and Loaiciga, 1992	DA	Plume characterization	Maximize coverage	Location problem, linear programming	Determine approximate plume location; 2D.
3	Hudak and Loaiciga, 1993	D	Detection	Maximize coverage	Location problem, linear programming	Determine approximate plume location; 3D (layered).
4	Hudak, 1994	S	Plume characterization	Cover a maximum number of siting horizons, and allocate wells to the single most effective horizon	Linear programming	Define plume cross-sections at siting horizons; 3D (layered).
5	Meyer et al., 1994	S	Detection	Maximize the number of realizations in which the plume is detected given the maximum number of wells.	Simulated annealing	Monte Carlo simulations to compute the probabilities of detecting concentrations above a threshold at each location; 2D
6	Cieniawski et al., 1995	S	Detection	Minimize number of undetected plumes and the average area of contamination of detected plumes.	Genetic algorithm	Monte Carlo simulations to compute the probabilities of detecting concentrations above a threshold at each location; 2D

#	Reference	Type	Application	Objective	Dual Control	Description
7	Lee and Kitanidis, 1996	DA	Groundwater quality monitoring	Minimize pumping and measurement cost	Dual Control	Monte-Carlo simulations with calibrated model to estimate uncertainty in model predictions to be used in extended Kalman filter updates; 2D
8	Nordquist and Voss, 1996	S	Model discrimination	Where results from different models differ the most	Ranking locations	Sensitivity-oriented simulations with each conceptual model; minimum and maximum values of parameters are combined in all meaningful combinations; 2D
9	Storck et al., 1997	S	Detection	Maximize detection probability, minimize cost, minimize volume of contaminated groundwater at the time of detection.	Simulated annealing	Monte-Carlo simulations with calibrated model to generate a list of potential monitoring wells that detect each contaminant plume as well as the contaminated aquifer volume when the plumes are detected first at each potential monitoring well; 3D
10	Grabow et al., 2000	DP	Plume characterization	None	None	Simulation to estimate the direction and movement of the plume, the rate of expansion, and moments of the contaminant mass distribution of the plume, 2D
11	Montas et al., 2000	S	Plume characterization	Minimize errors in contaminant concentration moments given the maximum number of wells selected from the original large set.	Partial enumeration	Monte Carlo simulations of flow for random conductivity fields; random walk for each simulation; design based on ensemble-averaged plumes

#	Reference		Application	Objective		
12	Bierkens, M. F. P., 2006	S	Detection	Minimize total monitoring strategy calculated by adding the risk of pollution to the costs of installing and maintaining monitoring wells and conducting routine chemical analyses.	None	Monte-Carlo simulations with calibrated model to generate probability distributions of contaminant concentrations; 2D
13	Dhar and Datta, 2007	DP	Groundwater quality monitoring	Maximize probability of detecting contamination at locations where a given pollution standard is exceeded, and minimize installation cost	Mixed integer programming	Monte-Carlo simulations with calibrated model to generate probability distributions of contaminant concentrations; 2D
14	Tonkin et al., 2007	DP	Groundwater quality monitoring	Minimize predictive variance	Ranking	Sensitivity-oriented simulations; sensitivities of simulated equivalents of both the observations and the predictions are found with respect to the model parameters
15	Kim and Lee, 2007	DA	Plume characterization	Minimize existing uncertainty due to well installation	Genetic algorithm	Monte-Carlo simulations with calibrated model to generate probability distribution functions of contaminant concentrations
16	Chadalavada and Datta, 2008	DP	Groundwater quality monitoring	Minimize probability of not monitoring pollutant concentrations at those locations where probable concentrations are large. Minimize estimation variances of pollutant concentrations at unmonitored locations.	Genetic algorithm	Monte-Carlo simulations with calibrated model to generate probability distributions of contaminant concentrations; 2D

#	Reference	DP/S	Category	Objectives	Method	Description
17	Datta et al., 2009	DP	Source identification + groundwater quality monitoring	Minimize normalized error of flux estimates; maximize the sum of trimmed means given a set of potential monitoring well locations, and a maximum number of monitoring wells to be installed.	Sequential quadratic programming	Monte-Carlo simulations with perturbed water fluxes to generate trimmed mean contaminant concentrations; 2D
18	Bashi-Azghadi and Kerachian, 2010	S	Groundwater quality monitoring	Minimize monitoring costs (number of monitoring wells), maximize the reliability of contaminant detection and maximize the probability of detecting an unknown pollution source.	Genetic algorithm	Monte-Carlo simulations with calibrated model to generate probability distributions of contaminant concentrations
19	Kollat et al., 2011	S	Plume characterization	Minimize monitoring costs, maximize information provided to the ensemble Kalman filter, minimize failures to detect the tracer, maximize detection of tracer fluxes, minimize error in quantifying tracer mass, and minimize error in quantifying the centroid of the tracer plume	Genetic algorithm	Monte-Carlo simulations with calibrated model to estimate uncertainty in the model predictions to be used in the Ensemble Kalman filter updates; 3D
20	Melles et al. 2011	S	Detection	Combined kriging-predicted error variance and population-weighted detection capability	Simulated annealing	Monte Carlo simulations based on multivariate distribution of measured weather data; 2D

| 21 | Reed and Kollat, 2012 | S | Plume characterization | Minimize monitoring costs, maximize information provided to the ensemble Kalman filter, minimize failures to detect the tracer, maximize detection of tracer fluxes, minimize error in quantifying tracer mass, and minimize error in quantifying the centroid of the tracer plume | Genetic algorithm | Monte-Carlo simulations with calibrated model to estimate uncertainty in the model predictions to be used in the Ensemble Kalman filter updates, 2D |

¶Studies in which modeling was used to develop a synthetic example for the optimization process have not been included;
§S – static network; DA - dynamic network with a single augmentation; DP – dynamic network with augmentations at end of predetermined phases of the project.

2.3 Integrating Model Abstraction into Monitoring Design

Systematic model abstraction can lead to a variety of conceptual models of potential use for optimization of the monitoring efforts. A conceptual model is always needed to address any of GMN objectives shown in Fig. 2-6. Therefore, as a result of the model abstraction process, multiple GMNs can be designed for purposes of aquifer characterization, parameter estimation, characterization of the background concentration, contaminant detection, compliance assessment, plume characterization, source identification, and model discrimination. Because each conceptual model has a flow and transport model associated with it, most of the above GMN optimization tasks can be approached with one of the methods shown in Fig. 2-1.

Abstraction of the model structure can, in particular, be used in GMN design efforts that address the model discrimination problem. The specific objective then will be to augment an existing well network for the purpose of optimizing model discrimination. It is generally accepted that the traditional approach of model uncertainty analysis, which considers only a single conceptual model, may fail to adequately sample the full space of plausible conceptual models. As such, the traditional approach is prone to modeling bias and underestimation of predictive uncertainty. The purpose of model discrimination is not to identify the best model but rather to provide insight into the behavior of the different models, thereby adding important information about subsurface properties controlling or affecting flow and transport. A reduction in the prediction uncertainty resulting from model discrimination can be expected in three categories of uncertainty as defined by Refsgaard et al. (2012): (a) the geological structure; (b) the effective model parameters; and (c) model parameters including local scale heterogeneity. Two approaches to GMN design for model discrimination have been used in subsurface contaminant hydrology so far. Knopman and Voss (1989) postulated that points of greatest difference in predictions can contribute the most information to the discriminatory power of a sampling design. They suggested the use of three objective functions in the design optimization: the sum of squared differences in predicted concentrations, the squared scaled difference, and the minimum squared difference. This approach has been successfully used in several field cases (Knopman et al., 1991; Nordquist and Voss, 1996). Another approach was proposed by Usunoff et al. (1992). They formulated the optimization problem within a maximum likelihood framework and defined the performance of each model in terms of its ability to fit existing experimental data as well as data predicted with other models at additional monitoring points. Both approaches do not impose any restrictions on model structure and hence can be applied with the abstracted models.

Several GMN designs are possible if alternative conceptual models are proposed as a result of model abstraction. Monitoring design optimization may then rely on weighted simulation results from the corresponding mathematical models to provide the required input for the optimization algorithms as outlined in Table 2-1. Bayesian model averaging (Ye et al., 2004) can be instrumental in determining model weights. Rojas et al. (2010) recently used a combined GLUE-BMA approach to analyze prediction uncertainties generated with eight different conceptual models. The models predicted substantially different flow processes outside of the studied area. Having an appropriate site-specific GMN design (or designs) is then important to decrease model uncertainty. If the scale of the prediction domain is substantially different from the scale of the domain of the initial observations, then the abstraction of scale change can provide additional conceptual and mathematical models to be included in the monitoring design using methods outlined in Table 2-1.

Metamodeling abstraction is currently already being used for GMN designs. As shown by Arndt et al. (2005), replacing time-consuming model runs with computations using regression equations appears to be very efficient in designing optimal monitoring schemes for contamination sources.

Abstraction of parameter determination includes the application of pedotransfer functions that allow one to generate model ensembles for application to flow in variably saturated soils (Guber et al., 2006; Guber et al., 2008). Model-based GMN design methods include the use of ensemble modeling with Monte Carlo-based randomly generated hydraulic and transport parameters (Table 2-1). It has been noted that such methods for generating parameters can be hampered by uncertainty about possible correlation between parameters and appropriate ranges of some parameter values. The use of pedotransfer functions may alleviate this problem since the parameters of water retention and (if available) the hydraulic conductivity are not randomly generated but fitted to measured data. Another abstraction of this type – zonation – has already been used successfully in the design of GMNs (Tsai et al., 2003).

Most of the model-based GMN design methods have been developed using assumptions that numerical calibration of the models can be performed, that parametric sensitivities can be determined, that measurement errors can be neglected, that the vadose zone can be neglected, and that a one-dimensional solute transport model is fully acceptable. One or more of these assumptions may not be applicable in most practically important cases. In this study the ARS study team develop or modify heuristic techniques that will avoid the need to make these types of assumptions (Section 6). The techniques will be applied to models that are obtained via model abstraction for full 3D flow and transport in a variably-saturated flow domain.

3   EXPERIMENT TO TEST AND CONFIRM MODEL ABSTRACTION FOR MONITORING FLOW AND TRANSPORT IN SOILS

In this study, the ARS study team applied a systematic abstraction process to base models accounting for heterogeneities in subsurface flow and transport. The abstraction will lead to often-used simplified models, with as overall objective the estimation of pattern and magnitudes of possible errors obtained with the simplified models. The purpose of this section is to provide an analysis of the exhaustive ARS OPE3 vadose zone flow and transport database within the context of model abstraction, and to describe the experiments and measurements carried out to develop the dataset and the comprehensive model of flow and transport. This model and its abstractions have been used to demonstrate the integrating model abstractions in developing monitoring designs for the ARS OPE3 site example.

3.1   Outline of Lateral Flow and Transport Experiment

A century-long history of experimentation on solute transport in soils has resulted in a multitude of advanced experiment setups and procedures, as well as methods of data analysis. Experiments may be carried out at different scales as illustrated in Fig. 2-3. Currently no single-source compendium of experimental methodologies exists to study flow and transport in soils, and to characterize the ability of soils to transmit and retain chemicals. One exception may be a recent monograph by Dane and Topp (2002). Applying a model to infer soil properties from experimental results is still the standard method of data analysis.

Generally, there are three sources of information about solute transport in soils: outflow breakthrough sampling, destructive soil sampling after transport has occurred, and monitoring of pore solution concentrations within the soil during the observed solute transport event. The applicability and relative reliability of those sources depends on the scale of the study.

Outflow breakthrough sampling is probably the most reliable methodology since it effectively integrates solute transport for the given support area or volume. However, its applicability is limited when no or little breakthrough is present. Such situations are typical for strongly absorbed chemicals, colloids and microorganisms, and generally for all solutes if the soil is unsaturated and transport is slow. Even if considerable breakthrough occurs, the presence of several alternative transport and retention processes may complicate the interpretation of results. This because multiple parameters describing those processes cannot be simultaneously inferred from a single breakthrough curve.

Destructive post-experiment sampling is a useful complement to breakthrough data in that this may help to diminish or eliminate ambiguity in parameterization of the retention processes in soil column experiments. As the scale coarsens, however, the resource demands may become daunting and sample cross-contamination may become an issue, while field variability may further compromise data reliability. A destructive sample may not contribute much to understanding the physical or chemical heterogeneity of a soil.

Monitoring the pore solution composition is in principle the best approach for obtaining data. Unfortunately, monitoring devices tend to distort the concentration field within a soil. It is often not clear what part of the soil solution affects the sensor reading, or what part of the pore solution is extracted. As the observation scale coarsens, field variability introduces additional uncertainty in the data.

Although the ambiguity of data and/or their analysis undoubtedly creates difficulties in solute transport experiments at the field scale, the importance of the vadose zone in controlling chemical fate and transport in the subsurface requires the development and application of field-scale predictive models. These field-scale models require data for their calibration and testing. A list of structural heterogeneities that are important to consider within the context of vadose zone modeling is given in Appendix A. Experimental setups in general should addresses two types of heterogeneities: differences in water mobility in different parts of the soil pore space, and the effects of restricting layers and the capillary fringe on lateral transport.

The field experimental setup was implemented in this study to evaluate the potential contribution of the variably-saturated zone and the capillary fringe to lateral chemical transport along a restrictive fine-textured layer. The lateral flow and transport processes were found to be relatively complex since the perched groundwater table fluctuated substantially in time. A schematic of the experiment is shown in Fig. 3-18. The tracer application area during the experiments was continuously irrigated, leading to a gradually expanding groundwater mound. The composition of groundwater during the experiment was monitored using observation wells located within the projected plume. The chemical application area had a radius of several meters to accommodate the scale of possible releases of contaminants from waste storage facilities. A conservative tracer was used. No soil solution sampling during the experiments was envisioned. The irrigation rate was selected such that the soils remained unsaturated. Soil moisture sensors had been installed to monitor the soil moisture content. Groundwater depths were measured, while surface runoff was monitored and, if found significant, intercepted and measured. All vegetation was eliminated from the chemical application area. Data obtained from soil surveys, water content measurements and the tracer experiment itself provided information for a base model that was developed to accommodate the encountered complexities as explained in the following chapter.

3.2 The OPE3 Experimental Site at the Beltsville Agricultural Research Center

3.2.1 General Description

The study area was the 22-ha USDA-ARS "Optimizing Production Inputs for Economic and Environmental Enhancement," or OPE3, experimental site in Beltsville, MD. An aerial view of the site is shown in Fig. 3-1. The watershed in the area drains into a riparian wetland forest that contains a first-order stream (Gish et al., 2002, 2005; Chinkuyu et al., 2008). The vadose zone at the site was formed in fluvial deposits. Soils are mostly sandy loams and coarse sands on a clay layer with surface slopes ranging from 1 to 4%. A description of the soil cover is given in Table 3-1. The watershed is in agricultural use and has been under corn since 1998. Tillage practices were the same each year, with fields typically being disked about one month prior to a second disking operation, followed by planting.

OPE3 is a small watershed site that is intensively instrumented, and where all data are geo-located with a differential or kinematic global positioning system. Data collection included the following:
- Over 40 km of ground penetrating radar (GPR) data were collected and analyzed.
- Every two years, soil cores were extracted to determine spatial correlations and distributions of organic matter, pH, and sand, silt, and clay percentages.
- Electromagnetic induction (EM-31 and EM-38) data were collected for two of the sites.
- 36,000 volumetric water content measurements at 48 locations were collected daily.
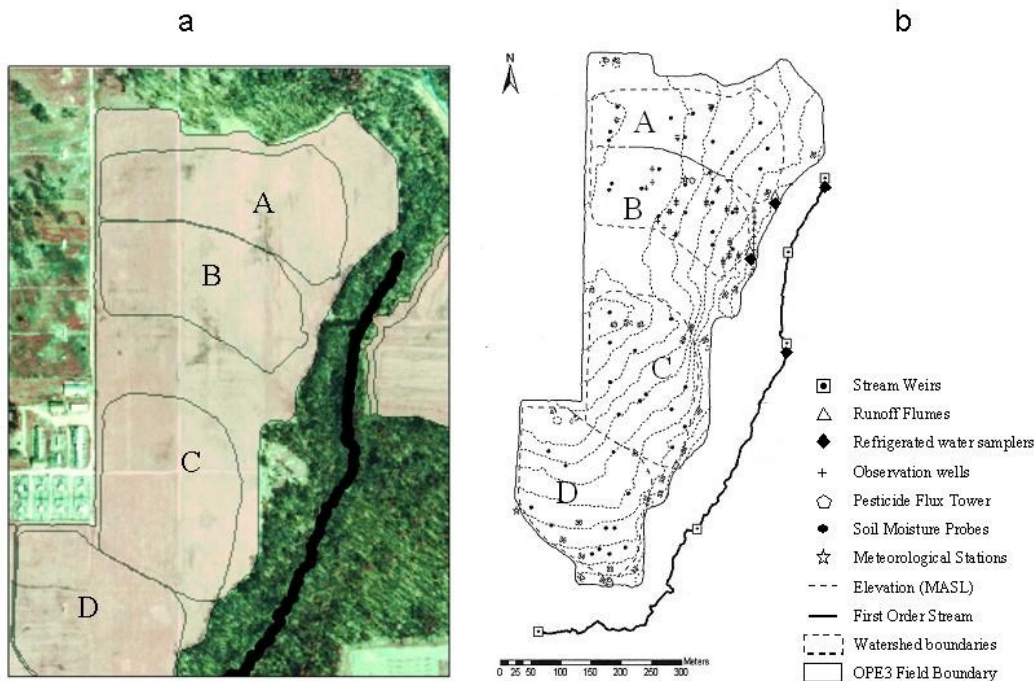
Figure 3-1   The USDA-ARS OPE3 research site; a – aerial view, b – instrumentation. A, B, C, D – research fields.

- Micro-meteorological stations with eddy covariance systems to monitor climatic conditions.
- Multiple pesticide vapor flux towers were operational after pesticide application.
- Water and chemical (N, P, and pesticides) runoff fluxes were collected from each field.
- Field B (Fig. 3-2) was instrumented with 52 groundwater observation wells.
- Corn grain yields were measured using a grain yield monitoring device.
- Aircraft and satellite remote sensing imagery were collected. Ground- and tower-based reflectances were collected as needed.
- Plant growth and development were measured periodically during the growing season.
- 180 observation wells in the riparian wetland were monitored for anions and pesticides.
- Stream flows and chemical fluxes in the stream were measured at five stations within the riparian wetland.
- Wetland soil cores were extracted up to one-meter depth and analyzed for grain size, bulk density, carbon content, hydraulic conductivity, water content, and denitrification potential.
- Dissolved gas was measured in groundwater samples throughout the wetland for evidence of denitrification and methanogenesis.
- Dissolved oxygen, dinitrogen, nitrous oxide, and methane were measured.

The experimental site of this particular study is located at the 3-ha field B of the OPE3 experimental site. The field is currently instrumented with 64 soil moisture sensors (EnviroSCAN, SENTEK Pty Ltd., South Australia) distributed among 12 capacitance probes (Fig. 3-2) that measure soil water content at depths from 10 to 180 cm every 10 min year-round. Table 3-2 shows installation depths of the soil moisture sensors. Field B was equipped with 52 groundwater monitoring wells, 41 of which are still active (Fig. 3-2). Some of the wells were instrumented with Divers (Van Essen Instruments, Giesbeek, The Netherlands) to measure groundwater levels every 30 min.

Other monitoring equipment at Field B included an automated Sigma 980 runoff flow meter (Hach Company, Loveland, CO) that measured water flow every 10 min and periodically collected surface runoff samples so that water and chemical fluxes could be determined. The flume was installed at the field surface outlet (Fig. 3-1b). The automated runoff sampler monitored fluxes on an event basis. An energy balance meteorological station with an eddy covariance tower was installed 20 m south of the boundary between Fields A and B (Fig. 3-1b).

3.2.2 Existing OPE3 Information Relevant to Flow and Transport in Soils

3.2.2.1 Small Scale Structural Units and Heterogeneities

The focus here was on heterogeneities that could be revealed at the core scale and could control the behavior of flow and transport at the field scale. The following information was used:
- Borehole data of soil texture and composition
- Soil water retention measurements on undisturbed samples,
- Field and laboratory soil hydraulic conductivity measurements.

3.2.2.1.1 Basic Soil Properties

Data on pH, organic matter content, Ca, Mg, K, P, sand, silt and clay contents had been collected during a soil survey performed by the USDA-NRCS at OPE3 site in 1997 using samples representing distinct genetic horizons. A detailed description of selected data is shown in Fig. 3-3. The soil pH was in a range between 4.3 and 7.5. The organic matter content varied from 0.12 to 4.65% and demonstrated a typical decrease with depth, indicating the absence of recent substantial changes in soil stratification caused by floods or erosion. The sand, silt and clay content exhibited large variations at each depth. Although the soils were characterized as having a coarse texture (Table 3-1), sand content varied in the range between 34 and 97 % (Fig. 3-3), with often drastic changes from one horizon to another in some profiles. Loamy sand and sandy loam classes dominated in soil texture. Only 9% of the samples belonged to sand, sandy clay loam and loam texture classes according to the survey (Fig. 3-4a). The relationship between soil clay and sand content was linear below 60 cm (Fig. 3-4b), which suggests a potential for having isolated lenses of predominantly coarse and predominantly fine material.
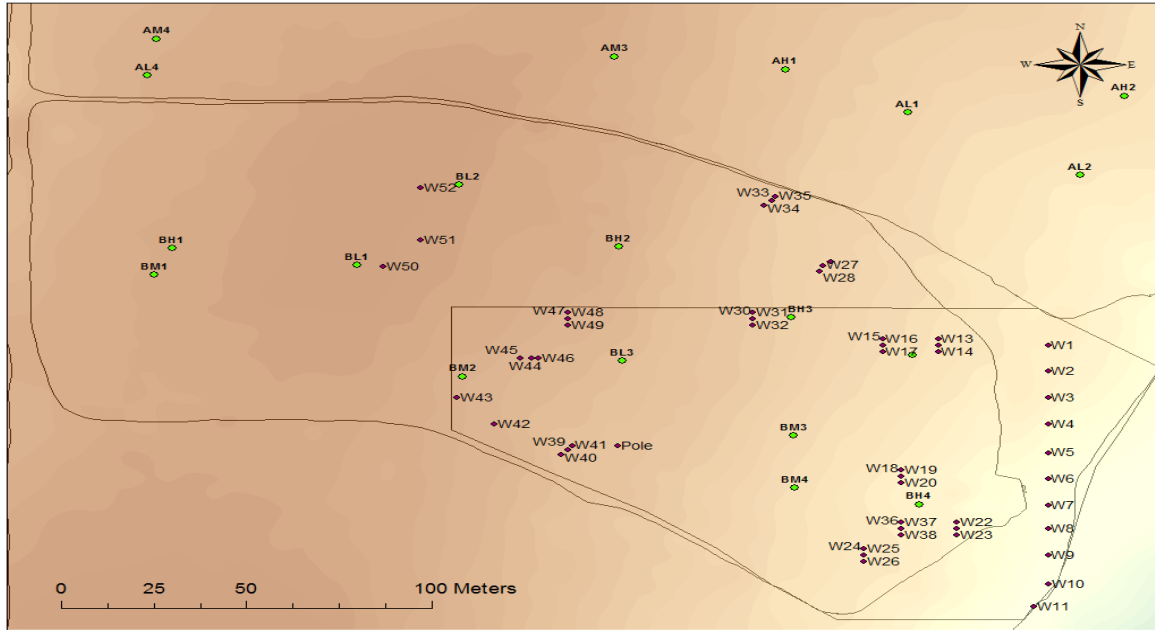
Figure 3-2   Locations of groundwater wells (brown circles) and soil moisture sensors (green circles) at field B.

Table 3-1 Soil cover of the OPE3 crop production area.

Soil associations and soil series	Coverage %	Topsoil texture	Description
Downer-Muirkirk-Matawan	49	Sandy loam	
Downer			Coarse-loamy, siliceous, semiactive, mesic Typic Hapludults
Muirkirk			Clayey, kaolinitic, mesic Arenic Paleudults
Matawan			Fine-loamy, siliceous, semiactive, mesic Aquic Hapludults
Bourne	23	Fine sandy loam	Fine-loamy, mixed, semiactive, thermic Typic Fragiudults
Matawan-Hammonton	23	Loamy sand	
Matawan			Fine-loamy, siliceous, semiactive, mesic Aquic Hapludults
Hammonton			Coarse-loamy, siliceous, semiactive, mesic Aquic Hapludults
Downer-Ingleside	5	Loamy sand	
Downer			Coarse-loamy, siliceous, semiactive, mesic Typic Hapludults
Ingleside			Coarse-loamy, siliceous, semiactive, mesic Typic Hapludults

Table 3-2   Installation depths of the soil moisture sensors at field B of OPE3 site.

Location	Depth (cm)						
	10	30	50	80	120	150	180
BH1	x	x	ns	x	ns	ns	ns
BH2	x	x	ns	x	ns	ns	ns
BH3	x	x	ns	x	ns	ns	ns
BH4	x	x	ns	x	ns	ns	ns
BL1	x	x	x	x	x	x	x
BL2	x	x	x	x	x	x	x
BL3	x	x	x	x	x	x	x
BL4	x	x	x	x	x	x	x
BM1	x	x	x	ns	x	x	x
BM2	x	x	x	ns	x	x	x
BM3	x	x	x	ns	x	x	x
BM4	x	x	x	ns	x	x	x

ns stands for no sensor.

3.2.2.1.2   Soil Hydraulic Property Measurements

To evaluate possible effect of the macroporosity on the soil hydraulic properties, undisturbed samples having a 20 cm^2 support area were taken for hydraulic conductivity measurements at the same support size in both the laboratory and the field. An example of water retention data from three depths with increasing clay content is shown in Fig. 3-5. Samples taken close to the soil surface showed well-defined macroporosity. A small decrease in the pressure head close to zero caused a substantial decrease in the water content of these samples. This decrease is a result of macropores being emptied. The effect is almost absent at 50 cm depth where the clay content is larger than close to the surface, and where root activity is also lower. Macropores may be less developed at this depth and hence are more difficult to detect here because of their infrequent appearance.

Data for the saturated hydraulic conductivity corroborated the water retention data in terms of a decrease in macroporosity with an increase in clay content and a decrease in root activity. As shown in Fig. 3-6a, the maximum values of the hydraulic conductivity were similar at depths of 25 and 40 or 45 cm. However, a substantial part of the soil, about 25% at well w50 and 50% at well w52, did not conduct water at a measurable rate at a depth of 40-45 cm where the effects of macroporosity on water retention were observed.

Anisotropy in the saturated hydraulic conductivity is another consequence of changes in soil structure. Fig. 3-6b shows that the horizontal hydraulic conductivity was smaller than the vertical conductivity at a depth of 25 cm, with about half of the samples showing distinct non-zero horizontal conductivities. At the depth of 45 cm, all samples had zero horizontal conductivity. The data are in agreement with visual observations of seepage along the walls of a soil pit where the measurements were taken (Fig. 3-7). Seepage can be seen at a depth of 60 cm where the lateral conductivity was substantially higher than the vertical conductivity.

3.2.2.2 Field Scale Heterogeneities

Considerable field-scale heterogeneity was evidenced by observed groundwater levels and chemical compositions, measured soil water contents, ground-penetration radar surveys, remote sensing imagery, and maps of measured corn yield.
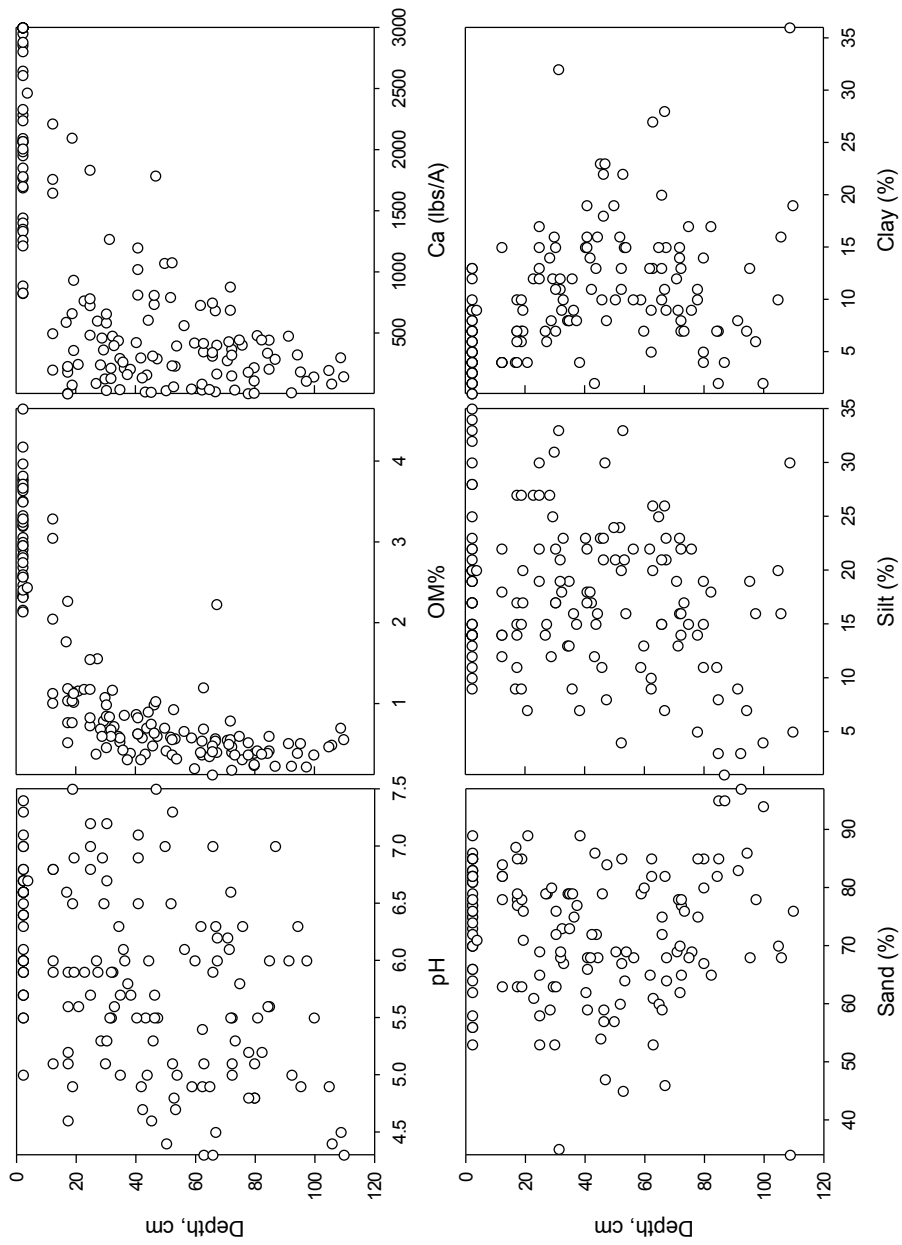
Figure 3-3   Basic soil properties at the OPE3 site based on data of soil survey conducted in 1999.
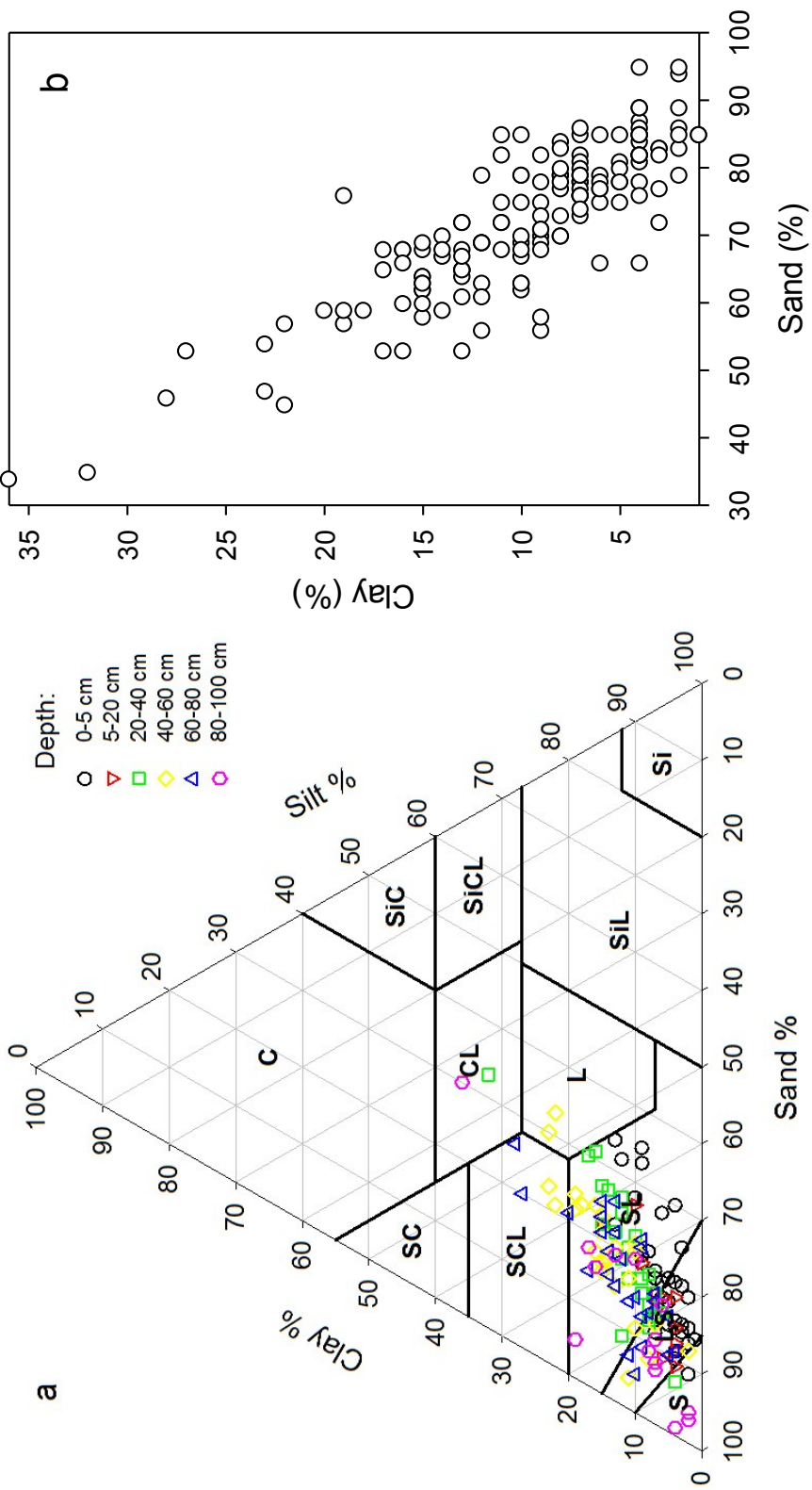
Figure 3-4   Soil texture at the OPE3 site based on data of soil survey conducted in 1999.
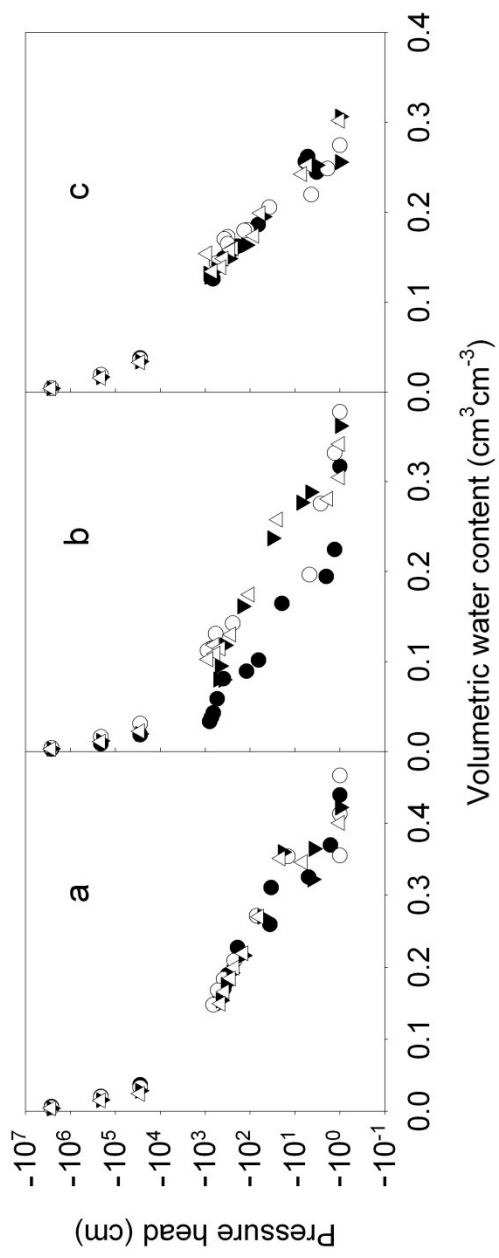
Figure 3-5 Soil water retention data of samples collected at locations near wells w50 and w52 of field B; 2-5 cm depth, b) 20-23 cm depth, and c) 47-50 cm depth (●,○ – well w50, ▽, ▼ - well w52).
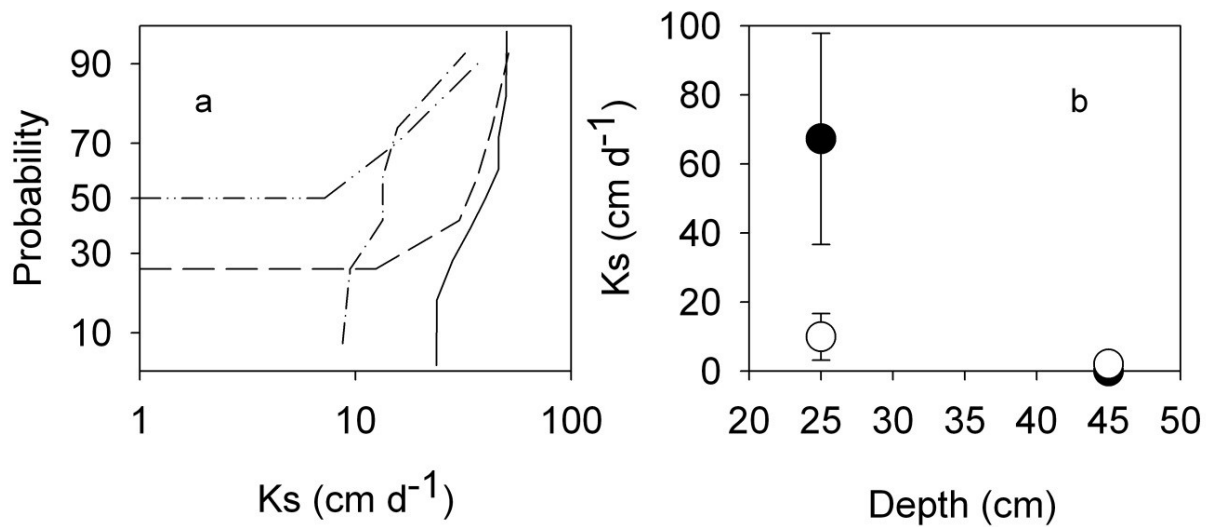
Figure 3-6 **Probability distributions of the saturated hydraulic conductivity for 20 cm2 support areas; a – field falling head measurements; ———— - well w50, depth 25 cm , — — —- well w50, depth 45 cm, — · — · — - well 52, depth 25 cm, — ·· — ·· — - well 52, depth 40 cm; b – constant head measurements of vertical (K_{SV}) and horizontal (K_{SH}) conductivities of undisturbed samples with 20 cm^2 support area in the laboratory (error bars show standard deviations).**

Figure 3-7   Water seepage from the wall above a locally restrictive soil layer near well w50.

3.2.2.2.1 Hydrologically Active Zones and Subsurface Flow Pathways

Initially two sets of independent data were collected to determine optimal locations for subsequent soil core sampling. Photogrammetric and geophysical techniques were first used to provide a general characterization of the experimental site. Because surface slope and aspect influence both surface and subsurface hydrology, a 0.25-m contour interval topographic map was derived from a stereo pair of aerial photographs and ground control points located with a submeter, a differential global positioning system (DGPS) receiver (Trimble 4000-SE, Sunnyvale, CA).

Electromagnetic induction (EM-38) was used to estimate relative clay contents near the soil surface in order to estimate infiltration rates (Doolittle et al., 1994). The results showed EM-38 values from 5 to 30. The EM data were divided into three equally sized populations, low (EM < 12), intermediate (17 >EM > 12), and high (EM >17) values.

The entire 7.5-ha site was next divided into different hydrologic groups based on surface topographic features and EM data. The ARS study team assumed that areas with EM > 17 and a slope >2% would have a relatively low infiltration capacity, whereas areas with EM <12 and a slope < 1% would have relatively high infiltration capacities. All other areas were assumed to have medium infiltration capacities. Within each hydrological region of potential infiltration, 7 or 8 plots having dimensions of 25 x 25 m^2 each were randomly selected for higher resolution surveys (Fig. 3-8).

A ground-penetrating radar survey was conducted in 2000. Data were acquired for the entire 7.5-ha field B along parallel north-south transects 25-m apart using a 150 MHz antenna. Within selected 25 by 25 m plots, additional GPR data were collected by towing the 150-MHz antenna along north-south transects that were 2 m apart. GPR data were acquired in digital form so that a trace of the subsurface reflections could be produced using RADAN software (Geophysical Survey Systems, Salem, NH). The GPR data were distance-normalized and low-pass filtered prior to data interpretation. The GPR trace followed the shallowest contrasting dielectric discontinuity. Strong dielectric reflections were considered to be manifestations of water holding capacity differences due to textural discontinuities such as the presences of a clay lens below a sandy soil. Generally, the clay lens (high dielectric) occurred below the C horizon, which frequently contained much gravel (low dielectric). Depths to the strongest reflection were as shallow as 0.9 m and as deep as 3.4 m, but the majority of data gave the strongest reflection at depths of 1.3 and 2 m. A GPR image profile is shown in Fig. 3-9. The first continuous restricting layer in this figure was situated immediately above the first continuous strong reflection, shown as a dotted line between depths of 1 and 2 m. The reflection of the restricting layer was not registered in about 5% of the total transect length. The topography of the first restricting layer was constructed using kriging assuming an omnidirectional variogram. The ARS study team used the software packages GEO-EAS (EPA, Las Vegas, NV) and GS_ (Gamma Design Software, Plainwell, MI) for this purpose. An example of the topography this obtained is given in Fig. 3-10. A first approximation of the subsurface flow pathways was constructed by subtracting the depth to the first continuous restricting layer from the surface elevation. The Arc/Info GIS hydrologic modeling tools FLOWDIRECTION and FLOWACCUMULATION were applied to a raster grid of the elevation-corrected subsurface topography to determine potential flow pathways (Fig. 3-10). The FLOWDIRECTION routine provided a grid of flow directions from one cell to its steepest downslope neighbor, while FLOWACCUMULATION determined the accumulated water from all cells that flowed into a particular downslope cell. The location of the restrictive layer was verified using multisensor capacitance probes installed at the blocks using a 2-m GPR survey, remote sensing imagery and a yield map (Gish et al., 2003, 2005). Sensors above the restrictive layer demonstrated the accumulation of water above the layer (Gish et al., 2002).
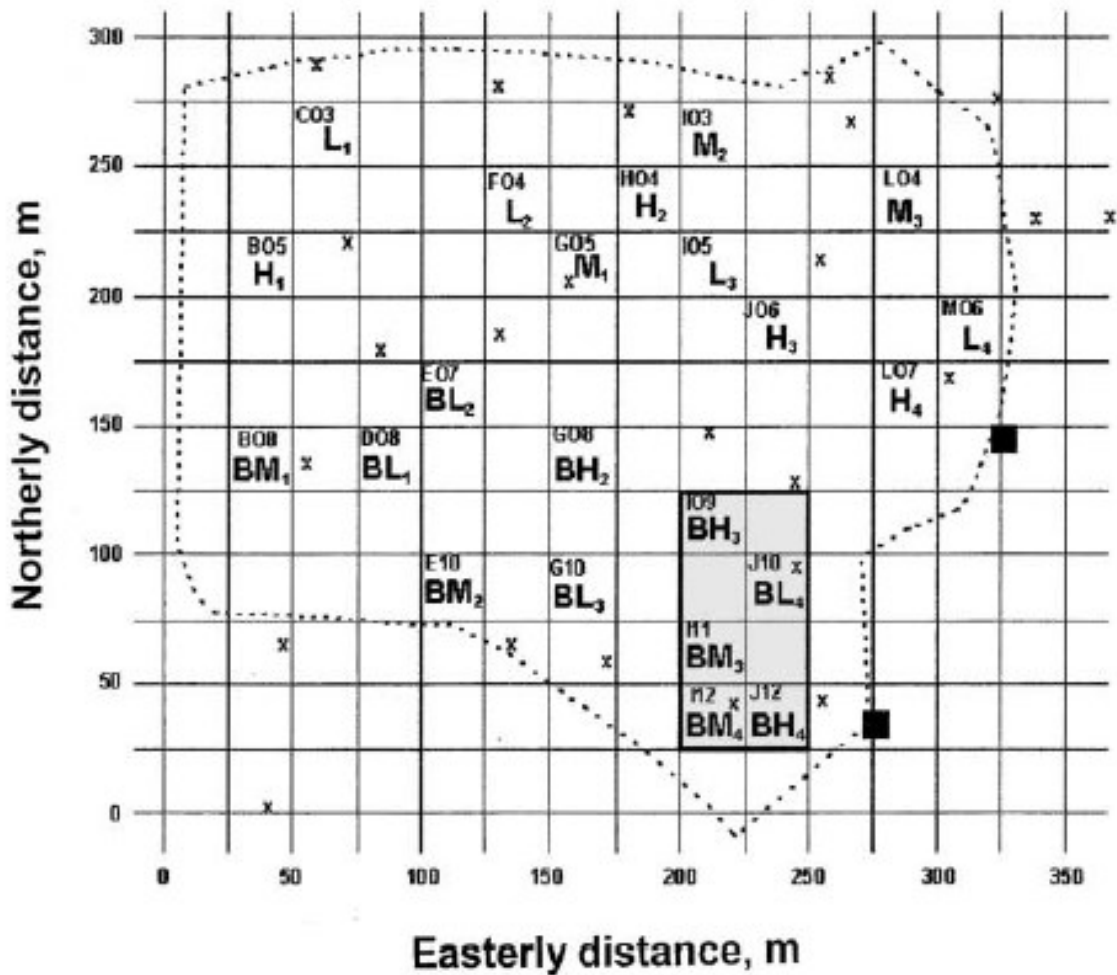
Figure 3-8   Layout of the 7.5-ha site overlain by a 25x25 m sample grid. Blocks with alphanumeric and infiltration capacity designations make up the stratified random sample of grid cells selected for soil moisture monitoring probes. The shaded rectangle in the southeast corner of the field site shows the 0.5-ha subsection that was used for soil moisture measurements and flow verifications (after Gish et al., 2002, with permission).
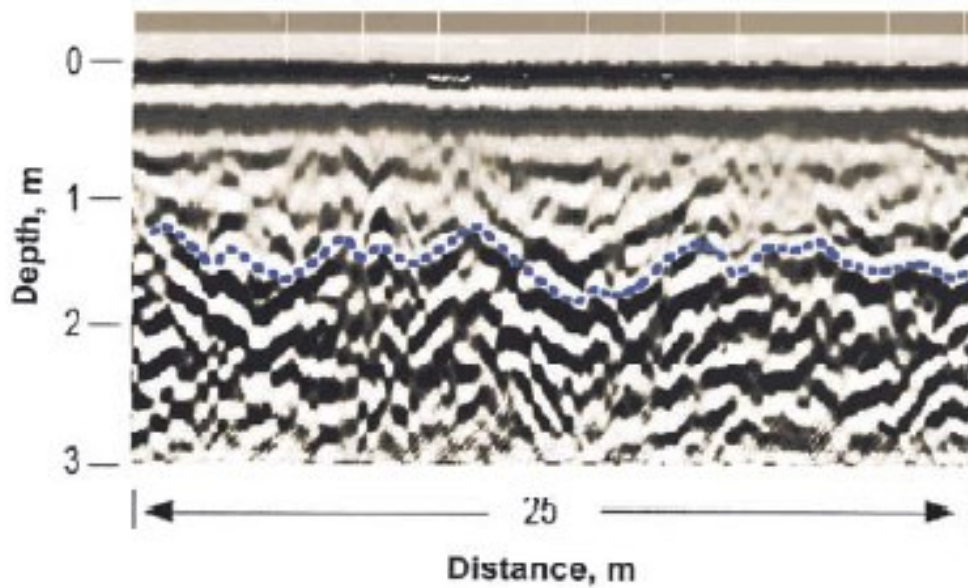
Figure 3-9 Ground-penetrating radar image with a digital trace of the first restricting layer, indicated by the dotted line (after Gish et al., 2002, with permission).
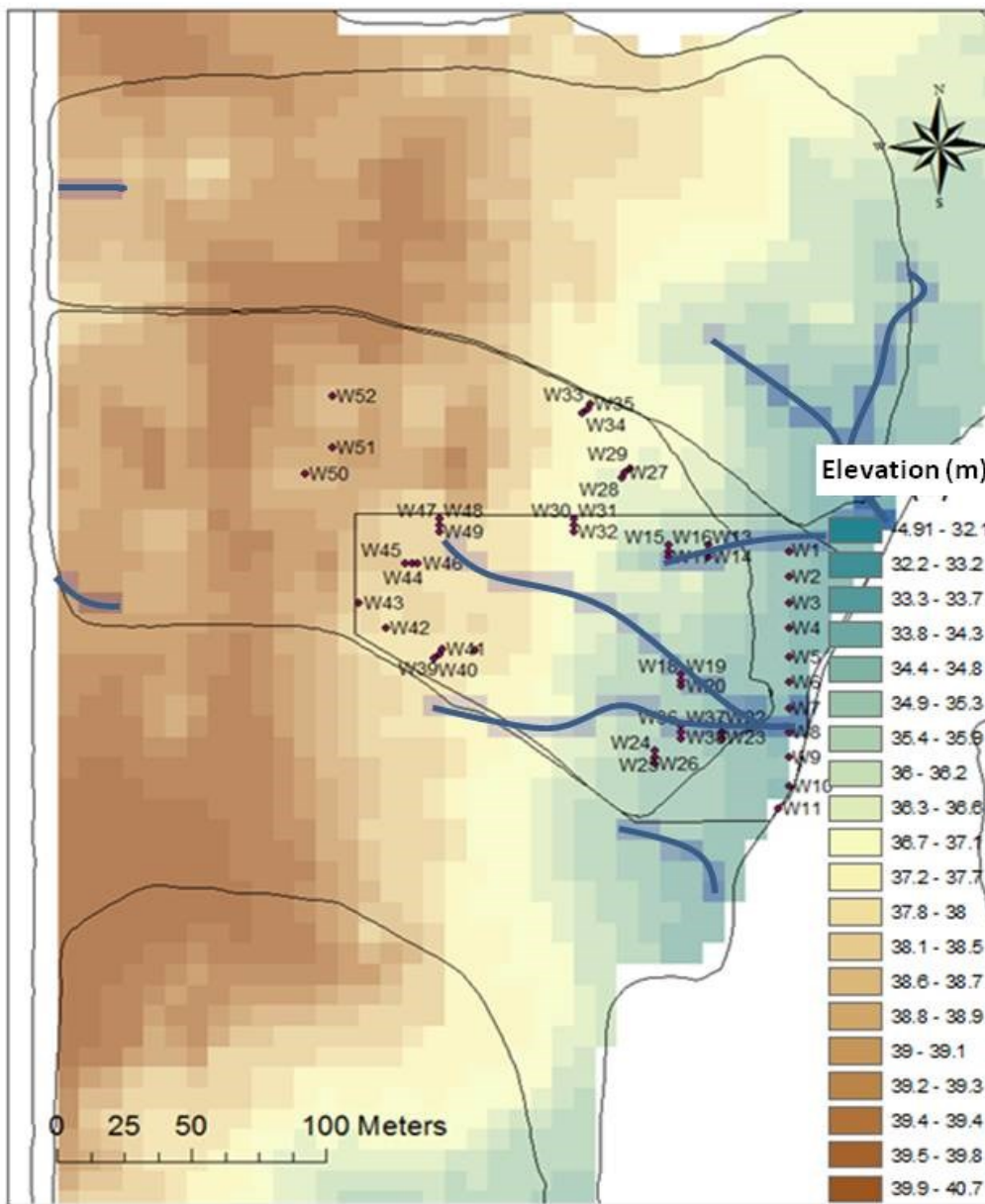
Figure 3-10 Interpolated elevation of the first continuous restricting layer with potential subsurface flow pathways (blue lines) identified using Arc/Info hydrologic tools.

Variability in the depth to the first continuous restricting layer also affected seasonal dynamics of the perched water tables. Groundwater depths monitored using Cera-Diver (Van Essen Instruments, Delft, The Netherlands) sensors installed in wells at a depth of 2.0 m in 19 locations across field B revealed high spatial variability in perched water levels (Fig.3-11). In some locations (e.g. w10, w11, w33, w39, w42-w44), the perched water level was consistently higher compared to other locations (e.g. w02, w12, w15, w18, w21, w36, w47). The amplitudes of groundwater level oscillations observed at locations with a shallow perched water were generally greater than in wells with relatively deep groundwater, thus indicating fast response of the water levels to rainfall in those locations. Since no differences in the oscillation phases were observed for the 19 observation wells installed at different locations across field B, one can conclude that the decline in groundwater level was controlled primarily by water permeability of the restrictive layer, and that locations with shallow groundwater were hydraulically more active, compared to the other locations. In other words water fluxes were greater in wells where the restrictive layer was close to the surface due to hydraulic connectivity between surface and groundwater. Preferential flow was also a factor, especially in the well 21.

Patterns of high crop density during water limited growing seasons were observed with remote sensing imagery using the Airborne Imaging Spectroradiometer for Applications (AISA) for areas corresponding to the GPR-identified subsurface flow pathways that could accumulate limited subsurface water at the site (Gish et al., 2003). During mildly dry years, corn yields decreased significantly ($P<0.01$) as the distance from the pathways increased (Gish et al., 2003).

3.2.2.2.2 Structural Units of Soil Cover

The multi-sensor capacitance probe (MCP) network at the site revealed the existence of areas where the soil was either consistently wetter or consistently dryer than the average of the study area. The different areas defined large spatial structural units in the top part of the vadose zone constituting the soil cover of the site. There was a noticeable similarity in the soil water storage time series measured at 8 locations of field B from May 1, 2001 through January 1, 2003 (Fig. 3-12). Time series from different locations appeared to be vertically shifted relative to each other.

The dense spatial coverage provided by remote sensing and yield mapping supported the existence of a spatial organization in soil cover at the site. Differences in soil water retention and hydraulic conductivity manifested themselves either by spatially different biomass values or yield variability, or both. Areas that produced low corn yields in 2001 (e.g., BL4, BM2, BH3 and BM3 in Fig. 3-13) supposedly experienced either more intensive waterlogging or higher soil water losses as compared to other locations. The spatial organization can be seen in Fig. 3-14, which shows an infrared image of corn biomass, superimposed on adjusted corn yields for 1999 (having a relatively dry growing season). Yields were greater than the mean inside the black polygons and less than the mean outside these polygons. Therefore, areas with high biomass (inside the black polygons/red colors) will likely produce higher evapotranspiration rates than the low biomass areas. Multiyear mapping of the yields supported the existence of zones where water availability to plants was consistently better or consistently worse than average (Fig. 3-15).
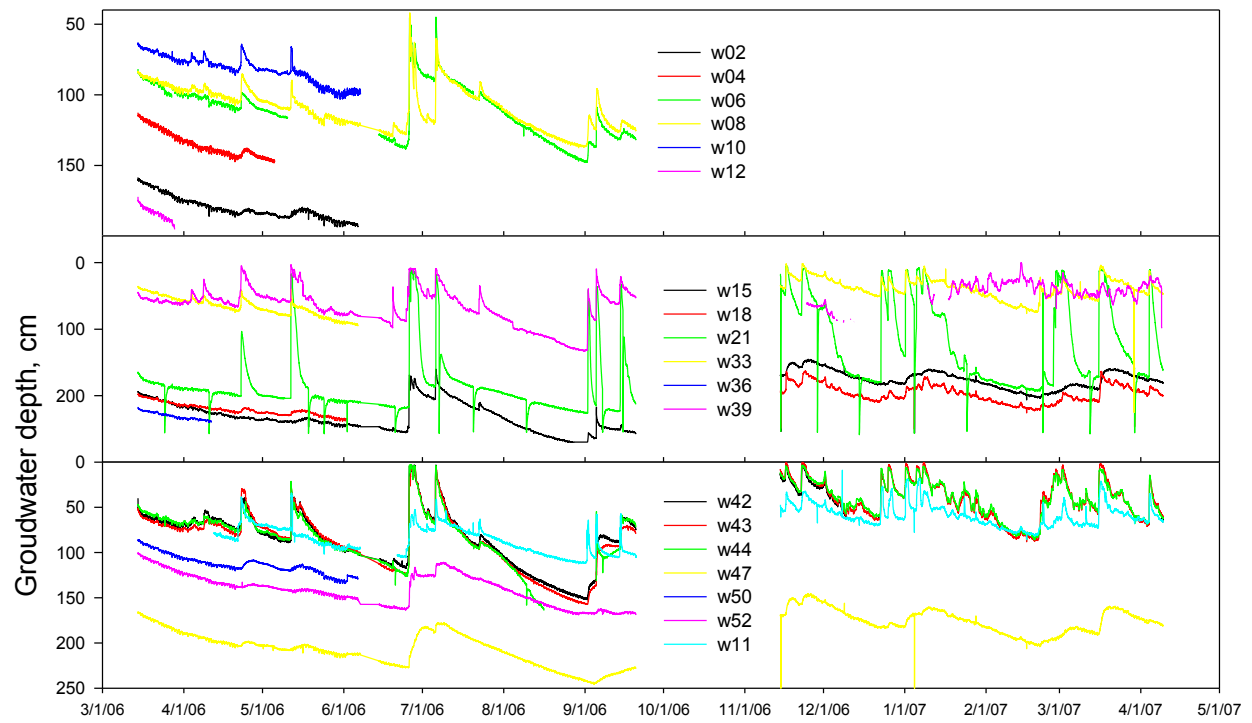
Figure 3-11   Time series of groundwater depths measured in field B in 2006-2007.
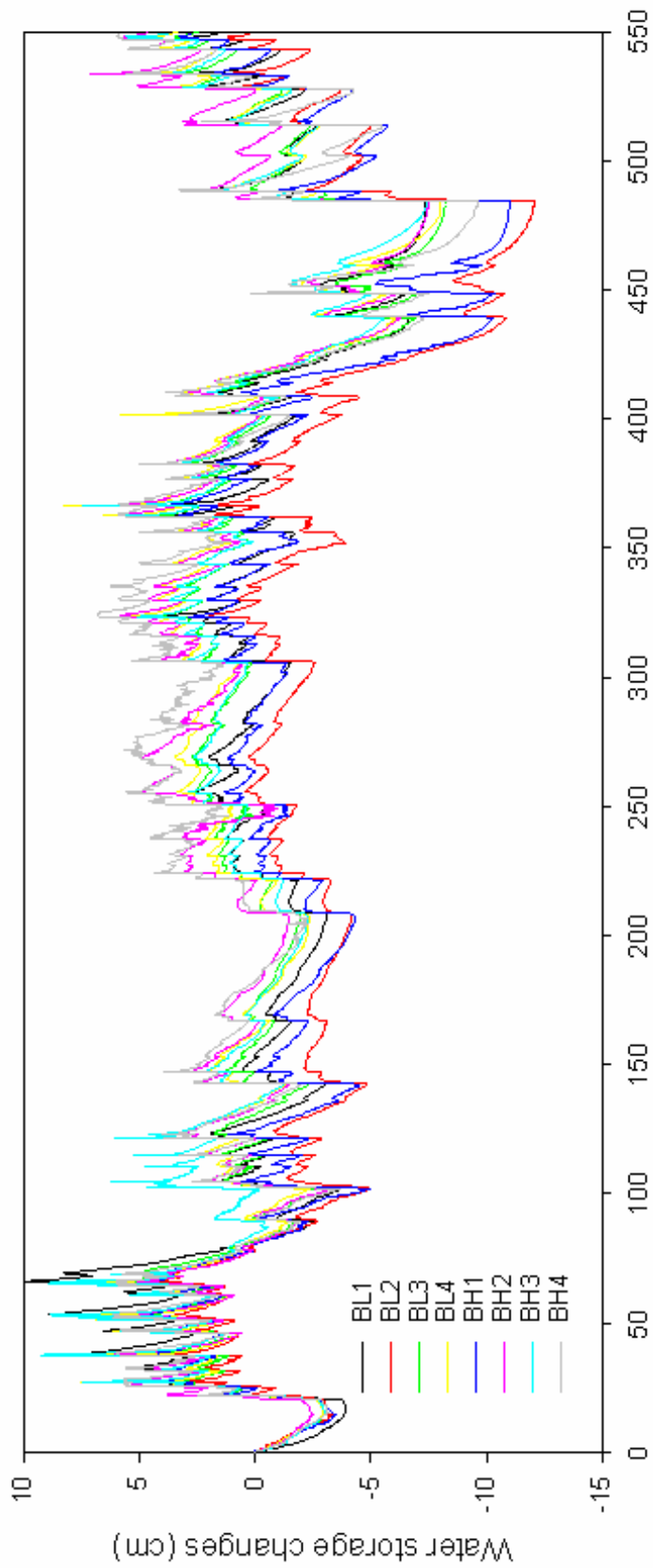
Figure 3-12 Changes in soil water storage of the 0-80 cm soil layer measured at 8 locations of field B from May 1, 2001 through January 1, 2003.
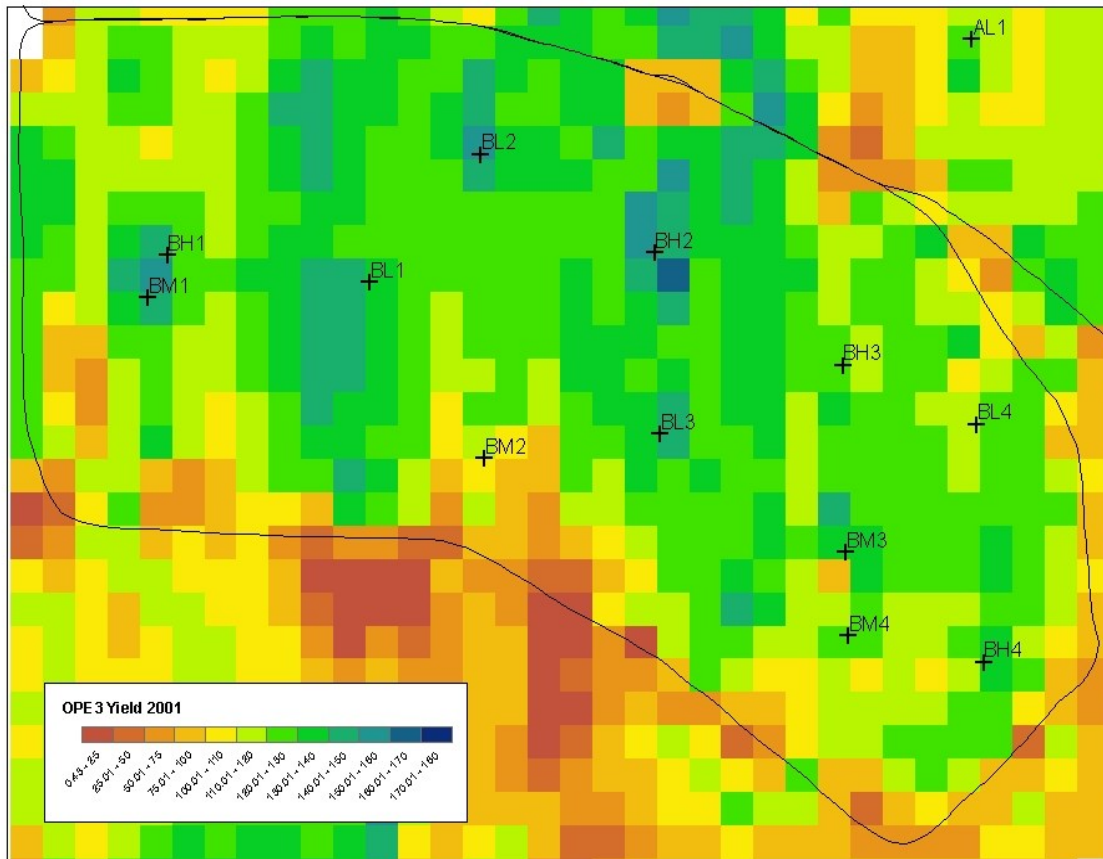
Figure 3-13 Spatial distribution of 2001 corn yields in field B. The red to yellow colors represent low corn yield areas and the dark blue colors high corn yield areas.
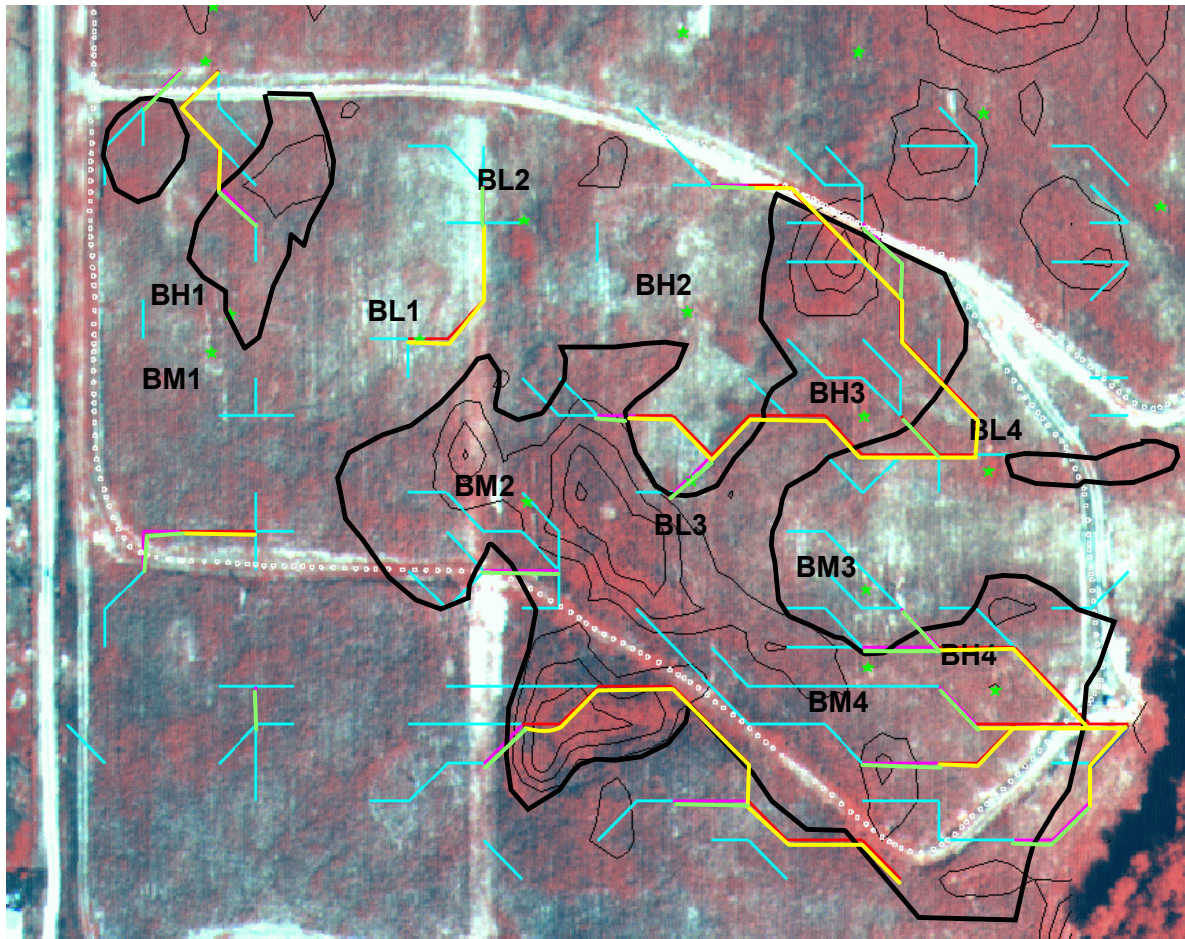
Figure 3-14 **Corn yields and color infrared image of corn biomass taken in August 1999 (in a relatively dry growing season). Green stars designate moisture probe locations, while the yellow, green and light blue lines designate subsurface flow pathway locations. The yellow color indicates regions with high biomass and the white regions low biomass, each reflecting extremes in the vegetative cycle of the corn plant. Corn grain yields greater than the mean are inside the black polygons, whereas areas outside the black polygons denote regions having yields below the mean.**
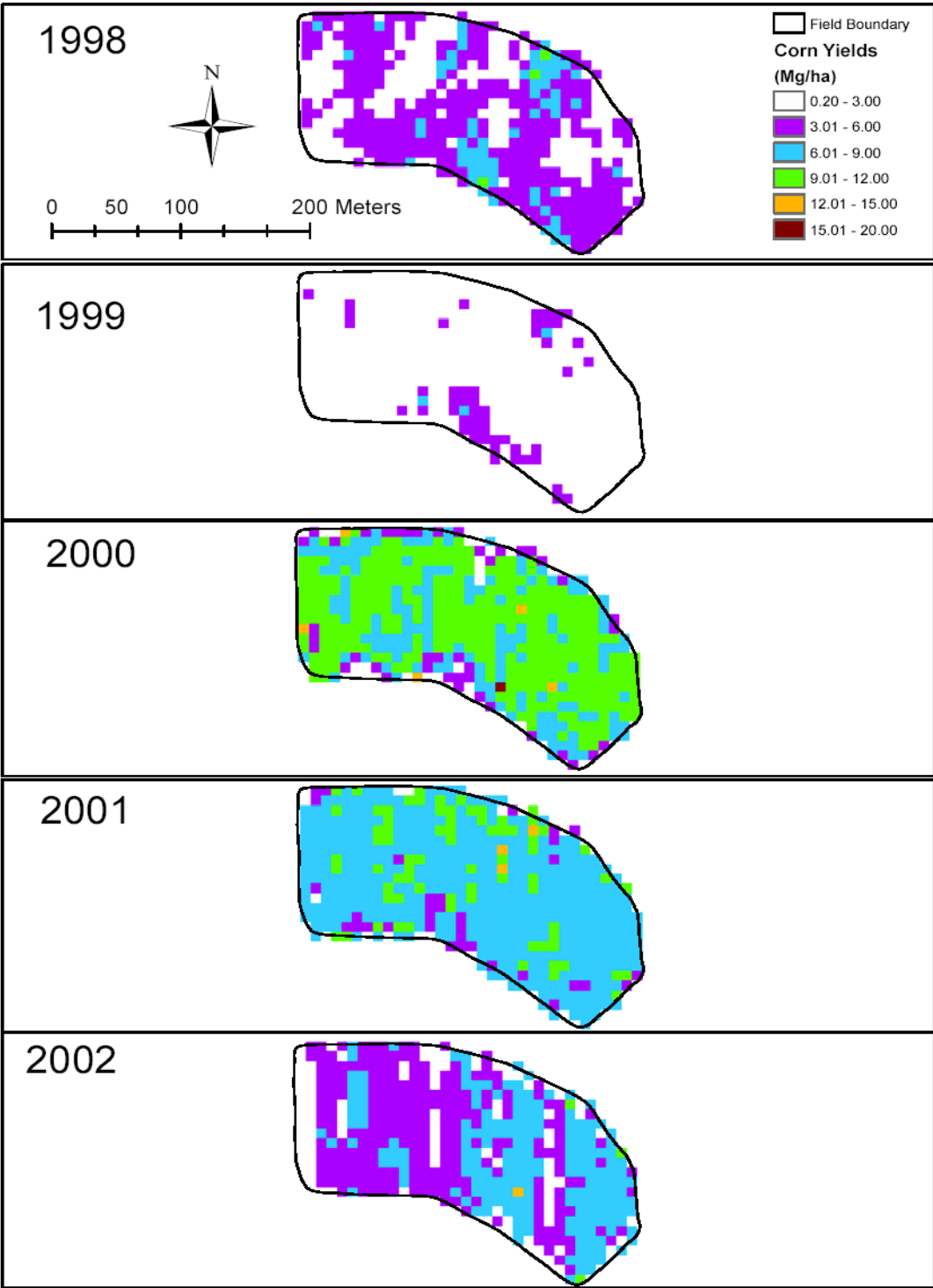
Figure 3-15 Spatial corn yield variations from 1998 to 2002.Note that 1998 and 1999 were relatively dry, 2000 was wetter than normal, 2001 was close to normal, and 2002 was dry.

Contaminant release typically occurs over an area much smaller than the biophysics-based unit. Still, it seems important to determine locations of different biophysics-based units, since the differences between units may reveal differences in the infiltration and groundwater recharge rates that will affect the transport of contaminant released over smaller area within a biophysics-based unit.

3.2.3 Surveying, Monitoring and Site-Specific Transport Experiments

The experiments as outlined in Section 3.1 envisaged tracer applications and monitoring in areas of 10x10 to 20x20 m^2. However, the OPE3 site had been surveyed in the past mostly at a much coarser scale. The flux measurement experiment used a location where more detailed surveys had been carried out in previous years.

An additional set of measurements was made for selection and characterization of the lateral flow experiment. The ARS study team decided to use for this experiment the intensively equipped South-Eastern part of Field B. Having a large number of groundwater wells (W18, 19, 20, 22, 23, 24, 25, 26, 36, 37, and 38) in this area was a major argument for selecting this location. However, the area had not been characterized at the fine scale during the ground penetration radar survey conducted in 2000 and described in section 3.2.2.2.1. Therefore, an additional fine-scale ground penetration radar survey was conducted to delineate boundaries of the soil hydrologic horizons for subsurface transport simulations.

3.2.3.1 Location Selection and Site Instrumentation

The location of the experimental site was selected to meet the following specific conditions:
  (a) The experimental area must be large enough to show spatial variability in soil properties;
  (b) The size of the irrigation plot should be reasonable for both a short-duration tracer application and a continuous uniform irrigation lasting several months;
  (c) The number of observation wells, and their spacing, should be large enough to reduce the influence of neighboring wells, while still capturing subsurface preferential fluxes within reasonable time intervals;
  (d) The site should exclude transit runoff and subsurface preferential pathways;
  (e) The site should not interfere with other ongoing long-term studies.

An experimental 10x10 m^2 site was selected in the intensively equipped South-Eastern part of Field B. Ground penetration radar data obtained in 2000 for this site indicated considerable variability in the depth to the restrictive soil layer (varying between 2.5 to 4.0 m), which was highest within the surveyed site. Two large branches of preferential subsurface pathways as identified by GPR survey started in this area, which potentially could be used for the field tracer experiment (Fig.3-10). The beginning of the main preferential surface flow pathway was found along the Southern boundary of the site on May 11 2006 after intensive rainfall. This suggests that preferential runoff would not affect the tracer study at the selected plot. Analyses of multi-sensor capacitance probe (MCP) monitoring data at locations BM2, BM3, BM4, BL3 and BH3
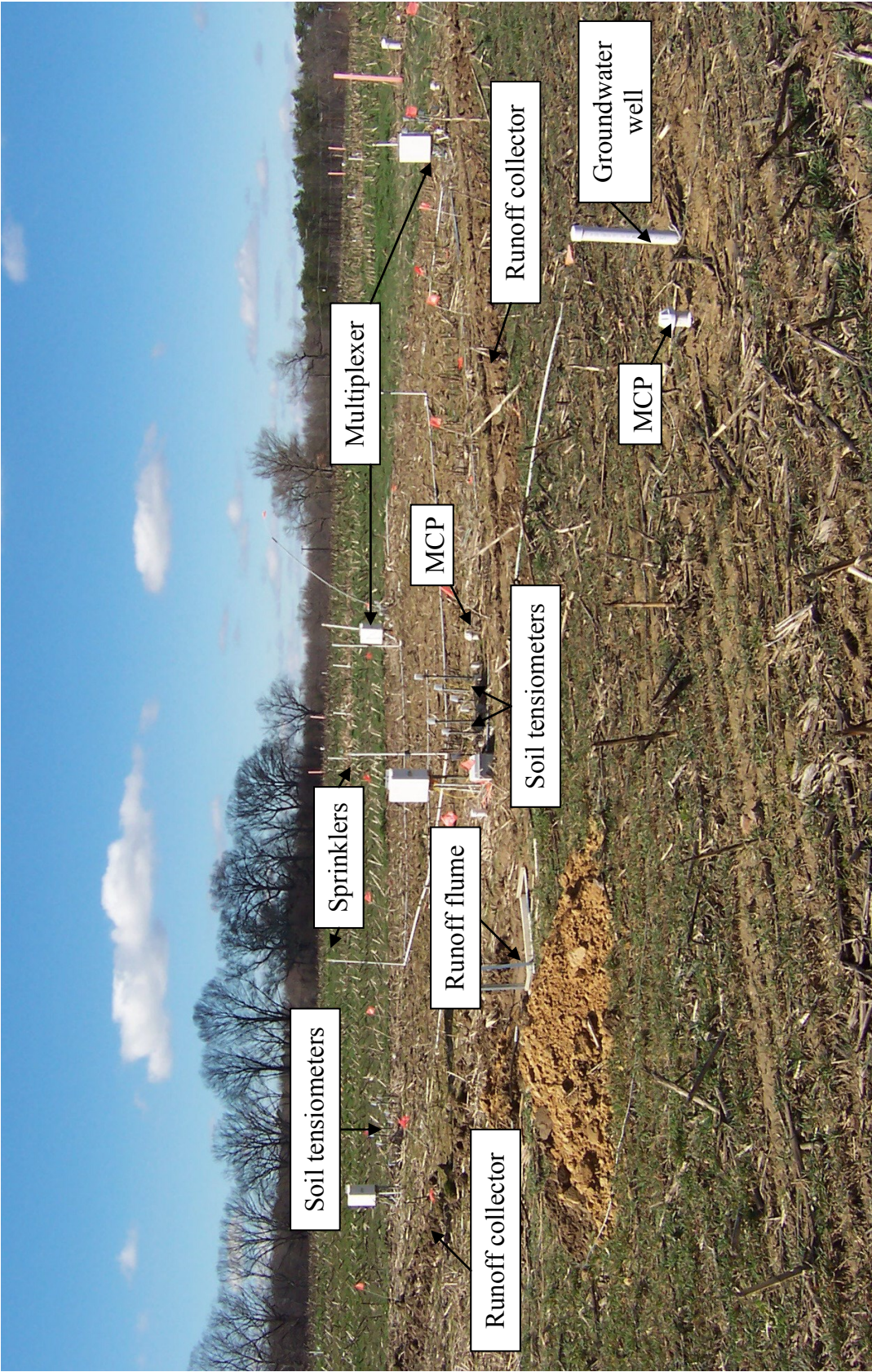
Figure 3-16   Setup of the lateral flow experiment.

showed high spatial variability in water content at different depths. The site was located 30 m from a nearby road, which did not interfere with the long-term ongoing studies at Field B.

Four plots (loc-1, loc-2, loc-3 and loc-4), each of 1 m2 and 10 m apart, were instrumented in the spring of 2006 to monitor soil water contents, pressure heads, and groundwater depths

 (Fig. 3-16). MCPs (EnviroSCAN, SENTEK Pty Ltd., South Australia) and soil tensiometers were installed at depths from 0.1 to 1.0 m at 0.1 m increments to monitor soil water contents and pressure heads, respectively. The MCP and tensiometers were connected to a CR-10X datalogger (Campbell Scientific, Inc., Logan, Utah) to collect data every 15 minutes. Two AM416 Relay Multiplexers (Campbell Scientific, Inc., Logan, Utah) controlled datalogger communication with 40 pressure transducers (PX26-005GV, Omega, Stanford, CT) installed in the tensiometers

(Fig. 3-17). Collected data were acquired from the Redwing 100 Airlink modem (Campbell Scientific, Inc., Logan, Utah) once a week. Five additional locations 7 m downgradient from the site, as well as three locations 14 m downgradient, were equipped with observation wells and MCPs for monitoring water contents and groundwater depths (Fig. 3-18). Groundwater depths were measured at each location using Cera-Diver (Van Essen Instruments, Delft, The Netherlands) sensors installed in wells at a depth of 2.0 m. Barometric pressures were recorded using a Baro-Diver sensor (Van Essen Instruments, Delft, the Netherlands). Volumetric water contents were measured periodically at locations loc-1 to loc-4 to correct MCP factory calibration.

A drainage line was constructed at a distance of 1m to the South and 1 m to the East of the plot boundary to collect runoff water. A ½' H-type flume was installed at the outlet of the drainage line and equipped with a 75 kHz Ultrasonic sensor connected to a Sigma 980 flow meter (Hach Company, Loveland, CO) for continuous water level measurements.

3.2.3.2 Basic Soil Properties at the Lateral Flow Experimental Site

Soil samples were taken from the soil surface to a depth of 2.0 m at each location during MCP and groundwater well installation in 2006. Soil texture was measured with the pipette method (Gee and Or, 2002) after dispersion with sodium pyrophosphate $Na_4P_2O_7$. Values of soil pH were measured at a solid to liquid ratio of 1:1 (Page et al., 1982). Organic carbon content was measured in the upper 0.6-m soil layer using the dry combustion method (Page et al., 1982).

Spatial variability in soil texture was found to be surprisingly high for the relatively small site. Soil texture was represented by six textural classes within the range from sandy loam to silty clay loam (Fig. 3-19). The particle size distribution generally did not change noticeably in the upper 1.3 m of the soil profile, where sand dominated over clay and silt. A distinct decrease in sand content and an increase in silt content can be seen below 1.3-m depth (Fig. 3-20), thus indicating a gradual transition in soil texture from sandy loam to loam and then silty clay loam. The horizontal variability in soil texture was of the same order of magnitude as the vertical variability. Coefficients of variation for sand and silt content increased with depth and were in the range from 0.06 to 0.92 and from 0.15 to 0.48, respectively. One could expect much variability in the soil water content at the studied site, since spatial variability in soil texture affects substantially the soil hydraulic properties.
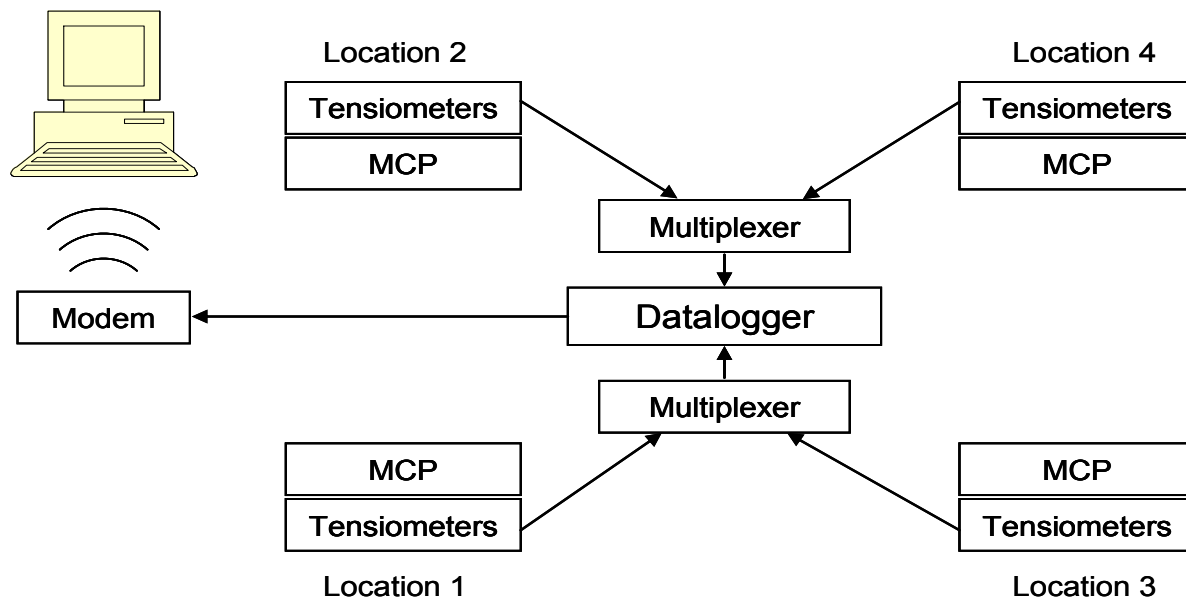
Figure 3-17 Schematic of the data acquisition and collection system.
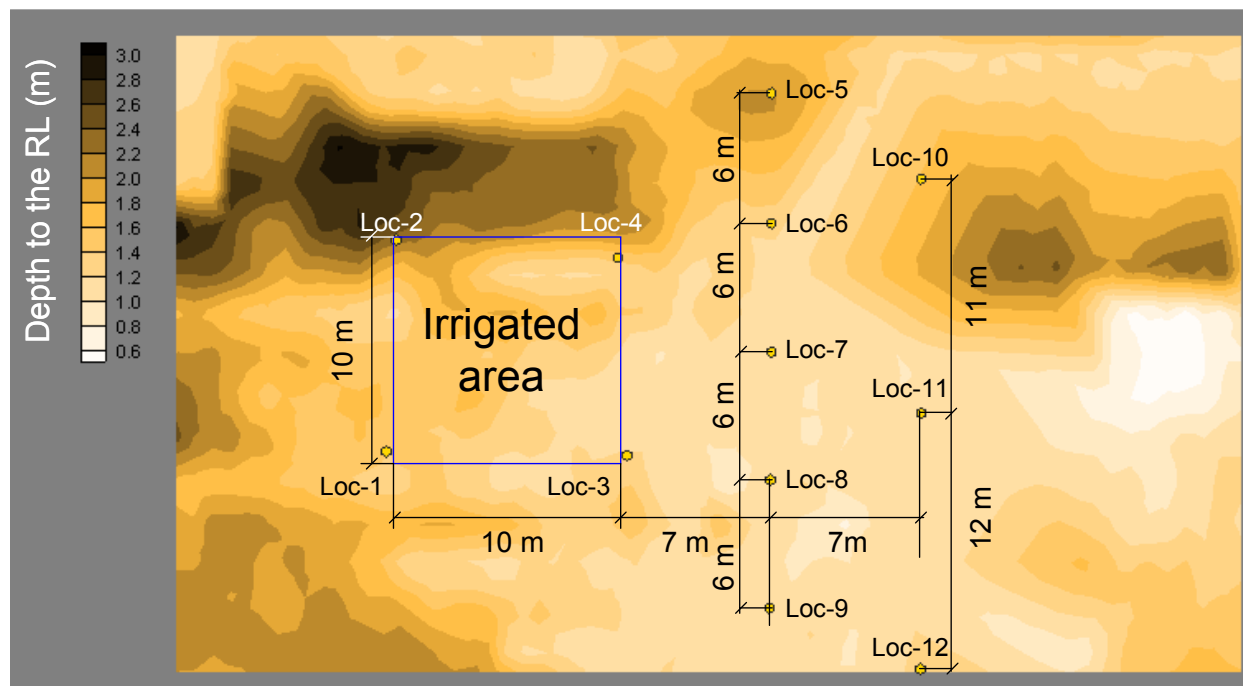
Figure 3-18 Locations of multi-sensor capacitance probes and groundwater wells (indicated by "Loc-" number) at the experimental plot.
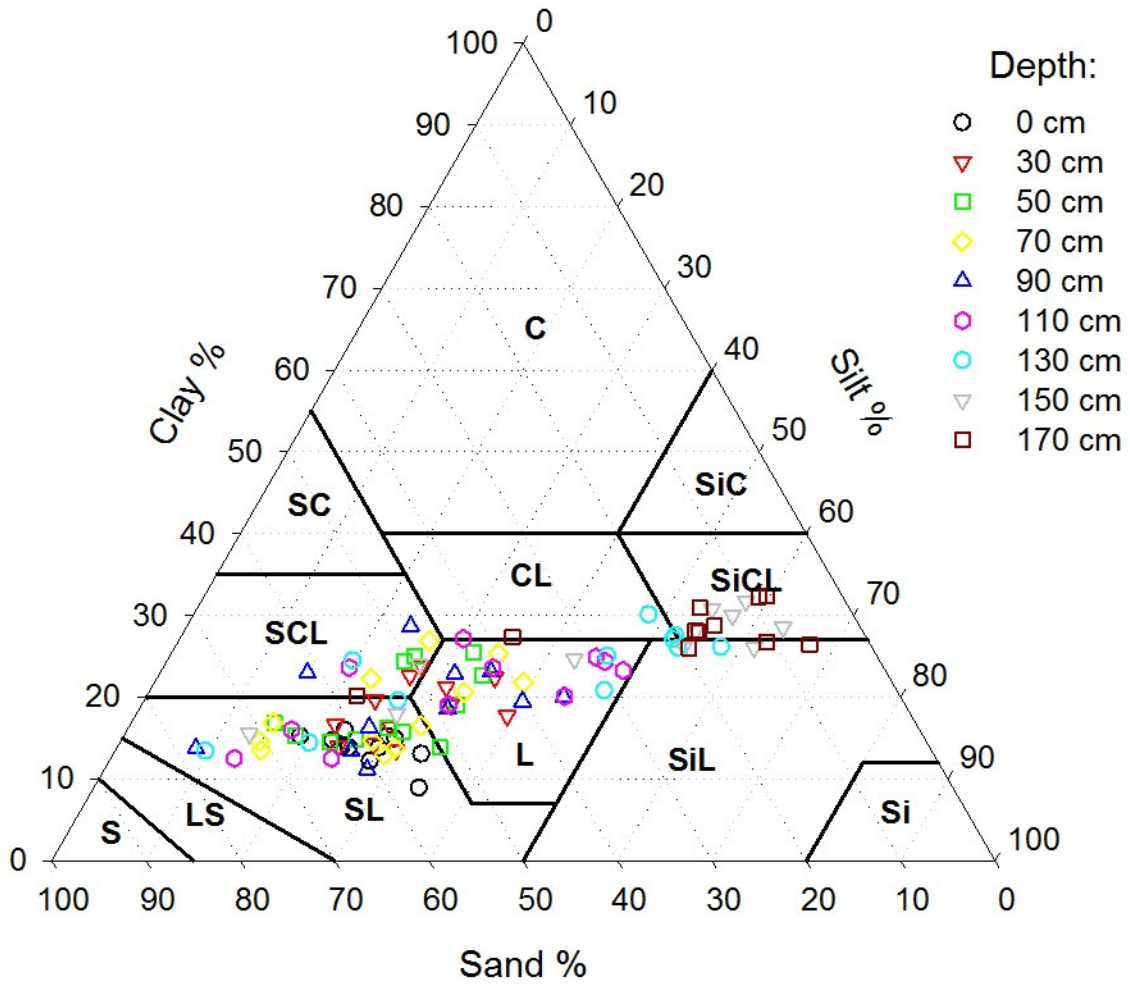
Figure 3-19 Soil texture measured at 10 depths in 12 locations.
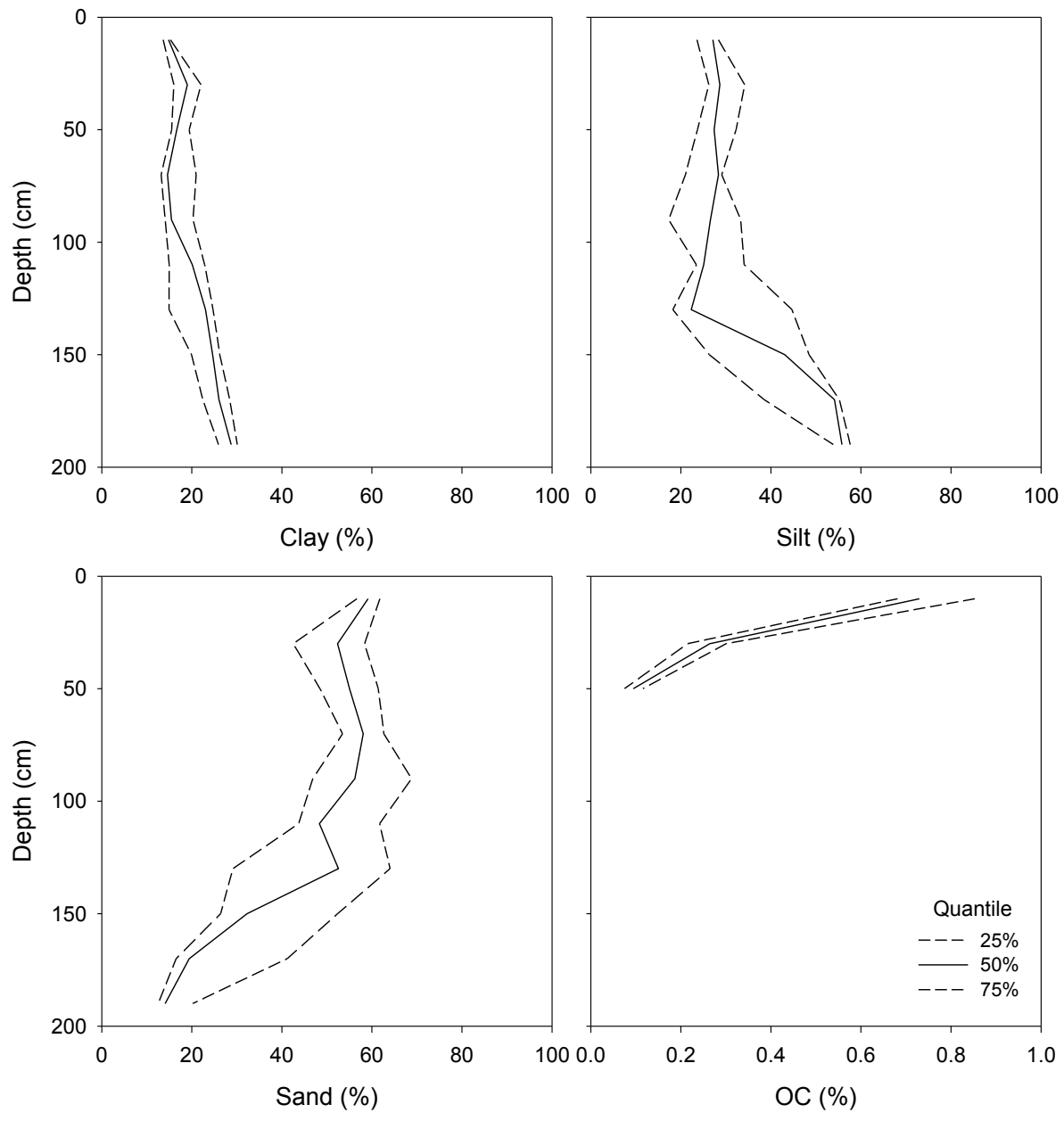
Figure 3-20 Measured clay, silt, sand and organic carbon distributions in the soil profiles.

The organic carbon content of the upper 0.6 m of the soil profile did not vary much at each depth; values decreased gradually from 0.4-1.2 % at 0.1 m, to 0.03-0.13% at 0.6 m depth. The soil pH was highest in the upper 0.2-m soil layer; average pH values gradually decreased with depth from 5.9 near the soil surface to 4.5 at 2.0-m depth (Fig. 3-20). Coefficients of variation for pH were within the range from 0.04 to 0.10, indicating less horizontal variability in pH as compared to the soil particle size distribution.

The ARS team also measured the soil bulk density on undisturbed samples taken from the upper 1.0-m soil layer at 0.1 m increments during MCP calibration at locations 1-4. The bulk density generally increased with depth from 1.3-1.6 to 1.6-1.9 g cm^{-3} in the upper 0.3 m of the soil profile, but did not change noticeably with depth in deeper layers (Fig. 3-21). Coefficients of variation for the bulk density were in the range of 0.05 to 0.09 at different depths, and did not correlate with average soil bulk density values and/or sampling depths.

Detailed soil texture measurements carried out in 2006 during the MCP and groundwater well installations at 12 locations were supplemented with data obtained in logs (wells w01-w52) installed in 2002 across field B. The soil texture was measured using the hydrometer method (Gee and Or, 2002) on soil samples taken from selected depths. Soil texture of the soil horizons represented 9 texture classes indicating high heterogeneity of soil properties (Table 3-2). In spite of the observed variability, four classes dominated soil texture. They were in the order of their frequency order: sandy loam (SL), loamy sand (LS), and silt loam (SiL) in layer 46-61cm; SL, LS, sand (S) and SiL in layer 107-122 cm; LS, SL and loam (L) in layers 168-183 cm and 229-244 cm; and LS, S and SL in layers 290-305 cm (Fig. 3-22).

Soil hydraulic properties were measured on undisturbed soil samples (5 cm ID, 5.1 cm height) taken from depths of 0 to 1.0 m in 0.1 m increments at locations 1, 2, 3, and 4 within the 10x10 m^2 experimental site. Soil samples were gradually saturated from the bottom. A constant head soil core method (Reynolds and Elrick, 2002) with a water head of 0.5 m was used to measure the saturated hydraulic conductivity of the undisturbed soil cores. The cores subsequently were placed into 1400 Tempe pressure cells (Soil Moisture Equipment Corp., Santa Barbara, CA) to measure soil water retention as described by Dane and Hopmans (2002). Results of the soil hydraulic properties measurements are presented in Figs. 3-23 and 3-24.

Measured saturated hydraulic conductivities (K_s) varied greatly vertically and laterally; they were mostly in the range from 0.003 to 27.1 cm d^{-1}. Generally smaller values were obtained at locations 1 and 3 compared to locations 2 and 4 (Fig. 3-23). Soil samples taken from depths of 20-50 cm at location 1 and from depths of 15-40 cm and below 80 cm at location 3 did not conduct water at 0.5 m pressure head. Low K_s values measured on 5-cm diameter samples presumably characterized the hydraulic conductivity of the soil matrix, while higher values could be measured on large undisturbed soil samples.

Soil water retention was found to vary with depths. Generally, the saturated water content decreased and the slope of the water retention curve increased with depth. The ARS study team did not observe high lateral variability in soil water retention, except for the 30-40 cm soil layer (Fig. 3-24), where the variability could be attributed to high variability in the bulk density (Fig. 3-21). Summarizing the basic soil properties, one could expect considerable variability in the
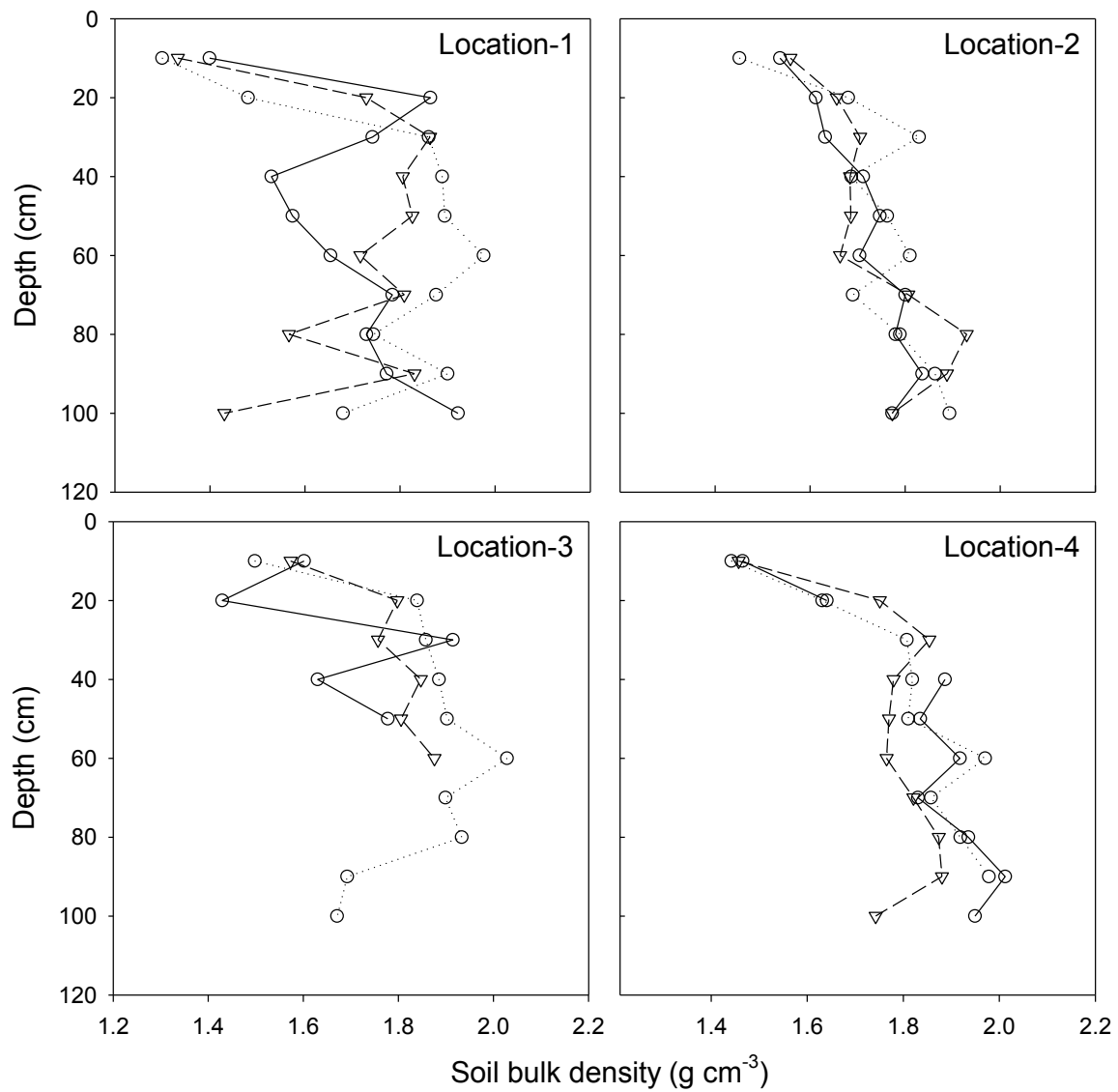
Figure 3-21 Measured soil bulk density distributions versus depth at locations 1 through 4.

Table 3-3   Soil texture measured in soil cores taken during installation of the groundwater wells at field B.

Well ID	Depth, cm					Well ID	Depth, cm				
	46-61	107-122	168-183	229-244	290-305		46-61	107-122	168-183	229-244	290-305
W01	SL	SiL	LS	SL	LS	W27	SL	SL			
W02	SL	SL	LS	SL	LS	W28	SiL	SL	LS	LS	LS
W03	LS	S	LS	LS	S	W29	SL	LS	SL	SL	SL
W04	LS	S	S	S	LS	W30	SL	SL	LS	SiL	
W05	SL	S	S	SL	SL	W31	SL	SL	L	SiL	
W06	SL	LS	SL	SL	LS	W32	SiL	SL	SiL	L	
W07	LS	S	LS	S	S	W33	SL	S	S	S	S
W08	LS	LS	LS	LS	LS	W34	SL	LS	S	LS	S
W09	LS	LS	LS	S		W35	LS	LS	LS	LS	
W10	SL	S	S	S	LS	W36	SiL	SiL	SL	LS	S
W11	L	LS	S	LS	S	W37	SL	L	SL		
W12	SL	SL	LS	LS		W38	SiL	SiL	SL	LS	SL
W13	SL	LS	LS	SL		W39	SL	SL	LS	SL	
W14	SL	SL	SL	LS		W40	LS	S	LS	LS	SL
W15	SL	SL	SL	SL	LS	W41	LS	SL	LS	SL	SL
W16	SL	SiL	SL	SL	SL	W42	SL	S	S	LS	S
W17	SL	LS	SL	LS	LS	W43	SL	SL	LS	SL	L
W18	SL	SL	LS	LS	LS	W44	LS	LS	L	SiL	
W19	SiL	SL	SL	LS		W45	SL	SL	SL	SL	L
W20	SiL	SL	S	LS	LS	W46	SL	LS	SL		
W21	SiL	SiL	LS	SL	SL	W47	CL	SiL	SL	LS	S
W22	SiL	SL	LS	LS	LS	W48	SiL	SiL	SiL	LS	LS
W23	SL	SiL	LS	LS	SL	W49	SiL	SiL	LS	S	S
W24	L	L	CL	CL	L	W50	S	LS	S	SL	LS
W25	SL	SiCL	CL	SCL		W51	LS	S	S	LS	SL
W26	L	SL	SiC	CL		W52	LS	S	S	S	S

Abbreviations used in the table:
S - sand
SL - sandy loam
LS - loamy sand
L - loam
SiL - silt loam
SiCL - silty clay loam
SiC - silty clay
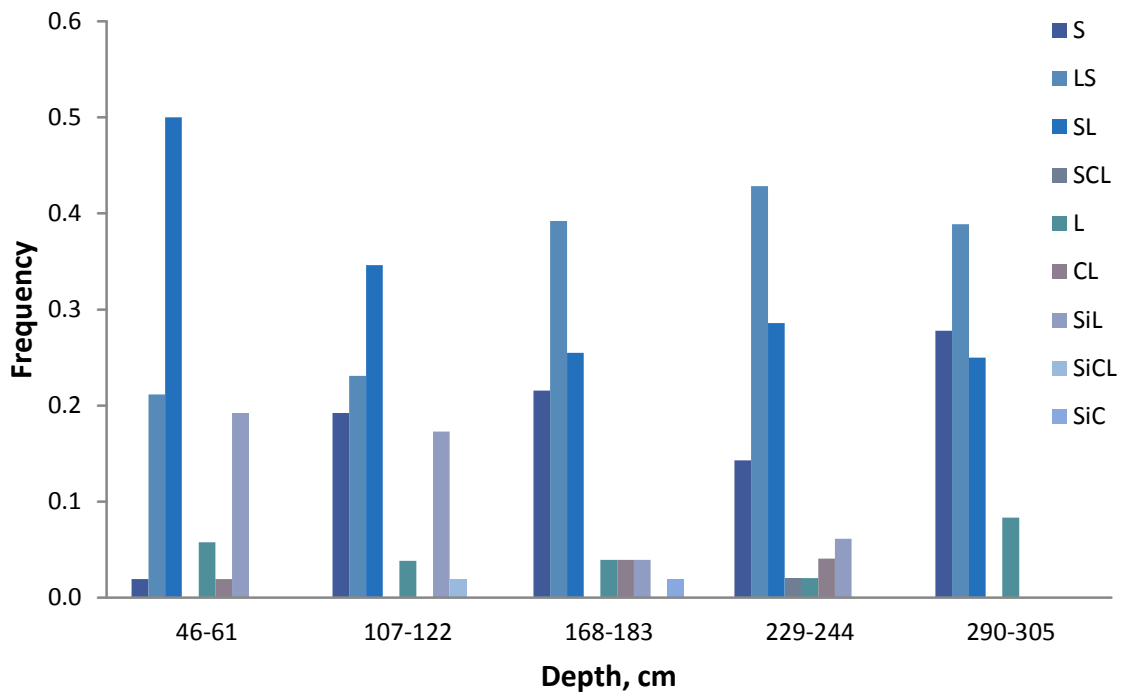
Figure 3-22 Soil texture measured in logs w01-w52 at field B.
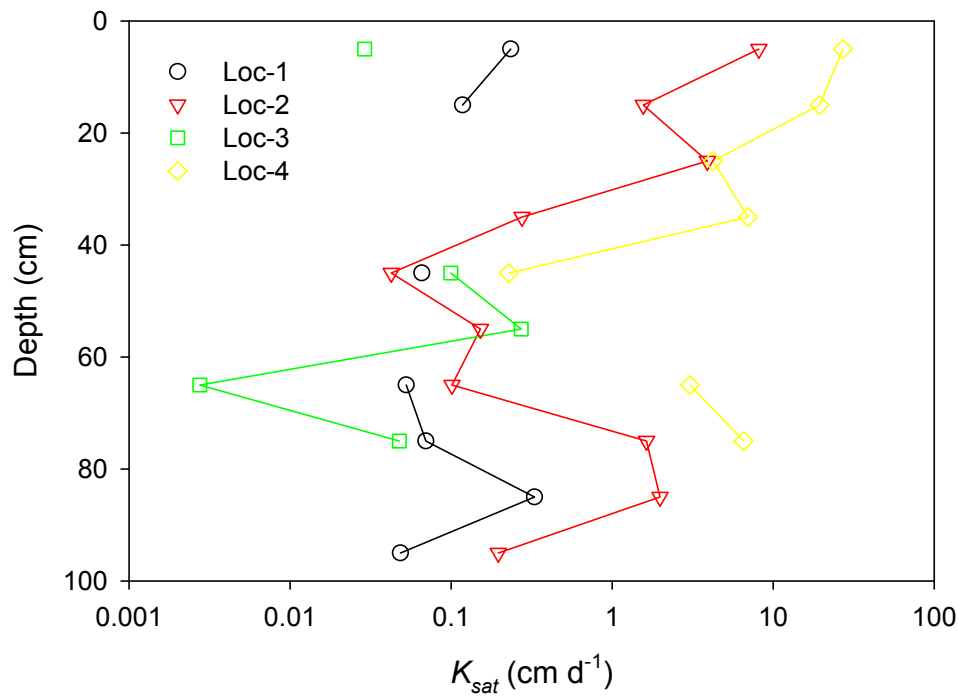
Figure 3-23 Laboratory measured saturated hydraulic conductivities, Ks.
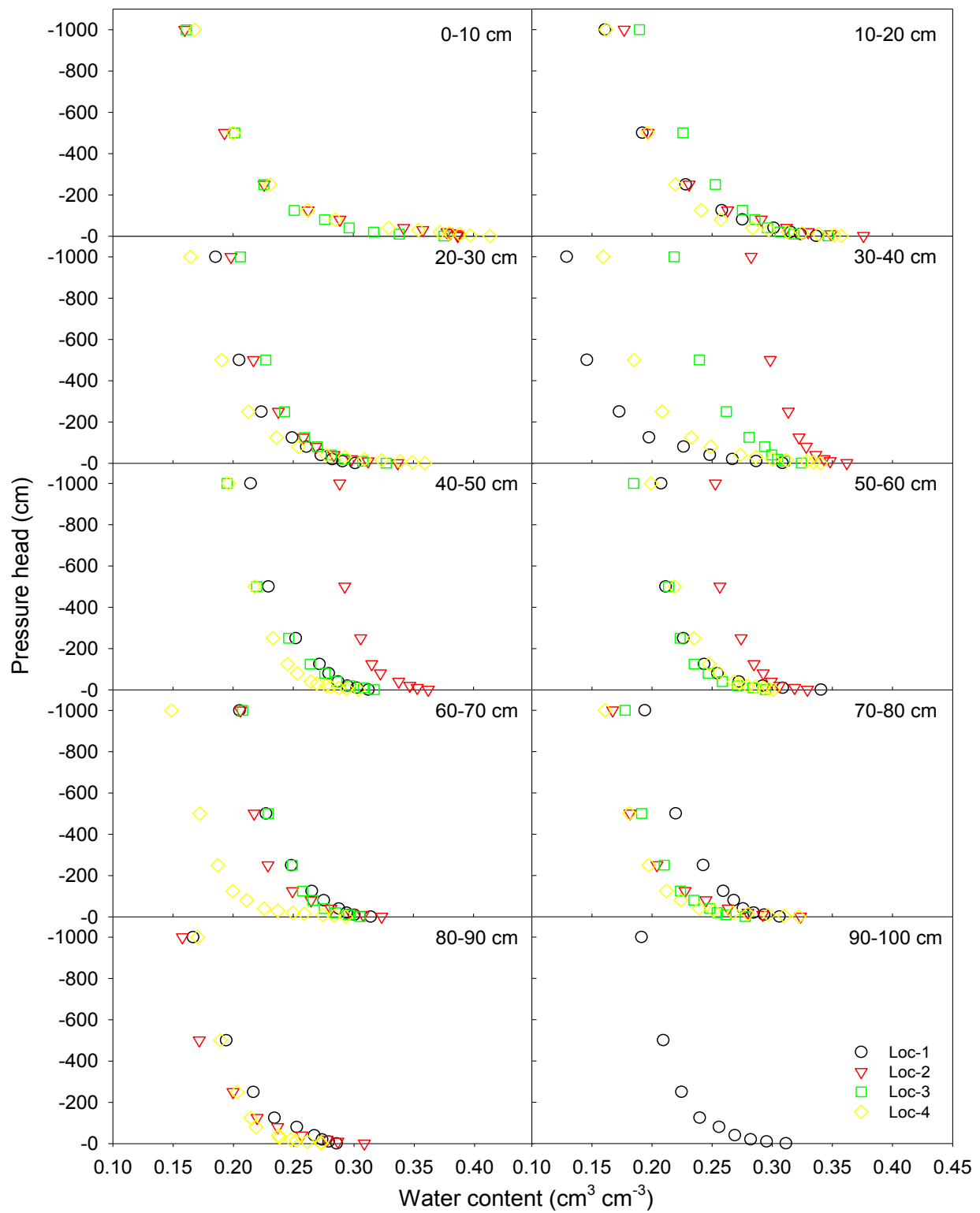
Figure 3-24 Laboratory-measured soil water retention curves.

spatial and temporal patterns of soil water content during the experiments, primarily because of high spatial variability in the saturated hydraulic conductivity.

3.2.3.3 Soil Moisture and Groundwater Monitoring

The monitoring equipment installed in 2006 at the experimental site was tested during July 1, 2006 – March 1, 2008. Soil water contents and pressure heads were monitored at the experimental site. A weather station located about 80 m from the site collected meteorological data at short time intervals (10 minutes). These data included: soil temperature, soil heat flux, air temperature, relative humidity, 3-D wind speed profile, rainfall, long and short wave solar radiation, net solar radiation, saturation and actual vapor pressure, evapotranspiration, and $CO2$ fluxes.Daily evapotranspiration (ET) rates were computed using the Penman-Monteith method as documented by FAO (Allen et al., 1998). The daily ET and rainfall rates were integrated over time to obtain cumulative ET and precipitation values as a function of time from the beginning of the observation period.

Monthly and cumulative ET and rainfall rates for the monitoring period July 1, 2006 - November 1, 2007 are shown in Fig. 3-25. Daily rainfall and ET values were in the range from 0 to 6.7 cm and from $5 \cdot 10^{-3}$ to 0.4 cm respectively. The total amount of precipitation was about two times larger than the total ET during the observation period. The rainfall and ET distributions were not uniform over the year. Relatively high rates of precipitation were observed in the months of April, September, October and November, and low rates in June and July. Contrary to precipitation, ET values were higher in the summer and lower in the winter.

Tensiometer and MCP readings were converted to volumetric water contents and pressure heads. The MCP factory calibrations were corrected separately for topsoil and subsoil to improve the accuracy of the water content measurements (see Appendix B). Tensiometer readings exhibited noise, which was removed using a multilevel 1-D wavelet decomposition with a Daubechies filter (db4). The ARS study team used the Wavelet toolbox of the MATLAB 6.5 Software for this purpose. Depending upon data quality, the multilevel parameter in the wavelet decomposition was fixed at values between 4 and 6.

The non-uniform distribution of precipitation and ET affected the dynamics of the water content and the pressure head in the vadose zone during the experiments. Periods of deep soil drying followed rainfall events when water contents approached saturation. Extremely low water contents were observed in September 2006, and in August and October 2007, at practically all depths (Fig. 3-26). The low water contents corresponded to pressure heads close to or below the ceramic's air entry value. For this reason, tensiometers installed at depths less than 0.8 m did not work properly in August and September 2007.

Temporal changes in the soil water contents varied with depth. Water contents in the topsoil (0-0.4 m) generally increased rapidly during and after the rainfall events, while relatively small changes occurred in the subsoil (0.4-1.0 m). Water contents in deeper soil layers were less sensitive to precipitation and evapotranspiration as compared to the upper layers. This was also true for the pressure head. There always was a delay in response of the deeper soil layers to changes in the flux at the soil surface.
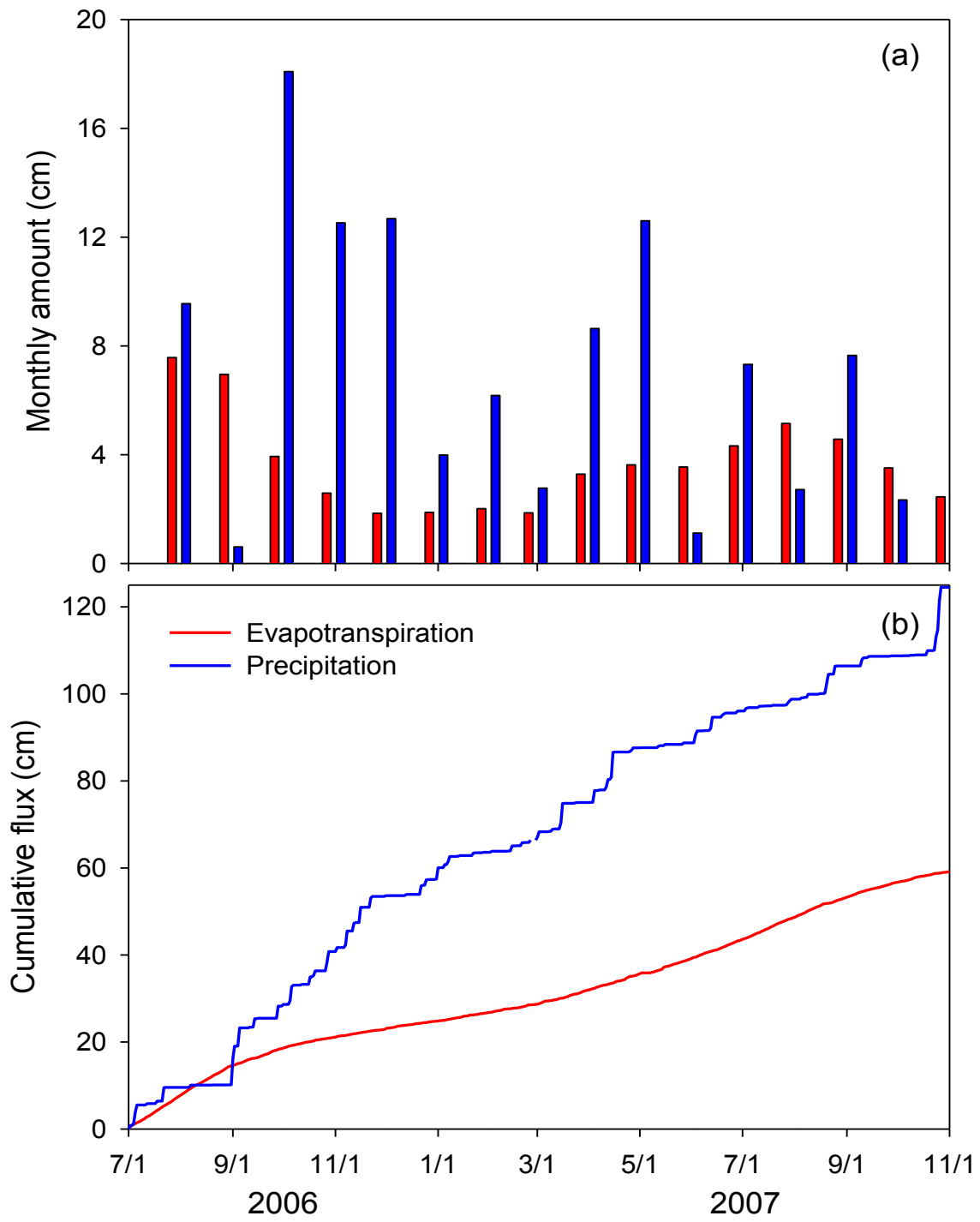
Figure 3-25 Monthly (a) and cumulative values (b) of evapotranspiration and precipitation measured at the experimental site during the monitoring period (during July 2006 – September 2007).
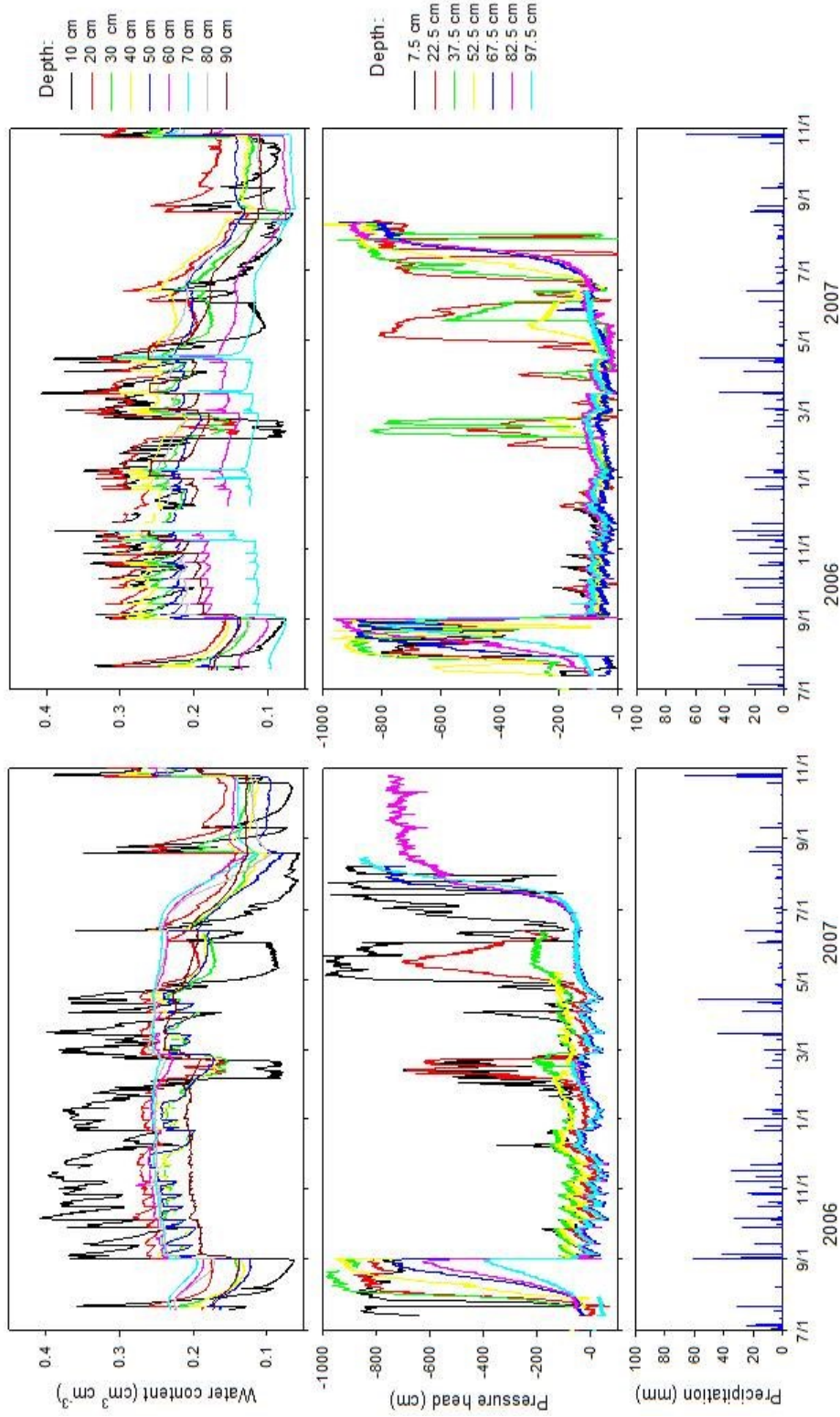
Figure 3-26 Examples of soil water content and pressure head time series measured at locations 1 and 2 in July 2006 September 2007.

Water contents at similar depths varied considerably among the four locations. The differences were most pronounced at depths between 20 and 50 cm, where water contents were consistently higher at locations 2 and 4 as compared to locations 1 and 3. Consistently low water contents were observed at a depth of 70 cm at location 2 during the entire monitoring period.

Groundwater depths (GWDs) were monitored at 12 locations starting November 1, 2007. Observed values varied between 0.0 and 200 cm during the monitoring period (Fig. 3-27). The GWD values were consistently smaller at locations 1, 3, 6, 9 and 11, and greater at locations 7, 8, 10 and 12, particularly during continuous drying periods. Maximum seasonal variability in the GWD was observed at location 5, 8 and 12, while variability was minimal at locations 1 and 3. The GWD decreased rapidly after intensive rainfalls at all locations. Changes in the GWD were more pronounced at locations 7, 8, 10 and 12 as compared to other locations during drying events.

Spatial differences in soil water content dynamics were reflected by different field soil water retention curves, obtained by plotting field-measured pressure heads versus water contents for 6 depths at 4 locations (Fig. 3-26). The ARS study team shows only data measured during continuous soil drying (drainage water retention curves). The data in Fig. 3-28 in general resembled the laboratory-measured water retention curves (Fig. 3-24). Similarly as the laboratory data, saturated water contents gradually decreased with depth, whereas the slope of the water retention curve increased with depth. The slopes of the laboratory-measured retention curves were steeper compared to the field-measured data, resulting in lower water contents at the same pressure heads for the field measurements. The field-measured water retention curves at similar depths differed among the 4 locations. The maximum differences occurred at depths between 60 and 75 cm between locations 2 and 3 and locations 1 and 4, with the water retention curves of locations 2 and 3 shifted to lower water contents (Fig. 3-28).

Overall, 44-month intensive equipment testing showed that in spite of some technical issues with MCP sensors and tensiometers the installed equipment appeared to be reliable and appropriate for the subsurface transport experiment outlined in part 3.1 of this chapter. Results of the soil moisture monitoring supported soil hydraulic and soil texture data and confirmed that the studied area represents a high heterogeneous layered system.

3.2.3.4 Fine-Scale Ground Penetration Radar Survey

A fine-scale GPR survey was conducted in April 2006 to obtain detailed information about the stratification of soil hydrological layers at the field site. A polygon of 16,617 m^2 in the South-Eastern part of Field B (Fig. 3-17) was surveyed with a subsurface interface radar system-2 (Geophysical Survey System, Inc., North Salem, NH). Ground-penetrating radar data were acquired for the entire site along parallel east-west transects 2-m apart. The RADAN (Geophysical Survey Systems, Inc) software was used to process the acquired information. A total of 60 scanned images were processed. A horizontal scale adjustment was used first to correct horizontal distances based on markers installed every 20 m along the scanned transects. Subsequently the diffractions were removed and dipping layers were corrected using the 2-D variable velocity migration method. This type of correction was necessary for images in which the deeper objects were obscured by numerous shallower objects that appeared as constructively

Figure 3-27 Groundwater depths measured at the experimental site during the period November 2007 - March 2008.
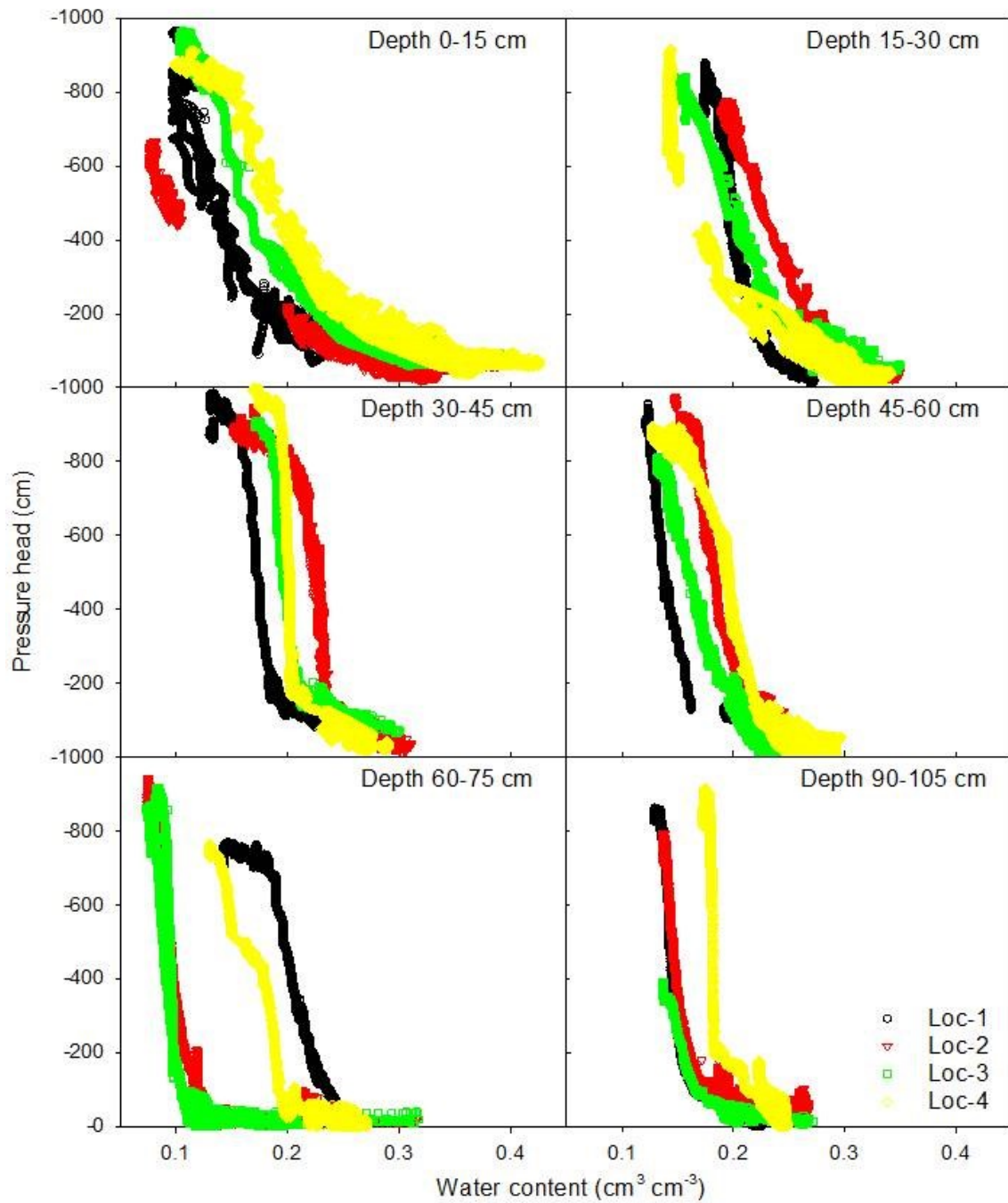
Figure 3-28 Plots of measured soil water pressure heads versus volumetric water contents at locations 1-4.
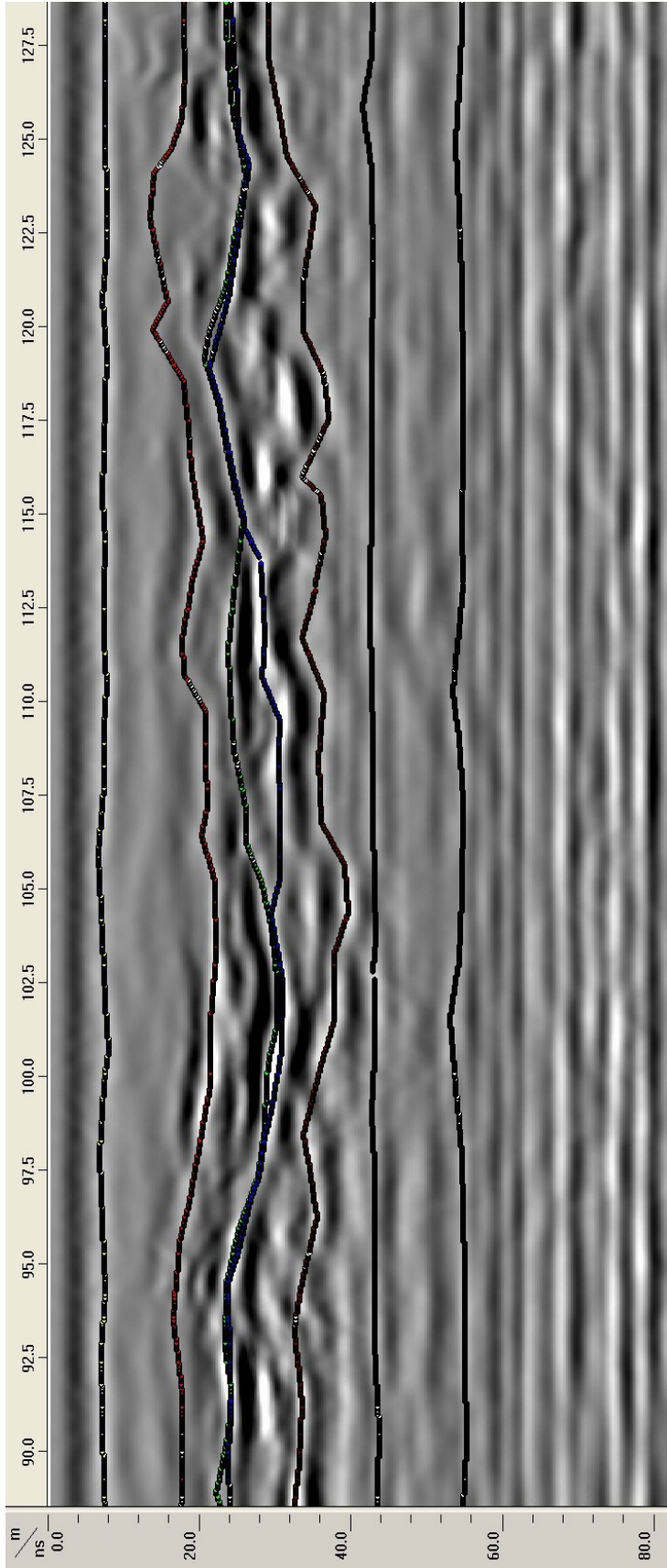
Figure 3-29 Example of ground-penetrating radar images used to delineate hydrological soil layers.

interfering hyperbolic reflectors. Zero depths on the images were corrected next to match the soil surface using the vertical scale adjustments, after which the reflective layers were digitized using the EZ tracker option of the RADAN software. Seven digitized layers from each GPR image were saved in ASCI format at 25-cm horizontal resolution. Depths to the reflective layers were expressed as two-way signal travel times in the digitized images (Fig. 3-29).

To define true depths to the reflective layers the ARS study team computed the GPR signal travel time т as a function of the travel depth z as follows:

$$\tau = 2 \int_0^{300} dz \big/ v(z) \tag{3-1}$$

where v is the GPR wave velocity in soil, which depends on the soil dielectric constant E.

Topp et al. (1980) showed that at frequencies between 1 MHz and 1 GHz values of E for sandy loam, clay and clay loam soils can be estimated from the volumetric soil water content θ as:

$$\mathbf{E} = 3.03 + 9.3\theta + 146.0\theta^2 - 76.7\theta^3 \tag{3-2}$$

The wave velocity U is then:

$$U(Z) = c/\sqrt{\mathbf{E}} \tag{3-3}$$

where c is the speed of light (0.3 m ns^{-1}).
Seven soil water content profiles measured at locations BH3, BH4, BL3, BL4, BM2, BM3 and BM4 (Fig. 3-8) during the GPR survey were used to compute profile distributions of the soil dielectric constant. Measured θ values varied between 0.167 and 0.347 at different depths and locations, and tended to decrease with depth (Fig. 3-30a). Variability in the water content translated into variability in the soil dielectric constant. Values of E also decreased with depth, ranging from 8.3 to 20.6 (Fig. 3-30b). The variability in soil water content and associated variability in the soil dielectric constant did not produce large differences in the signal travel time τ (Fig. 3-30b). The variability in τ increased with depth, but remained less than 8 ns at a depth of 200 cm. Although soil water contents were not measured at other locations during the GPR survey, the ARS study team assumed that the water content variability measured at the seven selected locations represented reasonably well the spatial variability in θ across field B. Since observed variability did not affect much the travel time τ, the average relationship $\tau(Z)$ was used.

The data for seven digitized layers were converted into depths to the soil horizons. To do this the ARS study team assumed that the vertical variability in soil texture affected the soil hydraulic properties and produced abrupt vertical water content changes. The layers between two consecutive reflective lines hence can be represented by soil horizons having specific soil hydraulic properties. To compute boundaries of these horizons the ARS study team interpolated digitized data using the inverse of the average relationship $\tau(Z)$. Results of the interpolation for the first six soil layers are shown in Fig. 3-31. The thickness of the first soil layer varied from 28 to 50 cm. This depth corresponded to the plough depth. The deepest layer detected by the GPR survey at this frequency was in the depth range from 211 to 243 cm.

3.2.3.5 Tracer Experiment Experimental Setup.

The lateral flow experiment was set up as shown in Figs. 3-16 – 3-18, and Fig 3-32. Four sprinkler nozzles (Olson Red Mini Rotor Style Drip Emitter, Olson Irrigation Systems, Santee, CA) were installed at heights of 1 m in the corners of the irrigation plot (Fig. 3.22a).

One additional sprinkler nozzle (Olson Green Mini Rotor Style Drip Emitter) was installed at a height of 10 cm above the soil surface in the center of the plot. The system was designed to provide continuous irrigation at a rate of 9.6 cm d^{-1}, with a uniformity coefficient of 0.86, to the 13x14 m study area. The sprinklers were connected to a water hose equipped with a DLJSJ50 water meter (Daniel L. Jerman Co., Hackensack, NJ) to measure the amount of applied water. Two 2.1-ton water tanks were installed for the tracer application (Fig. 3-22b).

The observation wells at locations 1-12 were equipped with samplers to collect water from 3 different depths (Fig. 3-32c). The samplers consisted of a 16-mm plastic OD tubes containing 4 rubber packers installed into the wells. The packers separated three 30-cm long sections inside of the wells. The 2-mm ID plastic tubing connected the center of each section with the sampler outlet. Water samples were taken from each section with a 50 ml syringe. The sample volume typically ranged between 10 and 14 ml. The samplers were installed so that the samples were taken at depths of 1.1 m at locations 1-4, and depths of 1.05, 1.35 and 1.65 m at locations 5 - 12. A windshield (160 cm height) was constructed around locations 1- 4 to minimize wind impacts on the irrigation uniformity (Fig. 3-22d).

A total of 42 kg KCl was dissolved in two 2.1-ton tanks with 4050 L of tap water to provide a Cl- concentration of 4940 ppm. The KCl solution was applied to the experimental site using the irrigation system. The application started at 10:30 am and ended at 5:05 pm on March 25, 2008. A total of 3861 liters was applied to the site within 6.6 hr. The irrigation system was reconnected to the tap water hose immediately after the application.

The irrigation was interrupted at 10:30 am on March 27, 2008, because of runoff at the site. Runoff losses were 39% of the irrigation water. The irrigation regime was adjusted to minimize runoff losses and to provide a nearly steady-state infiltration rate into the soil. Starting March 28, the site was irrigated twice a day, from 6 to 8 am and from 4 to 6 pm. Irrigation times were controlled by an irrigation timer and set such that well water could be sampled twice a day for Cl-. Runoff amounts from the experimental site were recorded continuously. Irrigation was interrupted only after intensive rainfalls. The irrigation regime was corrected one more time on June 1 because of an increase in evapotranspiration that caused a decrease in the groundwater levels at locations 1-4. New times were set up for three irrigations each day: from 5 am to 7 am, from 1 pm to 3 pm, and from 9 pm to 11 pm. This schedule was followed until the end of experiment.
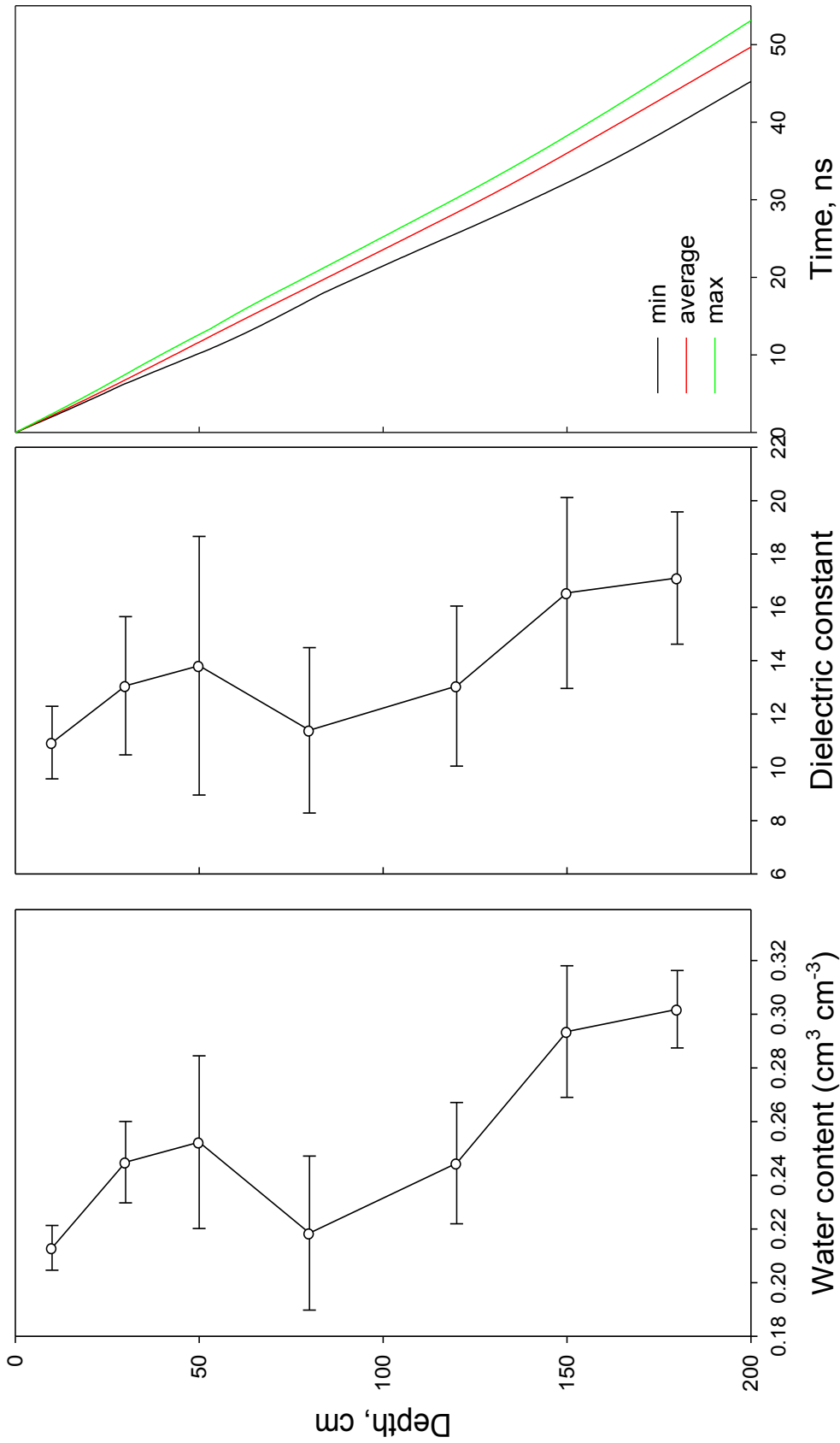
Figure 3-30 Plots of (a) the vertical water content distribution measured at locations BH3, BH4, BL3, BL4, BM2, BM3, and BM4 during the GRP survey, (b) the dielectric constant computed from average water contents among locations, and (c) the minimum, average and maximum time of GPR wave propagation computed from the soil dielectric constant profiles at seven locations.
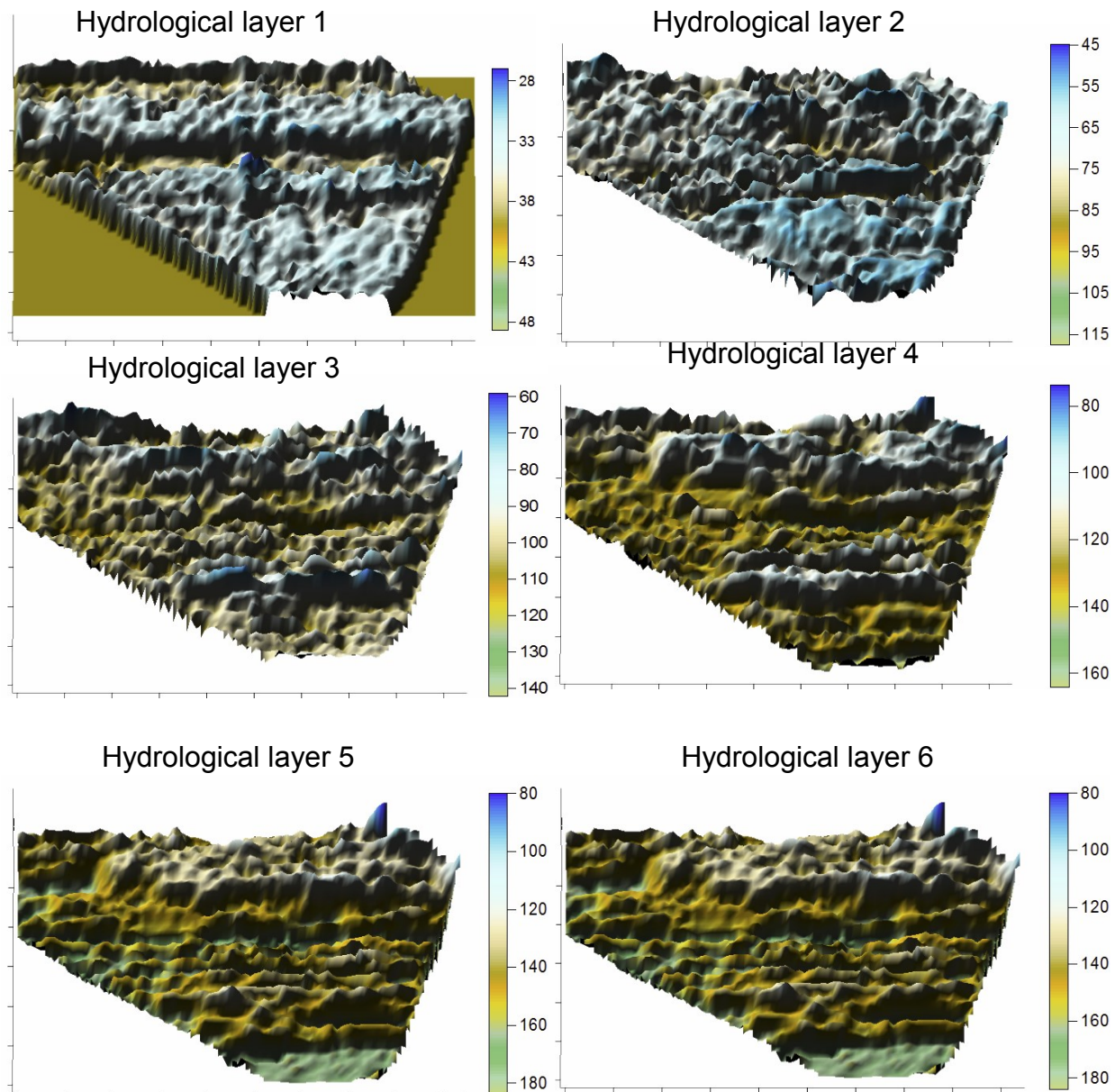
Figure 3-31 Depths to six out of seven soil hydrological layers delineated using the fine-scale GPR survey at the experimental part of field B in 2006. The color scale in the legend indicates depths in cm.
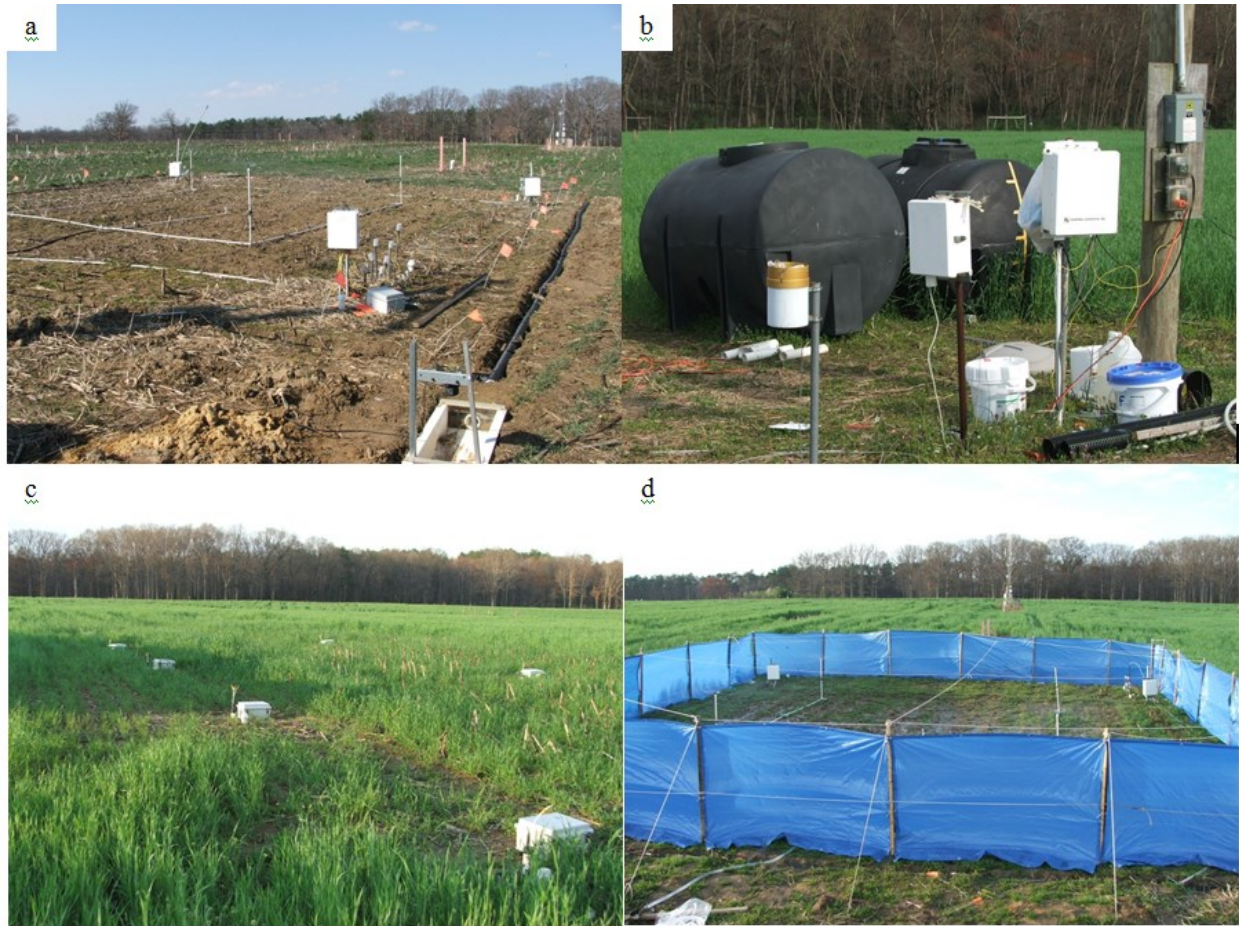
Figure 3-32 Plots of (a) the irrigated plot equipped with tensiometers, MCPs and a flume for runoff collection, (b) the tracer application tanks, (c) the observation wells with the tracer sampling outlets, and (d) the windshield around the tracer application area.

Water samples were taken from three depths at 12 locations one hour prior to the KCl application. Samples were taken twice a day within the first 40 days after application, and then once a day (at 5 pm) until the end of the experiment. Within 24 hours of their collection in the field, the samples were analyzed for Cl- by ion chromatography as described in chapter 3.2.4.1. Irrigation, while rainfall water was periodically analyzed also for Cl-. Soil water contents and groundwater depths were monitored using MCPs and Cera-Divers at 12 locations as described in chapter 3.2.3.3. The study was terminated when the groundwater levels decreased below the water sampling depths at most locations.

Cumulative irrigation, precipitation, evapotranspiration and runoff data are shown in Fig. 3-33. The daily irrigation rate was 9.5 cm d^{-1} during the first two days, 1.5 cm d^{-1} during the next 64 days, and 1.8 cm d^{-1} during the following 65 days. The first portions of runoff usually approached the flume 30-40 min after irrigation started. Runoff amounts were found to be in the range from 21 to 38% of the irrigation volumes. A total of 59 rainfall events were recorded during the 132-day experiment. The mean daily precipitation depth was 3.0 mm, which did not affect the soil water contents. Several events, however, did affect the water contents, the groundwater levels, as well as Cl- concentrations in the wells. Rainfall events on April 22-23 (54.6 mm), on May 10-14 (159 mm), on June 5-6 (43.4 mm), and on July 25 (50.0 mm) exceeded the daily irrigation rates considerably, and as such affected the soil water regime as shown below. The average ET values were 0.5, 0.8, 1.0 and 1.1 mm d^{-1} in April, May, June and July, respectively, which did not affect the soil water content dynamics.

Figure 3-34 illustrates the water table dynamics. Groundwater levels remained relatively high at locations 1-4 during the lateral flow experiment (Fig 3-34a). Minimum groundwater depths (GWDs) were observed at location 3, and the maximum levels at location 2 where they exhibited severe oscillations. Severe oscillations were also observed at location 4, but the phreatic surface there was closer to the soil surface as compared to location 2. Peaks in the GWD at locations 2 and 4 generally corresponded to the more intensive rainfall events. The GWD at location 1 was affected less by rainfall as compared to locations 2 and 4. Observed differences in the water content and GWD dynamics at locations 1 through 4 implied different hydrological regimes in different parts of the irrigated site. The same is obviously true for locations 5 through 12. However, GWD differences between those locations remained between 5 cm after rainfall events, and 119 cm after continuous soil drying. These differences were largest between the wells at adjacent locations 11 and 12. Considerable temporal variably in the GWD was also observed at locations 5 through 12. Average GWDs varied from 5 to 161 cm during the lateral flow experiment. Spatial and temporal variability in the GWD created unsteady subsurface flow from the irrigated plot toward locations 5-9 and locations 10-12.

Plots of the soil water content during the experiment are shown in Fig 3-35. Soil water contents before the experiment were in the range from 0.148 to 0.403 cm3 cm-3 and gradually decreased with depth. Three days after the experiment started, water content approached maximum values at practically all depths at locations 1 and 3, and below 25 cm depth at location 4, while remaining relatively constant until the end of the experiment. The water contents oscillated noticeably in the upper 20 cm at location 4 and at all depths at location 2.
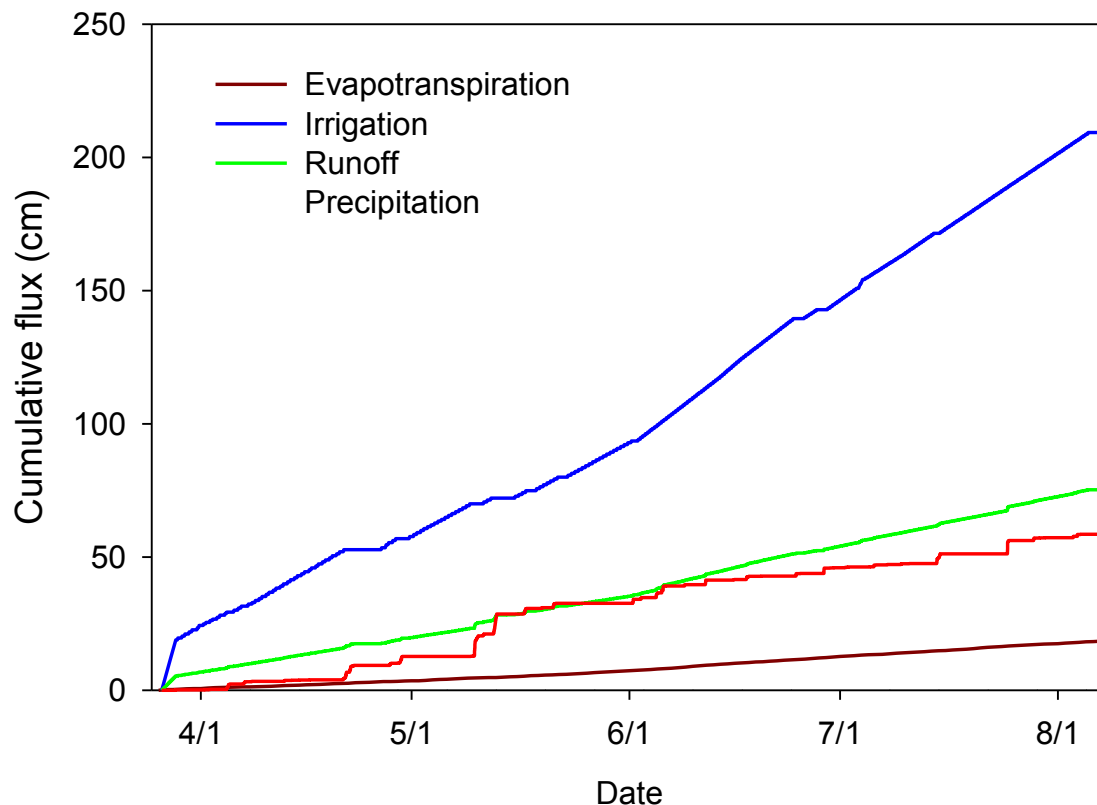
Figure 3-33 Cumulative water fluxes measured during the lateral flow experiment.

Considerable water content oscillations were observed also at non-irrigated locations 5 - 12 in the upper 40-cm soil layer. Sharp increases and gradual decreases were observed in this soil layer immediately after the rainfall events (Fig. 3-35). Water contents at depths of 50 and 60 cm showed different trends at these locations. Sharp increases were followed by sharp decreases at locations 7-9, whereas water contents at locations 11 and 12 did not change noticeably with time. Differences in the soil moisture dynamics were caused by different soil texture and changes in the groundwater depth at these locations. Soil texture at depths of 50-60 cm was sandy loam at locations 7-9, and loam at locations 11-12. Water contents in sandy loam soils near saturation tend to decrease faster with an increase in the absolute value of the pressure head, compared with loamy soils. Rapid changes in the GWD, as well as transitions from saturation to unsaturated conditions and vice-versa, likely contributed further to the abrupt water content changes at locations 7-9.

Differences in soil properties, hydraulic regimes and unsteady subsurface flow at the 12 locations affected Cl- transport. Cl- concentrations in groundwater prior to irrigation varied within the range from 6 to 24 ppm, but did not differ significantly between the different depths and locations. Because of small spatial variability in the initial Cl- concentration at the different locations, the ARS study team did not expect any effect of the initial Cl- distribution on the results of the lateral flow experiment. Cl- concentrations in the irrigated water were 10.2±3.0 ppm, whereas rainfall Cl- concentrations (0.6±0.6 ppm) were one order of magnitude less than the irrigation water and the groundwater concentrations. Hence, the effect of rainfall on the groundwater concentrations could be quite pronounced, especially after intensive rain events.

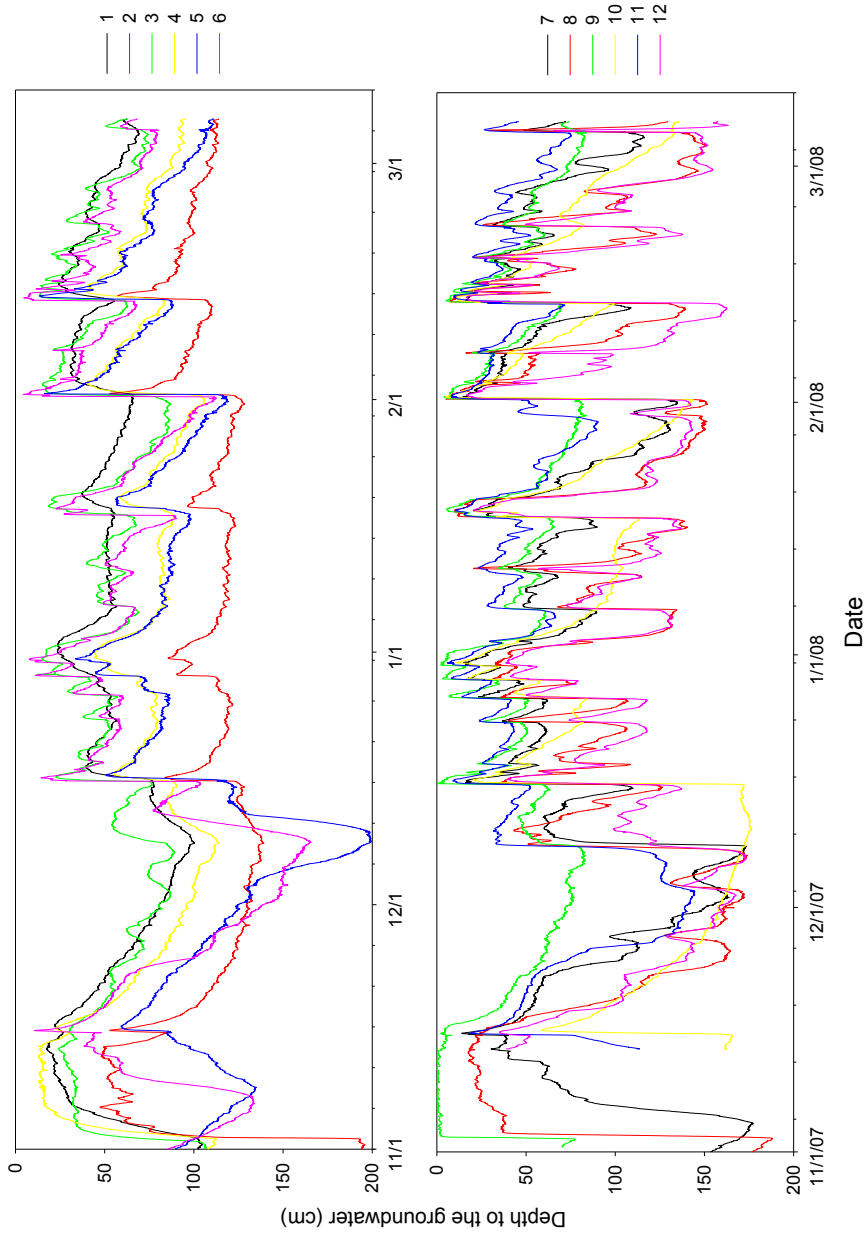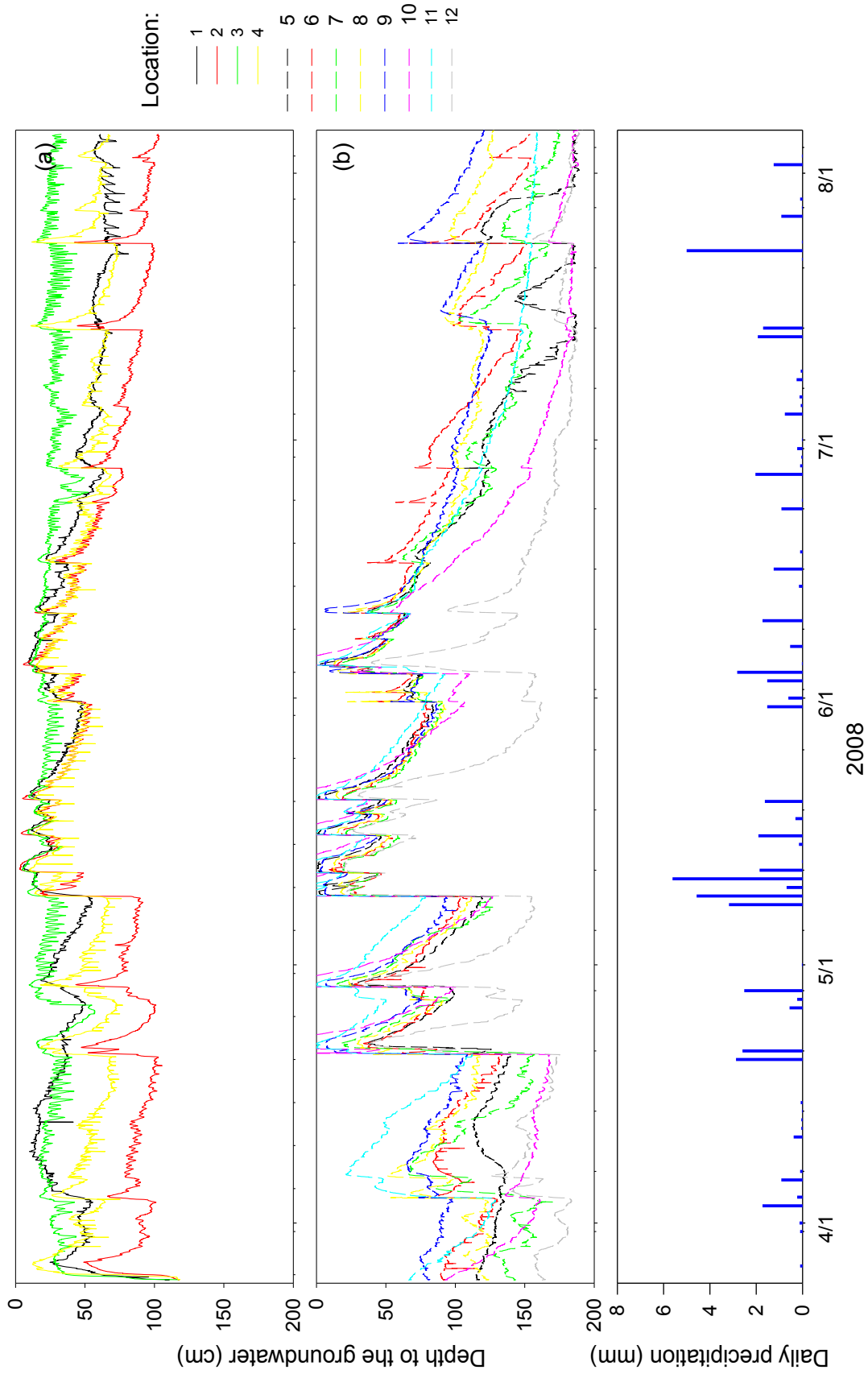**Figure 3-34   Groundwater depths measured at the experimental site during the period November 2007 – March 2008.**
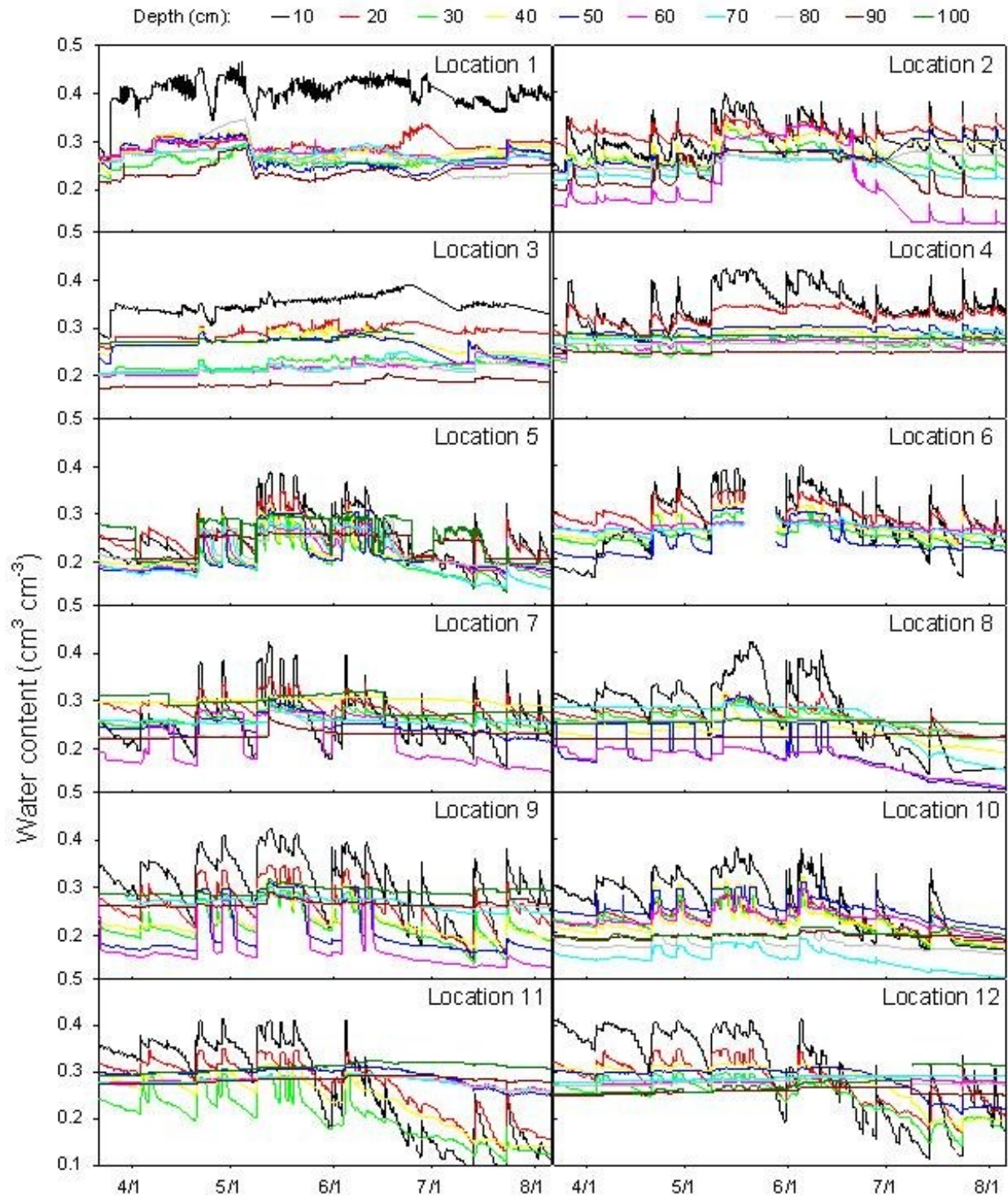
Figure 3-35 Soil water content measured at 12 locations during the lateral flow experiment.
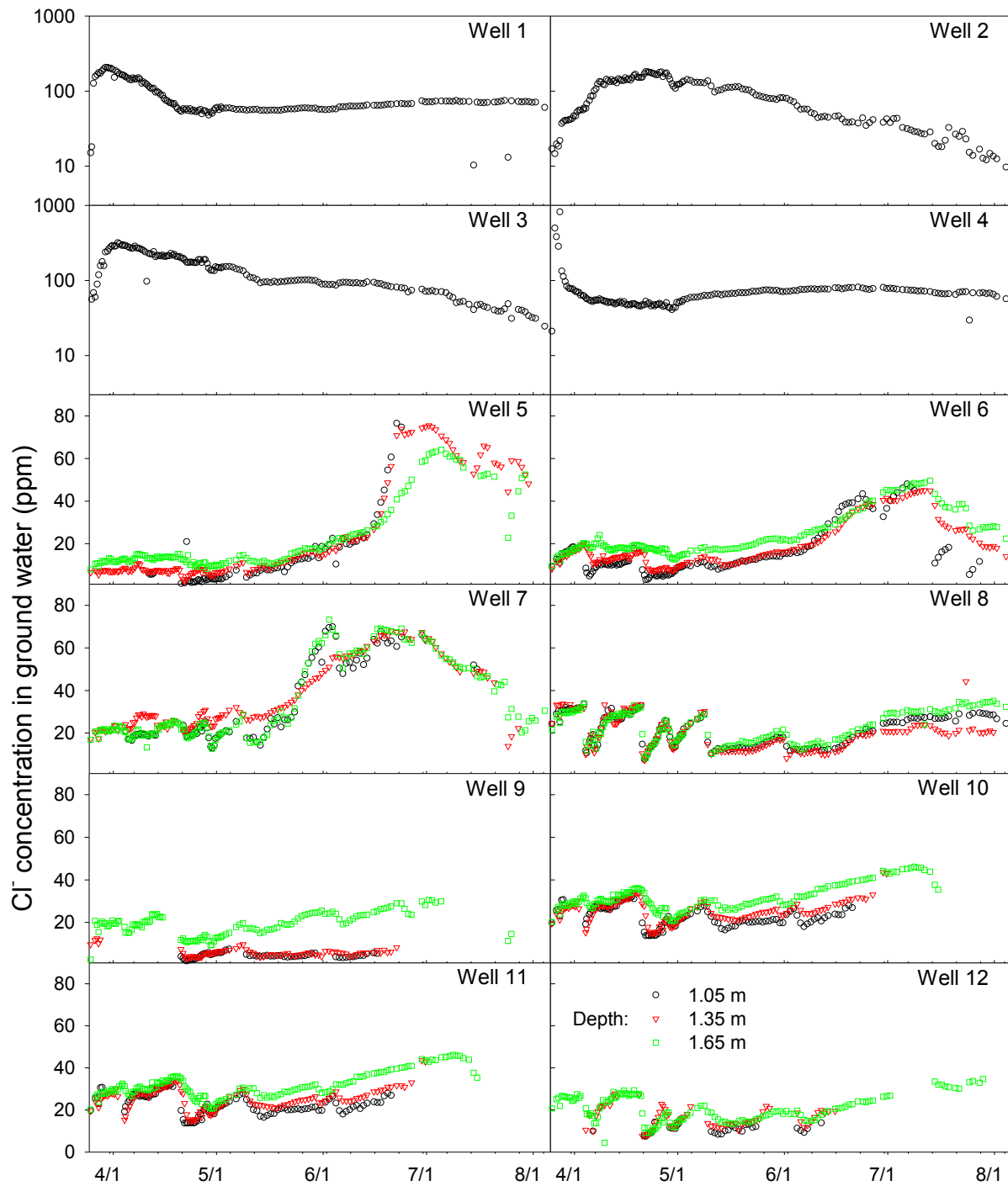
Figure 3-36 Chloride contents of groundwater measured in 12 wells during the lateral flow experiment.

The Cl- time series measured in 12 wells are shown in Fig. 3-36. The ARS study team   observed two distinctly different breakthrough curves at the irrigated site. A fast increase in Cl-concentration was observed in wells 1 and 4 at the initial stage of the breakthrough curve soon after the tracer was applied.

A gradual decrease and another increase in Cl- concentration followed the initial stage. Contrary to wells 1 and 4, a relatively slow increase was followed by a gradual decrease in Cl- in wells 2 and 3. These dramatic differences in the vertical transport rates of Cl- at the irrigated site were likely caused by different transport conditions. The water contents were relatively constant and groundwater was closer to the soil surface at locations 1 and 3 compared to locations 2 and 4 where GWDs and the water content in the upper 0.50-m soil layer varied significantly during the later flow experiment. The water content and GWD data indicated that flow conditions were close to saturation at locations 1 and 3, but far from saturation in the topsoil layer at locations 2 and 4.

The tracer arrived at locations 5 through 9 at different times. It took approximately 60 days for the Cl- to appear in noticeable amounts at three depths at location 7, and 80 days at locations 5 and 6. The distance between the irrigation site and locations 5 through 9 was 7 m, which means that the flow velocity should be within the range from 9 to 12 cm d^{-1}. The tracer arrived at locations 5 and 7 at higher concentrations than at location 6. Surprisingly, the ARS study team did not observe Cl- breakthrough at locations 8 and 9, which were also located 7 m downgradient from the irrigation site. Locations 5 through7 were probably closer to location 4, where vertical Cl- transport was faster than at location 3.

An increase in the Cl- concentration was observed at a depth of 1.65 m at locations 10, 11 and 12 at the end of the experiment. However, the increases were not so pronounced as those at locations 5, 6 and 7 where the concentrations remained within the range observed at locations 8 and 9.

The Cl- concentration time series differed at three depths at locations 5 through 12. Cl- concentrations were generally higher at depth of 1.6 m, than at the 0.5-m and 0.8-m depths, except at locations 8 and 12 where the Cl- concentrations did not change with depth, and location 7 where the highest Cl- concentrations were observed at a depth of 1.35 m. Noticeable differences in the Cl- breakthrough curves were observed at three depths at locations 5, 6 and7. The peak concentrations at location 6 did not differ among three depths. At location 5, larger values were observed at depths of 1.05 and 1.35 m, whereas at location 7 the peak Cl- concentrations at a depth of 1.35 m were smaller than those at depths of 1.05 and 1.65 m. Differences in the arrival times and the Cl- concentration at three depths and 12 locations imply that the tracer was transported preferentially in hydraulically active soil zones.

An interesting phenomenon the ARS study team observed was the synchronous dilution of groundwater at different depths during rainfall. This follows from the oscillations in the Cl- concentration time series at locations 7, 8, 10, 11 and 12. These oscillations corresponded to the GWD oscillations caused by fast decreases in the GWD after intensive rainfalls (Fig. 3-36). Smaller Cl- concentrations were measured when the GWDs were smaller. An increase in the GWD always caused an increase in the Cl- concentration. The decrease in Cl- concentration can be explained as being the result of ground water dilution by infiltrated precipitation (C_{Cl}=0.6 ppm), while an increase could be caused partly by evaporation. However, it remains to be explained why Cl- changes occurred synchronously at three different depths without changes in the concentration with depth.

4   DETAILED FLOW AND TRANSPORT MODELS OF
THE OPE3 WATERSHED SITE

4.1 Description of Models

Two fully 3D numerical codes, HYDRUS-3D and TOUGH2, were used for the water flow and contaminant transport simulations in the subsurface at the OPE3 site. Advantages of these two codes are their built-in mass conservation and having the same data preprocessing and post-processing through the entire flow domain. The codes are, however, computationally intensive and not attractive for parameter estimation, uncertainty, sensitivity evaluation, and monitoring optimization unless massive parallel computing is used (e.g., Coumou et al., 2008).

The complexity of subsurface flow systems creates a need and various opportunities for conceptual simplifications leading to a multiplicity of flow models. One common simplification is based on the assumption that lateral flow and transport in the unsaturated zone is not significant, except for the capillary fringe (Abit et al., 2008) or when perched temporary saturated layers appear are present (Twarakavi et al., 2009). If lateral flow is ignored in the unsaturated zone are, flow and transport can be simulated as a one-dimensional (1D) process, while flow and transport in groundwater remains a fully 2D or 3D phenomenon. The scarcity of information on the anisotropy of hydraulic properties in the vadose zone also favors the 1D approximation above the groundwater table. Such anisotropy has been demonstrated (Zhang et al., 2007), but is very rarely characterized to the extent of providing accurate estimates of the lateral flow process. As soon as the relative insignificance of lateral flow and transport in the vadose zone is surmised for the expected hydrologic forcing, a more efficient computational scheme will result, which is important when multiple simulations are needed to evaluate uncertainty or sensitivity.

Barthel (2006) discussed alternative integrated modeling approaches that are based on coupling of existing modeling concepts for the saturated and unsaturated zones. A problematic issue with coupling existing models is the question of how to tailor their output in a consistent manner. Twarakavi et al. (2008) noted that evaluation of interactions between near-surface and groundwater flow processes using coupled models has been a desirable but difficult goal. These authors coupled the HYDRUS-1D code for the vadose zone with the MODFOW code for groundwater zone. The two computer codes interact (i.e., exchange information about groundwater recharge and groundwater level) only at the end of each MODFLOW time step, during which HYDRUS may perform multiple time steps to simulate the unsaturated zone flow. The MODFLOW code receives the recharge flux from HYDRUS and evaluates a new water table depth for the next time step. A new water table depth is then estimated and assigned as the pressure head bottom boundary condition for the HYDRUS code at the next time step of the MODFLOW code. A similar method was applied by Facchi et al. (2004). This approach is known to create mass balance error. Twarakavi et al. (2009) noted that for cases where the coupled mass balance is not achieved satisfactorily, a version of the HYDRUS package is available that can attain complete mass balance between the saturated and unsaturated zones by suitably altering the unsaturated profiles for each zone at the end of each time step. The need for modifications of the profiles can create difficulties in using water flow simulation results in subsequent simulations of solute transport.

In efforts to control mass balance errors, Stoppelenburg et al. (2005) used an iterative approach to gradually adjust the recharge rates, while Liang et al. (2003) searched iteratively for the groundwater level until the mass balance requirements were met. In the latter case, the amount

of water in the unsaturated zone was used to adjust the position of the groundwater table. Shen and Phanikumar (2010) further used the mass balance equation to provide a linkage between the unsaturated zone and groundwater. The saturated zone vertically was represented as a single cell extending from the bottom of the unconfined aquifer to the lower boundary of the first element (cell) in the unsaturated zone above the water table. In other work, Yakirevich et al. (1998) coupled a 2D Boussinesq-like equation for the saturated zone with the 1D Richards model for the unsaturated zone. Unlike the Boussinesq equation for a leaky phreatic aquifer, the developed model did not contain a storage term with specific yield and a source term for natural replenishment. Instead, the equation included a water flux in the phreatic surface term through which the Richards equation was linked with the groundwater flow equation. Xu et al. (2012) integrated the Soil–Water–Atmosphere–Plant (SWAP) package into the MODFLOW groundwater flow model in such a way that the SWAP package calculates vertical fluxes for MODFLOW, while MODFLOW provides averaged water table depths to determine the bottom boundary condition for the SWAP zones.

Instead of coupling different models for areas above and below the groundwater table, one can also formulate one model for the entire continuous flow domain including both groundwater and the unsaturated zone. An efficient approach would be to develop a modeling approach that would (a) use the same model for the entire flow domain and (b) ignore lateral flow in the unsaturated zone or outside of the capillary fringe. The ARS study team addressed this by developing in this chapter the QUASI-3D model based on the Richards equation assuming 1D vertical flow in the unsaturated zone above the capillary fringe, together with 3D flow in groundwater and the capillary fringe in the unsaturated zone just above the groundwater table. As a first step the ARS study team implemented the quasi-3D approach only for water flow; the QUASI-3D code hence was not used for simulations of flow and transport of the OPE3 field experiment. The ARS study team benchmarked the code for simulating flow problems; some results are presented in Appendix B.

4.1.1 The HYDRUS-3D Code

The HYDRUS software package consists of a computational module, and an interactive graphics-based user interface. (Šimůnek et al., 2006). The package can be used to simulate a wide range of problems involving water, heat and solute movement in two- and three dimensional variably-saturated porous media.

The mathematical model for simulating water flow and contaminant transport is based on the Richards equation for flow and the advection-dispersion equation for transport. Assuming a conservative tracer (no sorption and decay), the governing equations are, respectively,

$$\frac{\partial \theta}{\partial t} = \nabla[\mathbf{K} \cdot (\nabla h + 1)] - S \qquad (4-1)$$

$$\frac{\partial \theta C}{\partial t} = \nabla(\mathbf{D} \cdot \nabla C - \mathbf{q}C) \qquad (4-2)$$

where θ is the volumetric water content (L^3/L^3), t is time (T), h is the pressure head (L), $\mathbf{K}$ is the hydraulic conductivity tensor (L/T), C is the solute concentration (M/L^3), $\mathbf{D}$ is the hydrodynamic dispersion tensor (L^2/T), $\mathbf{q}$ is the Darcy flux vector (L/T), S is a sink/source term accounting for water uptake by plant roots or the presence of wells (1/T), and t is time (T). The transport equation can be extended to include provisions for nonlinear nonequilibrium reactions between the solid and liquid phases, linear equilibrium reactions between the liquid and gaseous phases, zero- and first-order production and degradation reactions. In addition, physical nonequilibrium transport can be accounted for by assuming a two-region, dual-porosity type formulations.

The soil water retention and the unsaturated hydraulic conductivity functions needed for solution of Eq. (1) were described in this study using the van Genuchten (1980) relationships:

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[1 + \left(\alpha |h| \right)^n \right]^{-m}$$

(4-3)

$$K(S_e) = K_s S_e^{\,l} \left[1 - \left(1 - S_e^{\,1/m} \right)^m \right]^2$$

(4-4)

respectively, where S_e is relative saturation, θ_r is the residual water content (L^3/L^3), θ_s is the water content at saturation (L^3/L^3), K_s is the saturated hydraulic conductivity (L/T), α (L^{-1}) and n (-) are shape parameters, $m = 1 - 1/n$, and $l = 0.5$.

The components of the hydrodynamic dispersion tensor for the case of an isotropic porous medium are expressed as

$$D_{ij} = a_T v \delta_{ij} + \left(a_L - a_T \right) \frac{v_i v_j}{v}$$

(4-5)

where i,j = 1, 2, 3 correspond to the x, y, z coordinate directions, respectively, a_L and a_T are longitudinal and transverse dispersivities (L), respectively, v_i are components of the flow velocity (L/T), and $v = \sqrt{\sum_i v_i^2}$ (L/T)

The governing flow and transport equations in HYDRUS-3D were solved numerically using Galerkin-type linear finite element schemes. Application of equations (4-1) to (4-5) to the field tracer experiment requires definition of the domain boundaries, the initial and boundary conditions, a numerical finite element mesh, and soil hydraulic and solute transport parameters.

4.1.2 Modified TOUGH2 Code

The TOUGH2 code, developed by Lawrence Berkeley National Laboratory (Pruess et al., 1999), is a numerical model for multidimensional fluid and heat flow of multi-phase, multi-component fluid mixtures in porous and fractured media. TOUGH2 has been applied to a wide range of problems, including geothermal reservoir engineering, nuclear waste disposal, environmental assessment and remediation, and unsaturated and saturated zone hydrology (Pruess et al., 1999). An advantage of the TOUGH2 simulator is that it can simulate flow in fractured porous media using double-porosity, dual-permeability, and multiple interacting continua (MINC) formulations.

Similarly as HYDRUS-3D, the physical processes of fluid flow and tracer transport in TOUGH2 are governed by Equations (4-1) and (4-2), respectively. An integral finite-difference method is used to discretize the governing equations in space. TOUGH2 provides multiple equation-of-state (EOS) modules to handle different fluid mixtures, which are described by their components and phases with related thermophysical properties. The EOS9 module of TOUGH2 is a saturated-unsaturated flow module applicable to variably saturated media (Wu et al., 1996). EOS9 only considers a single aqueous phase and a single water component for variably saturated flow (Pruess et al., 1999). Like HYDRUS-3D, the EOS9 model ignores any effect of the air phase on fluid flow. The module is very

efficient numerically and well-suited to simulate water flow in the very heterogeneous variably saturated subsurface of the OPE3 site.

For tracer transport, the ARS study team used the T2R3D module of TOUGH2 code, which is based on Equation (4.2) assuming general multi-phase subsurface flow conditions (Wu et al., 1996). The current version of TOUGH2 did not have provisions for directly coupling the EOS9 module with T2R3D (only the numerically more cumbersome EOS3 multiphase module enabled this coupling). It was therefore necessary to develop a fully coupled EOS9 and T2R3D model to simulate fluid flow and tracer transport in 3D variably-saturated media. For this reason the ARS study team incorporated the T2R3D numerical solution for tracer transport directly into the original EOS9 code. The resulting scheme first numerically solved the governing flow equation for each grid block, followed by solution of the transport equation using simulated values for fluid flow obtained at the same time step.

The newly coupled model was first tested for flow simulations using a case of the Yucca Mountain project as solved with the EOS9 module. The flow results obtained with the two models were very close, indicating that the new coupled model could be used reliably for more general flow simulation in variably saturated porous media. The model was tested next for tracer transport using an example provided by the coupled EOS3 and T2R3D modules of TOUGH2. Differences between simulated tracer concentrations using the two models were within 5%, thus demonstrating that the new coupled model is suitable for simulations of tracer transport in variably saturated media.

4.1.3 FULL-3D and QUASI-3D Models

In this section, the ARS study team describes the FULL-3D and QUASI-3D finite difference codes based on complete and approximate solutions of the Richards equation for 3D variably-saturated media. The ARS study team is especially interested in QUASI-3D since this code provides an attractive alternative via model abstraction of the computationally more demanding HYDRUS-3D and TOUGH2 codes. QUASI-3D assumes that flow in the unsaturated zone above the capillary fringe occurs only in the vertical direction, but is fully 3D elsewhere. FULL-3D is an extension of QUASI-3D in that flow in the entire domain is described using the Richards equation.

Figure 4-1 shows a schematic of the modeled hydrologic system, which includes both the unsaturated zone and the underlying groundwater zone. The boundary $z_{cf}(x,y,t)$ between the two zones, referred to as the specific surface, is characterized by a prescribed pressure head h_{cf}. Let us introduce the following simplifying assumptions: 1) Flow in the unsaturated zone is in the vertical direction only; 2) 3D flow is considered in the groundwater zone, which can include the capillary fringe; 3) Water is incompressible and its density does not depend on solute concentration or temperature; and 4) the porous matrix is rigid.

Water flow in the unsaturated zone and in groundwater is governed by the general 3D Richards equation

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_x}{\partial x} - \frac{\partial q_y}{\partial y} - \frac{\partial q_z}{\partial z} + S \tag{4-6}$$

where θ is the volumetric water content [L^3L^{-3}], t is time [T], S is a sink/source term [T^{-1}], x and y are horizontal coordinates [L], z is the vertical coordinate, directed upward [L] and, q_y and q_z are Darcy-Buckingham fluxes [LT^{-1}] given by
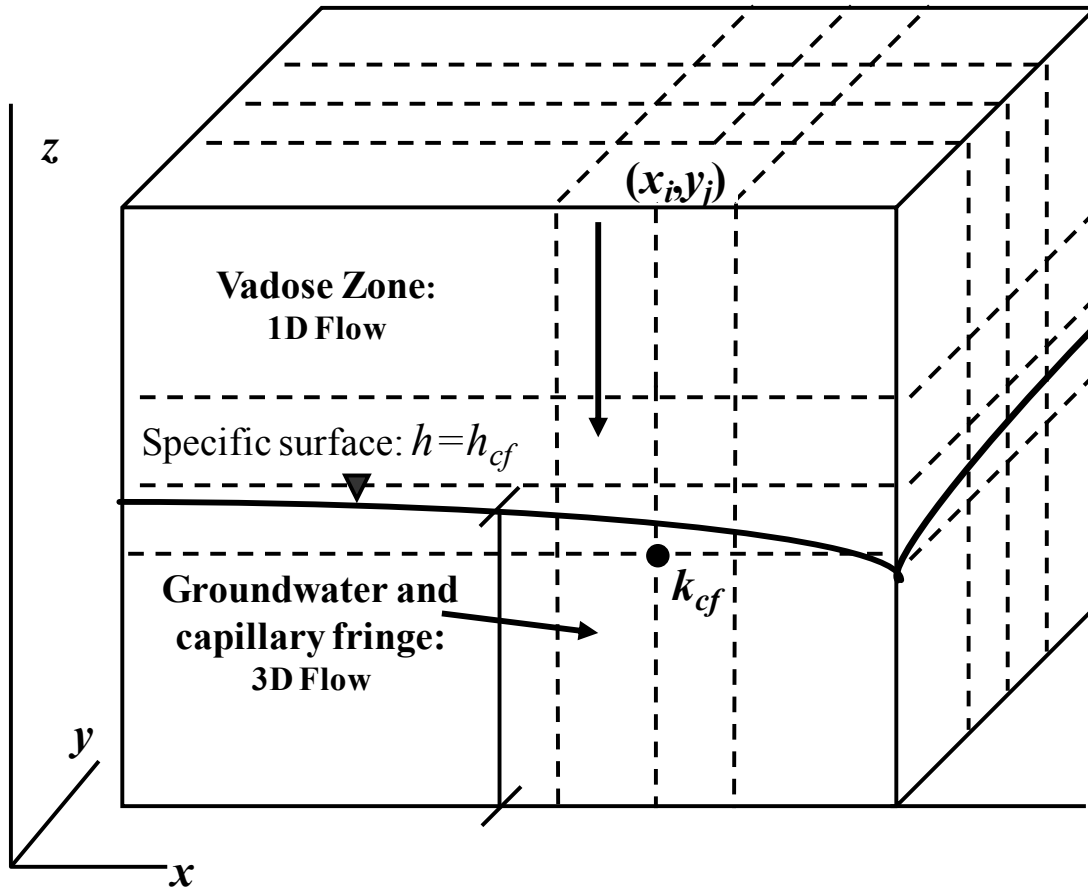
Figure 4-1   Modeled variably-saturated subsurface system with a section of the finite difference (FD) grid (dashed lines).

$$q_x = -K_x \frac{\partial H}{\partial x} \qquad q_y = -K_y \frac{\partial H}{\partial y} \qquad q_z = -K_z \frac{\partial H}{\partial z}$$

in which K_x, K_y and K_z are hydraulic conductivitites in the x, y, and z directions [L], $H=h + z$ is the total hydraulic head [L], and h is pressure head [L].

Using the assumption that lateral water flow is negligible in the unsaturated zone above capillary fringe ($q_x = q_y = 0$), a series of 1D Richards equations can be used to describe vertical flow at selected points within an areal plane as follows

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_z}{\partial z} + S \tag{4-7}$$

Flow in groundwater and the capillary fringe is governed also by (4-6), but with $\partial \theta / \partial t = 0$ assuming fully saturated flow. At the specific surface (4-6) and (4-7) are coupled subject to conditions of continuity for the hydraulic head and the vertical water fluxes, i.e.,

$$H\big|_{z=z_{cf}^+} = H\big|_{z=z_{cf}^-} \tag{4-8}$$

$$q_z\big|_{z=z_{cf}^+} = q_z\big|_{z=z_{cf}^-} \tag{4-9}$$

respectively, where z_{cf} is the coordinate of the moving specific surface evaluated as a height where the hydraulic head is equal to the prescribed value to give

$$h(x,y,z,t)\big|_{z=z_{cf}} = h_{cf} \tag{4-10}$$

in which h_{cf} characterizes the thickness of the capillary fringe above the water table (phreatic surface). Lateral flow in the capillary fringe is not considered if $h_{cf}=0$, in which case the specific surface coincides with the water table.

The model given by (4-6) through (4-10) was solved numerically using a finite-difference (FD) approximation on a rectangular grid. While not mandatory, the ARS study team considers here for simplicity, a uniform grid with nodes $x_i=(i-1)\Delta x$, $i=1,2,\ldots,M_x$, $y_j=(j-1)\Delta y$, $j=1,2,\ldots,M_y$, and $z_k=(k-1)\Delta z$, and $k=1,2,\ldots,M_z$.

The general FD approximation of (4-6) and (4-7) is

$$\frac{\theta_{i,j,k}^{n+1} - \theta_{i,j,k}^{n}}{\Delta t^n} = \Lambda_{x_{i,j,k}}^{n+1} + \Lambda_{y_{i,j,k}}^{n+1} + \Lambda_{z_{i,j,k}}^{n+1} + S_{i,j,k}^{n+1} \tag{4-11}$$

where the superscripts n and $n+1$ denote old and new time steps, respectively, and

$$\Lambda_{x_{i,j,k}}^{n+1} = -\frac{q_{i+1/2,j,k}^{n+1} - q_{i-1/2,j,k}^{n+1}}{\Delta x} , \Lambda_{y_{i,j,k}}^{n+1} = -\frac{q_{i,j+1/2,k}^{n+1} - q_{i,j-1/2,k}^{n+1}}{\Delta y} , \Lambda_{z_{i,j,k}}^{n+1} = -\frac{q_{i,j,k+1/2}^{n+1} - q_{i,j,k-1/2}^{n+1}}{\Delta z}$$

$$q_{i+1/2,j,k}^{n+1} = \begin{cases} -K_{i+1/2,j,k}^{n+1} \dfrac{H_{i+1,j,k}^{n+1} - H_{i,j,k}^{n+1}}{\Delta x} & \text{for} \quad k \le k_{cf}(i,j) \\ 0 \quad \text{for} \quad k > k_{cf}(i,j) \end{cases}$$

$$q_{i,j+1/2,k}^{n+1} = \begin{cases} -K_{i,j+1/2,k}^{n+1} \dfrac{H_{i,j+1,k}^{n+1} - H_{i,j,k}^{n+1}}{\Delta y} & \text{for} \quad k \le k_{cf}(i,j) \\ 0 \quad \text{for} \quad k > k_{cf}(i,j) \end{cases}$$

$$q_{i,j,k+1/2}^{n+1} = -K_{i,j,k+1/2}^{n+1} \frac{H_{i,j,k+1}^{n+1} - H_{i,j,k}^{n+1}}{\Delta z}$$

in which $k_{cf}(i,j)$ refers to the first node on the vertical grid line (x_i,y_j) below the moving boundary defined by (4-6) (Figure 4-1). Values of the hydraulic conductivity ($K_{i+1/2,j,k}^{n+1}$) are calculated as the arithmetic mean of the hydraulic conductivities at the neighboring cells (i,j,k) and $(i+1,j,k)$. The value of $k_{cf}(i,j)$ meets the conditions

$$\left| H_{i,j,k_{cf}} - z_{k_{cf}} - h_{cf} \right| < \Delta z/2 \tag{4-12}$$

The system of equations given by (4-11) is strongly nonlinear due to the nonlinearity of the $\theta(h)$ and $K(h)$ functions). A common method to solve (4-11) is linearization of the equations leading to a linear set of equations at each iteration of the form (Celia et al. 1990)

$$\mu_{i,j,k}^{(s)} \frac{H_{i,j,k}^{(s+1)} - H_{i,j,k}^{(s)}}{\Delta t^n} + \frac{\theta_{i,j,k}^{(s)} - \theta_{i,j,k}^{n}}{\Delta t^n} = \Lambda_{x,i,j,k}^{(s+1)} + \Lambda_{y,i,j,k}^{(s+1)} + \Lambda_{z,i,j,k}^{(s+1)} + S_{i,j,k}^{(s)} \tag{4-13}$$

where $\mu = d\theta/dh$ is the soil water (or specific) capacity, $s = 0, 1, 2,..$ is the iteration number in the iterative solution process, with $H^{(0)} = H^n$ and $\theta^{(0)} = \theta^n$.

The assumption that lateral water flow is negligible in the unsaturated zone above the capillary fringe can substantially decrease the number of computations needed to solve (4-13). Specifically, the system of the FD equations given by (4-13) can be decomposed into two simpler systems that are linked to each other at only one node on each vertical grid line. One possible decomposition is to use the following 1D equations for the unsaturated zone above the specific surface

$$\mu_{i,j,k}^{(s)} \frac{H_{i,j,k}^{(s+1)} - H_{i,j,k}^{(s)}}{\Delta t^n} + \frac{\theta_{i,j,k}^{(s)} - \theta_{i,j,k}^{n}}{\Delta t^n} = \Lambda_{z_{i,j,k}}^{(s+1)} + S_{i,k,j}^{(s)}, \quad k > k_{cf}(i,j) \tag{4-14}$$

while the full 3D equations are used for the medium below the specific surface:

$$\mu_{i,j,k}^{(s)} \frac{H_{i,j,k}^{(s+1)} - H_{i,j,k}^{(s)}}{\Delta t^n} + \frac{\theta_{i,j,k}^{(s)} - \theta_{i,j,k}^{n}}{\Delta t^n} = \Lambda_{x,i,j,k}^{(s+1)} + \Lambda_{y,i,j,k}^{(s+1)} + \Lambda_{z,i,j,k}^{(s+1)} + S_{i,j,k}^{(s)}, \quad k \leq k_{cf}(i,j) \tag{4-15}$$

Equations (4-14) and (4-15) can be rewritten as follows

$$A_{i,j,k} H_{i,j,k}^{(s+1)} - D_{i,j,k} H_{i,j,k-1}^{(s+1)} - G_{i,j,k} H_{i,j,k+1}^{(s+1)} = W_{i,j,k}^{(s)} \tag{4-16}$$

$$A_{i,j,k} H_{i,j,k}^{(s+1)} - B_{i,j,k} H_{i-1,j,k}^{(s+1)} - C_{i,j,k} H_{i,j-1,k}^{(s+1)} - D_{i,j,k} H_{i,j,k-1}^{(s+1)}$$
$$- E_{i,j,k} H_{i+1,j,k}^{(s+1)} - F_{i,j+1,k} H_{i,j+1,k}^{(s+1)} - G_{i,j,k} H_{i,j,k+1}^{(s+1)} = W_{i,j,k}^{(s)} \tag{4-17}$$

At internal nodes of a regular grid, the coefficients of (4-16) and (4-17) become

$$A_{i,j,k} = B_{i,j,k} + C_{i,j,k} + D_{i,j,k} + E_{i,j,k} + F_{i,j,k} + G_{i,j,k} + \frac{\mu_{i,j,k}^{(s)}}{\Delta t^n} \tag{4-18a}$$

$$B_{i,j,k} = \frac{k_{i+1/2,j,k}^{(s)}}{\Delta x^2} \qquad C_{i,j,k} = \frac{k_{i,j+1/2,k}^{(s)}}{\Delta y^2} \qquad D_{i,j,k} = \frac{k_{i,j,k+1/2}^{(s)}}{\Delta z^2} \tag{4-18b}$$

$$E_{i,j,k} = B_{i+1,j,k}; \; F_{i,j,k} = C_{i,j+1,k}; \; G_{i,j,k} = D_{i,j,k+1} \tag{4-18c}$$

$$W_{i,j,k}^{(s)} = S_{i,j,k}^{(s)} + \frac{\mu_{i,j,k}^{(s)}}{\Delta t^n} H_{i,j,k}^{(s)} - \frac{\theta_{i,j,k}^{(s)} - \theta_{i,j,k}^{n}}{\Delta t^n} \tag{4-18d}$$

The tridiagonal system of equations given by (4-16) are most easily solved using the standard Thomas algorithm:

$$H_{i,j,k+1} = \alpha_{i,j,k} H_{i,j,k} + \beta_{i,j,k} \tag{4-19}$$

for $k=k_{cf}, k_{cf+1}, \ldots M_z\text{-}1, M_z$. and

$$\alpha_{i,j,k} = D_{i,j,k} / (A_{i,j,k} - \alpha_{i,j,k+1} G_{i,j,k}),$$

$$\beta_{i,j,k} = (\beta_{i,j,k+1} G_{i,j,k} - W_{i,j,k}^{(s)}) / (A_{i,j,k} - \alpha_{i,j,k+1} G_{i,j,k}) \tag{4-20}$$

Now consider the subset of equations given by (4-15) pertaining to the combined saturated and capillary fringe region below the specific surface. These equations also contain values of H evaluated at nodes (i,j,k_{cf+1}). The ARS study team used (4-14) with $k=k_{cf}$, and substitute this equation into (4-15) to eliminate the variable $H_{i,j,k_{cf}+1}^{(s+1)}$, thus obtaining

$$A_{i,j,k_{cf}} H_{i,j,k_{cf}}^{(s+1)} - B_{i,j,k_{cf}} H_{i-1,j,k_{cf}}^{(s+1)} - C_{i,j,k_{cf}} H_{i,j-1,k_{cf}}^{(s+1)} - D_{i,j,k_{cf}} H_{i,j,k_{cf}-1}^{(s+1)}$$
$$- E_{i,j,k_{cf}} H_{i+1,j,k_{cf}}^{(s+1)} - F_{i,j+1,k_{cf}}^{(s+1)} H_{i,j+1,k_{cf}}^{(s+1)} - G_{i,j,k_{cf}} \left(\alpha_{i,j,k_{cf}} H_{i,j,k_{cf}}^{(s+1)} + \beta_{i,j,k_{cf}} \right) = W_{i,j,k_{cf}}^{(s)} \tag{4-21}$$

The system of equations (4-21) allows one to obtain the pressure head values $H_{i,j,k}^{(s+1)}$ in the sub-domain $k \le k_{cf}$. Having evaluated the $H_{i,j,k_{cf}}^{(s+1)}$, values of $H_{i,j,k}^{(s+1)}$ for $k > k_{cf}$ are estimated using (4-19) at each (i,j) node. The iteration step is completed by checking the condition stated in (4-7) and determining the new position of the specific surface for each pair (i,j). Iterations for each time step end when the following conditions are fulfilled

$$\max(H_{ijk}^{(s+1)} - H_{ijk}^{(s)}) < \varepsilon_H \tag{4-22a}$$

$$\max(\theta_{ijk}^{(s+1)} - \theta_{ijk}^{(s)}) < \varepsilon_\theta, \quad \text{for } h < 0 \tag{4-22b}$$

in which ε_H and ε_θ are prescribed error tolerances for the total head and the water content, respectively.

The numerical model above is complemented with initial and boundary conditions that prescribe heads and/or fluxes at the assumed domain boundaries. The algorithm given by (4-11) to (4-22) was implemented by modifying the MODFLOW code (Harbaugh et al., 2000), and applying the Groundwater Modeling Software (GMS 6.0, 2002) software for pre- and post-processing. The MODFLOW model was modified for this purpose in two ways. First, the composite 1D unsaturated zone and 3D groundwater flow model was combined into the QUASI-3D code. The ARS study team next solved the complete 3D Richards equation to yield the FULL-3D code using procedures very similar as implemented for the VSF model of Thoms et al. (2006). The FULL-3D was later used as a benchmark for testing the computationally more attractive QUASI-3D model.

The ARS study team noted that the manner in which the saturated and unsaturated zones in QUASI-3D are coupled, including the invoked numerical solution algorithm, is different from the

methods presented by Yakirevich et al. (1998) and Twarakavi et al. (2008). QUASI-3D implicitly couples the vadose zone and groundwater flow equations, which implies that calculations of the phreatic surface water flux at each time step are not required as a source term for the groundwater flow equation. The QUASI-3D also does not require introducing the specific yield parameter, unlike HYDRUS-MODFLOW (Twarakavi et al., 2008).

4.2 Model Calibration

4.2.1 Geometry of the Simulation Domain and Soil Material Distribution

Based on experimental data that include the location of the tracer application plot, observation (sampling) wells, the topography of the soil surface and of the low permeability bottom boundary (estimated from GPR survey data), the ARS study team considered for HYDRUS-3D simulations a 3-D layered domain that extended laterally for 55 m in the x (east-west) direction and for 50 m in the y (north-south) direction. The thickness (z direction) of the domain varied from 3.12 m to 5.18 m. The domain was smaller than the domain used for the preliminary simulations (Pachepsky et al., 2011) to reduce the simulation time. The unstructured (triangular prisms) finite element mesh consisted of 24,780 nodes and 45,640 3D elements (Fig. 4-2). The simulation domain was subdivided into 15 layers. The vertical size of the finite element prisms varied depending upon location, with the elements following the surface of the soil surface and the surface of the lower sub-layer. In the upper 1.65-m soil layer the vertical size of the prisms was about 20 cm, while the size was about 40 cm below this depth (Fig. 4-2). The horizontal mesh size decreased from 5 m near the downgradient domain boundary to 0.7 m close to the irrigation plot where most of the transport processes were expected. Using this geometry and finite element mesh, a 125-day long simulation generally took between 3 to 6 h with a Dell PC Dimension DXC061, Intel Core™2 Duo CPU 6700 @2.66 GHz. Further refinement of the finite element mesh required longer simulation times.

Twenty observation nodes (Fig. 4-2) were used to record the simulated breakthrough curves at specific sampling locations in observation wells L1 through L12. Although groundwater was sampled at three depths in wells L5 through L12, only two observation nodes were located at these wells since the grid was very coarse in the lower soil horizon.

The distribution of soil materials with different textures within the domain was set based on logs obtained during installation of observation wells L1-L12 (Fig. 4-3). First the ARS study team interpolated measured soil textures to depths corresponding to each computational layer. A cluster analysis (SPLUS 2000) was subsequently used to group the interpolated soil textures into 9 classes so that each class was represented by soil texture averaged for the samples belonging to this class (Fig.4-4). These classes included: three sandy loam classes (SL1, SL2 and SL3), two sandy clay loam classes (SCL1 and SCL2), two loam classes (L1 and L2), and two silty clay loam classes (SiCL1 and SiCL2). The soil water retention and hydraulic conductivity parameters of the van Genuchten equations (4-3) and (4-4) were obtained for the 9 texture classes with the ROSETTA software using data for soil texture and bulk density (Fig. 4-5, Table 4-1). The boundaries between different soil materials of each computational layer were set at about equal distances from the observation wells where materials were measured (Fig. 4-6). Soil texture measured at a depth of 210 cm was mostly presented by the SiCL2 texture class. Since the ARS study team did not have measurements below this depth, the SiCL2 texture at this depth was extended to the bottom of the flow domain. Subdividing soil hydraulic properties into 9 groups was expected to reflect hydrological complexity of the simulated domain and to provide flexibility in the model calibration process.

4.2.2 Initial and Boundary Conditions

As initial condition for the water flow simulations the ARS study team used an equilibrium pressure head distribution. This distribution was obtained from groundwater depths measured in the observation wells at the beginning of the tracer experiment (Fig. 4-7a). Spatially variable, but constant in time, pressure head profiles were specified along the lateral boundaries based on groundwater monitoring data obtained in the summers of 2006-2007 (Fig. 4-7b). The initial Cl^- distribution varied between 10 to 25 mg/l depending upon location as measured in the observation wells prior to the experiment (Fig. 4-7c). A third type (Cauchy) boundary condition was used along the soil surface to simulate solute transport. Cl^- concentrations of groundwater along the lateral boundaries were set to the initial concentrations observed before the tracer experiment.

Daily-averaged atmospheric boundary conditions were used at the soil surface. These conditions simulated a variable flux due to evapotranspiration and prolonged infiltration observed after precipitation and intense initial irrigation within the irrigated plot area and along the remaining part of the soil surface (Fig. 4-8). The ARS study team used a zero flux boundary condition along the bottom of the domain.

The imposed flow rate during each irrigation event was decreased by 28% to account for water lost to overland flow that was collected. This runoff water had no effect on the subsurface flow and transport processes. The Cl^- concentration of the applied tracer solution was 4940 mg/l, while the Cl^- concentrations of the rainwater and irrigation water were 8 mg/l.

Plot size: 55 x 50 m

Nodes: 247,80

3D elements: 45,640
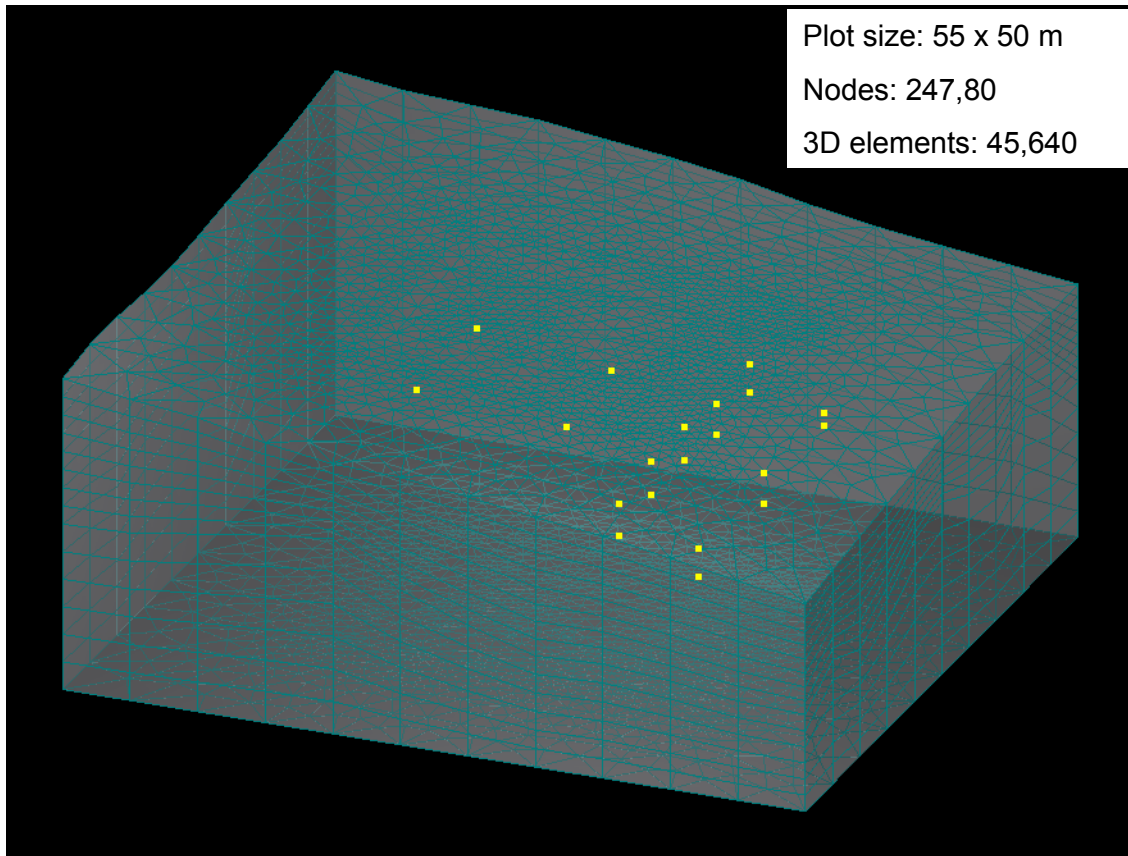
Figure 4-2   Mesh generated for HYDRUS-3D simulations. Observations nodes are shown in yellow.
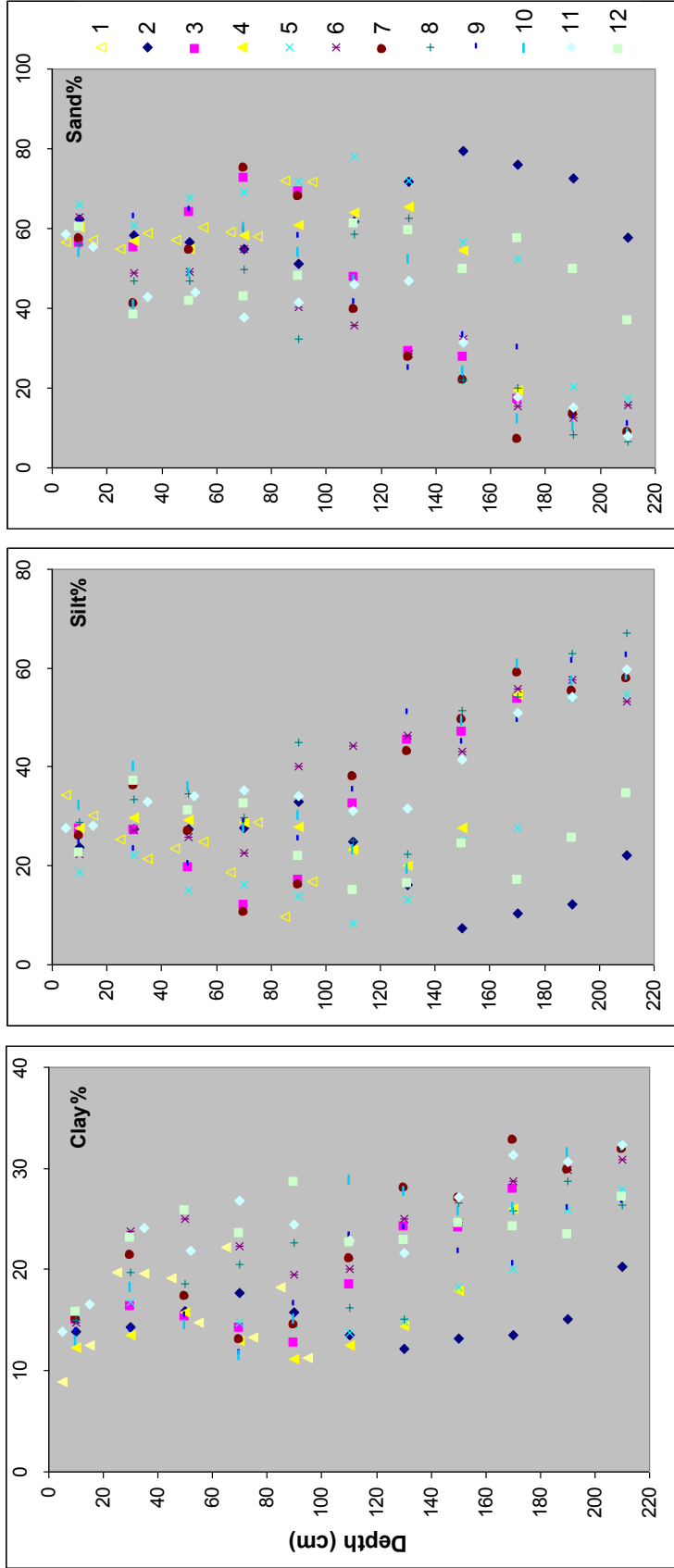
Figure 4-3   Vertical distribution of clay, silt and sand contents measured at 12 locations.
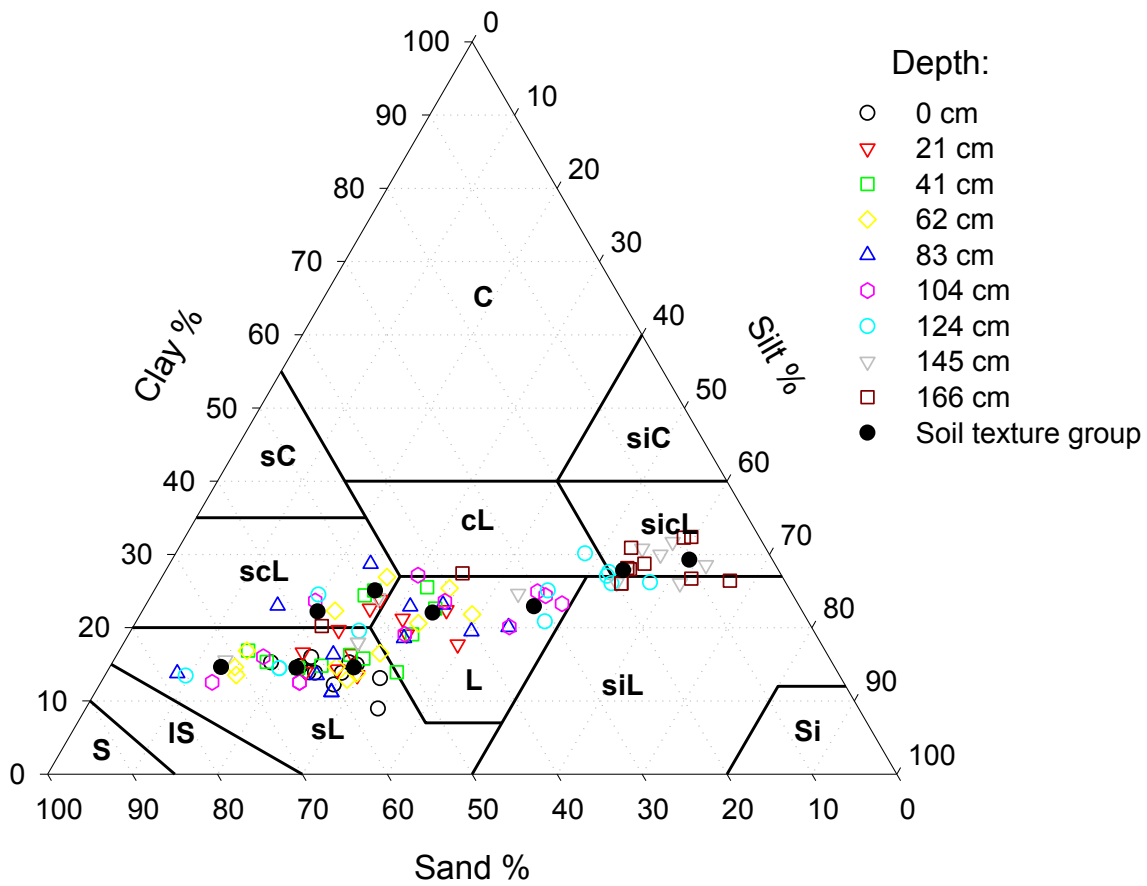
Figure 4-4   Soil texture groups obtained with cluster analysis.
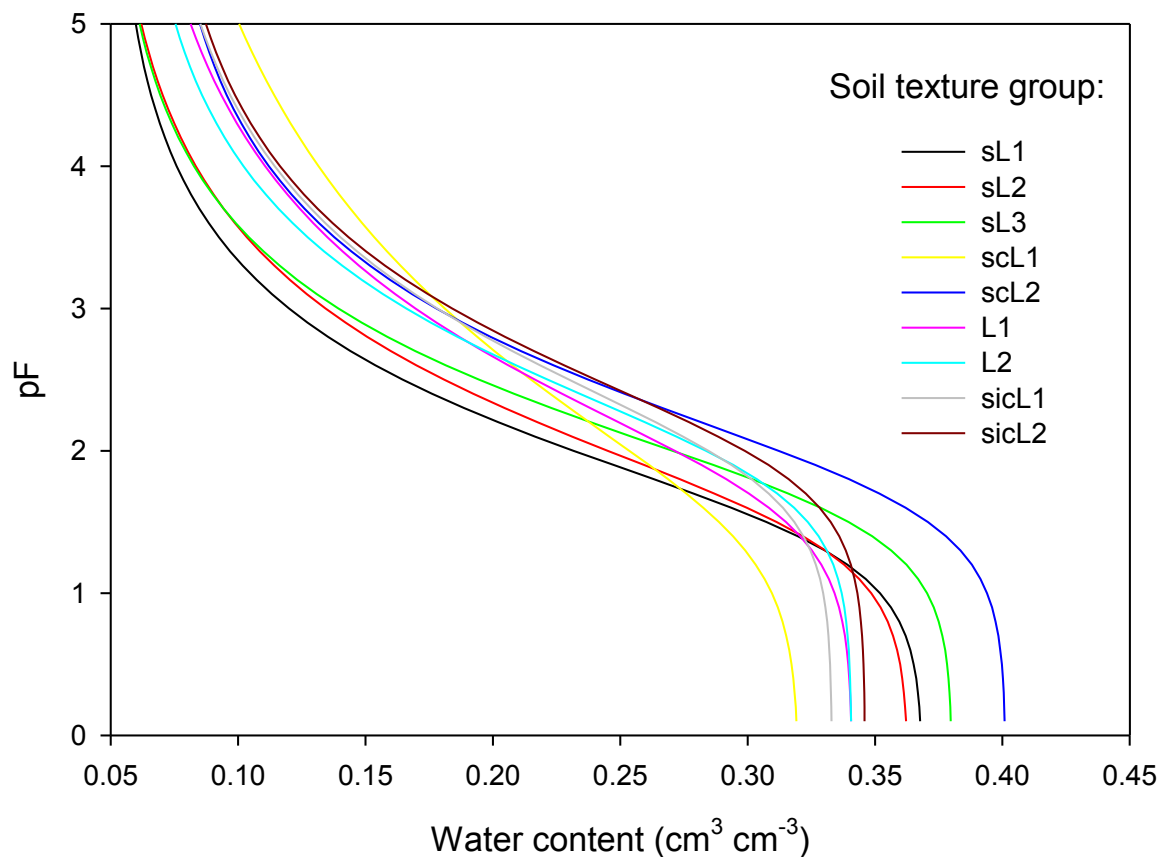
Figure 4-5   ROSETTA estimated soil water retention curves for selected soil texture groups.

Table 4-1   Soil properties used in simulations of Cl subsurface transport.

Texture	θr	θs	α	n	K_{s1}	K_{s2}	L	BD	Disp-L	*Disp-T*
	cm^3cm^{-3}	cm^3cm^{-3}	m^{-1}	-	$m\ d^{-1}$	$m\ d^{-1}$	-	$g\ cm^{-3}$	m	m
SL1	0.051	0.369	3.20	1.440	0.32	2.50	0.5	1.6	1.0	0.2
SL2	0.047	0.500	2.97	1.378	0.21	0.10	0.5	1.6	1.0	0.2
SL3	0.049	0.380	2.08	1.426	0.24	2.50	0.5	1.5	1.0	0.2
SCL1	0.048	0.320	3.34	1.204	0.04	0.45	0.5	1.8	1.0	0.2
SCL2	0.067	0.401	1.65	1.391	0.11	0.11	0.5	1.5	1.0	0.2
L1	0.053	0.341	1.85	1.309	0.04	2.00	0.5	1.7	1.0	0.2
L2	0.057	0.341	1.17	1.386	0.03	0.70	0.5	1.7	1.0	0.2
SiCL1	0.062	0.333	1.10	1.348	0.02	2.00	0.5	1.8	0.1	0.02
SiCL2	0.066	0.346	0.98	1.375	0.01	0.15	0.5	1.8	1.0	0.2

K_{s1} and K_{s2} are Rosetta-estimated and fitted values of saturated hydraulic conductivity.
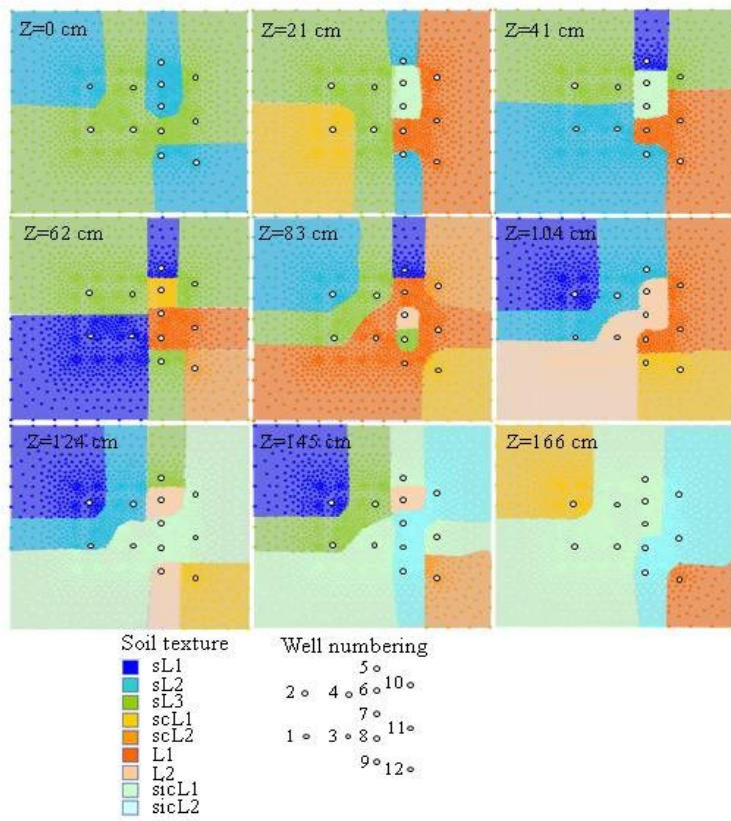
Figure 4-6   Material distribution used for the HYDRUS-3D simulations.

a) Initial pressure head distribution   b) Constant pressure at the domain boundaries

pressure head (m)
3.000
2.500
2.000
1.500
1.000
0.500
0.000
-0.500
-1.000
-1.500
-2.000
-2.500

c) Initial Cl distribution

Cl concentration (ppm)
25.000
15.000
5.000
0.000
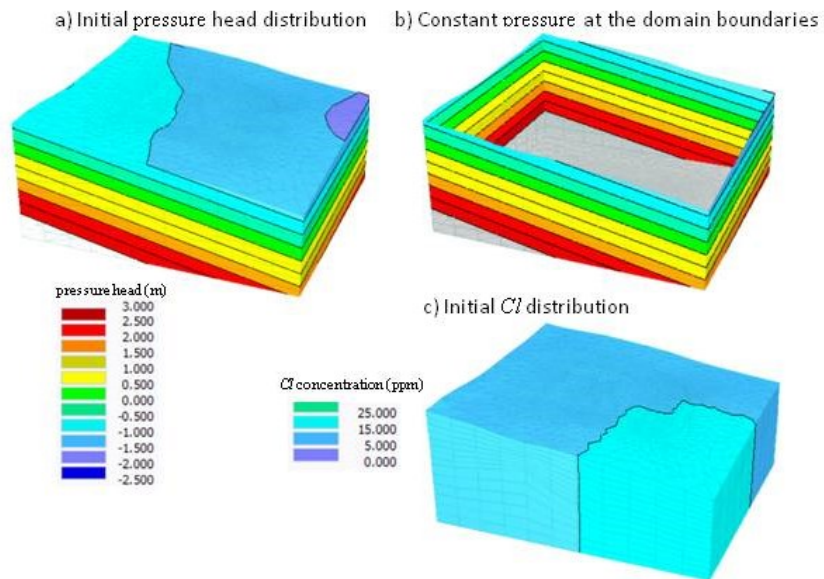
Figure 4-7   Initial and boundary conditions used for HYDRUS-3D simulations.
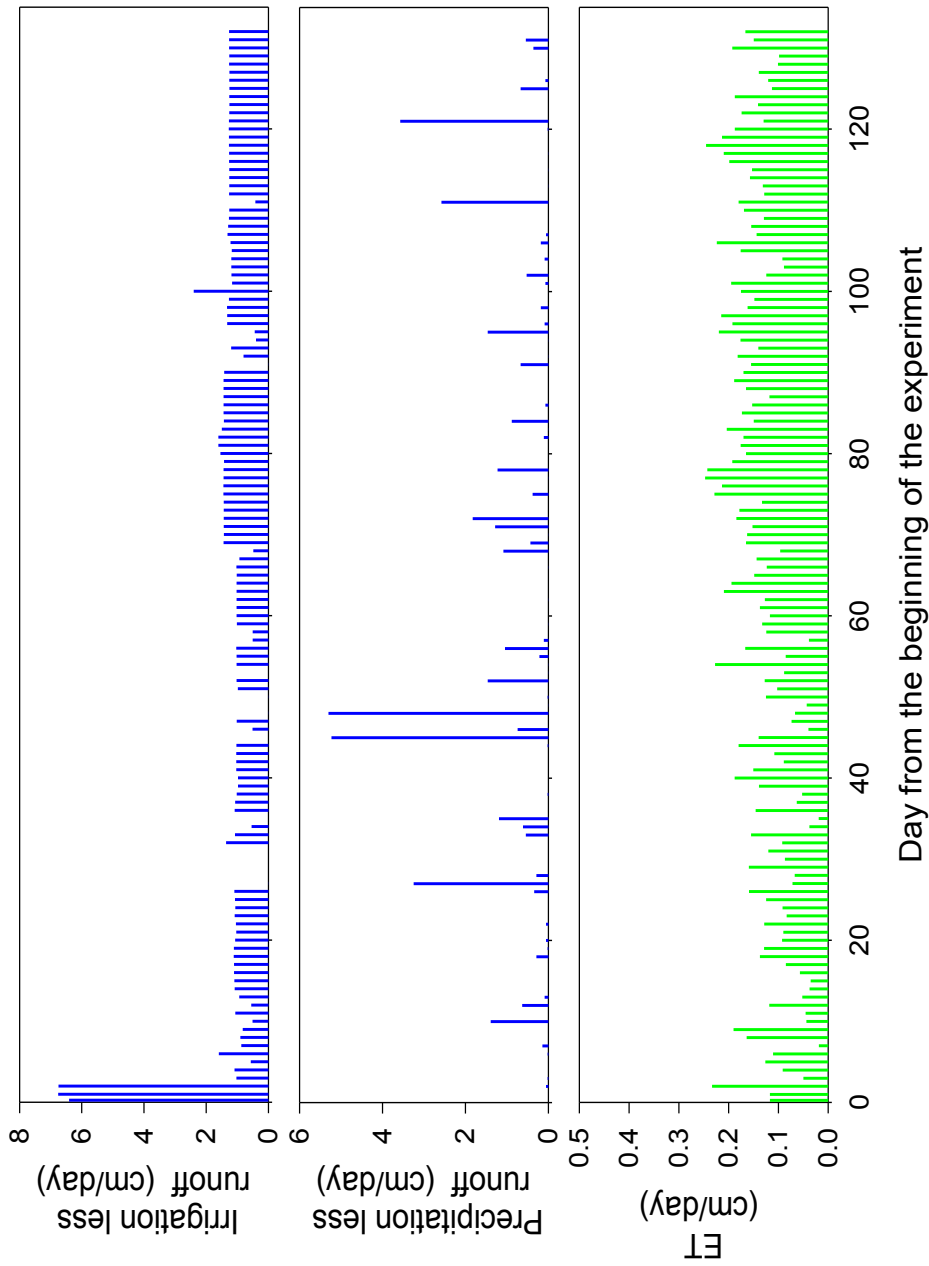
Figure 4-8 Daily irrigation, precipitation and evapotranspiration data used for the HYDRUS-3D simulation.

4.2.3 Model Calibration

4.2.3.1 HYDRUS-3D Simulations

The HYDRUS-3D software in its current version does not have an easily implemented option for automatic model calibration. Such an optimization would take an unreasonable amount of time for the fully 3-D variably-saturated field scale problem. Therefore, model calibration was performed using a trial-and-error method. Our approach involved the sequential fitting of flow and transport parameters to available data. At a first step the ARS study team estimated the velocity field by varying the saturated hydraulic conductivities of the various soil materials. The second step was the calibration of the transport parameters (dispersivities for the conservative tracer) through a series of trial-and-error runs. Much of our attention was focused on observation wells L5 through L7, where well-defined breakthrough curves (BTCs) were observed, and where both the arrival front and the receding (tailing) part of the breakthrough curve had been recorded.

The ARS study team started HYDRUS-3D simulations using the hydraulic parameter values estimated with the ROSETTA software (Table 4-1). The initial values of the longitudinal dispersivity were set at 0.5 m, while the ratio of longitudinal to transversal dispersivity was fixed at 5 for all simulations. The simulation results indicated that soil hydraulic conductivity was not sufficient to allow infiltration of the amounts of water precipitated during rainfall and irrigation events. Stable model performance and realistic groundwater depth fluctuations were achieved only after the K_s values were increased tenfold. The model reasonably well reproduced the phases of the groundwater fluctuations, but not their amplitudes (Fig. 4-9). This could be caused by the adopted coarse vertical discretization of the simulated domain, ignoring macropore flow, and imposing time-independent pressure heads along the domain boundaries in the simulations.

The trial-and-error procedure continued until a better fit for L5 through L7 was obtained. During these runs the ARS study team varied the values of K_s, and the longitudinal and transversal dispersivities, to obtain more rapid vertical Cl$^-$ transport at the irrigated plot, and an increase followed by a decrease in Cl$^-$ concentrations in wells L5-L7.

Simulated and observed breakthrough curves are compared in Fig. 4-10. The ARS study team noticed that the simulated BTCs for wells L1 and L4 exhibited much later arrival times of the tracer than the observed BTCs. Simulated tracer arrived earlier or at about the same time with measured tracer in wells L2 and L3, respectively (Table 4-2). The vertical flow during ponded infiltration was not simulated properly. Simulated peak concentrations also deviated from the observed in L1-L4. Better agreement of simulated and measured peaks was obtained for wells L5-L9.

Partitioning soil texture into 9 classes provided flexibility for model calibration, particularly at depths from 100 to 180 cm. Introducing a higher value of the saturated hydraulic conductivity K_s for materials L1 and SiCL1 caused HYDRUS-3D to predict more rapid tracer transport toward wells L5-L8 at depths 110 and 160 cm (Fig. 4-10). The applied tracer initially was pushed by the incoming fresh water through the high water conductive materials SL1 and SL3 to deeper soil layers. Unsaturated soil in the top 100-cm soil layer prevented lateral plume propagation beyond the irrigation plot. Vertical Cl$^-$ transport dominated at this stage. When the plume approached the saturated zone with the more permeable material L1 at a depth of 104 cm and SiCL1 at depths of 124 cm and 145 cm, vertical flow turned horizontal at these depths, also due to low K_s values in material SiCL2 underlying the SiCL1 at depths of 145 cm and 166 cm (see Fig. 4-6).

Thus, the incoming irrigation water pushed the plume laterally in the highly permeable L1 and SiCL1 materials (see Fig. 4-6). Thus, the incoming irrigation water pushed the plume laterally at depths 110 and 160 cm. After some period of time, the plume had a toroid shape with the concentration in the center of the plume being lower than at the peripheral part. This shape developed due to due to the application of clean irrigation water after Cl- injection The migration of the plume and its expansion in different directions was controlled by the spatial distribution of the hydraulic properties. Near the end of the simulation the advancing front plume had passed observation wells L5-L7.

In spite of having different transport distances from the irrigated plot to the transects crossing wells L5-L9 and wells L10-12 the observed increase in Cl⁻ concentration in the more distant wells (L10-L12) occurred almost synchronously with increases in the proximate wells (L5-L9). This indicated preferential transport toward wells L10-L12 through material SiCL2 at a depth of 166 cm. However, an increase in K_s for this layer resulted in earlier tracer arrival to wells L7-L9 at a depth of 145 cm. This explains why the ARS study team were not able to simulate the BTCs in those three distant wells.

The simulated depths differed from the observation depths because of coarse vertical discretization of the simulated domain. To compare results of the simulation with measured Cl⁻ data, the BTCs were averaged at each location. Cl was measured in the wells at depths of 105 cm, 135 cm and 165 cm. The computational observation HYDRUS 3d nodes were set at depths of 110 cm and 160 cm, due to the vertical discretization used in the simulations. The ARS study team compared measured concentrations averaged across the three depths (105 cm, 135 cm and 165 cm) with the simulated Cl concentrations averaged across the two depths (110 cm and 160 cm). Results of the comparison are shown in Fig. 4-10. Overall, the model was not able to simulate fast Cl⁻ transport within the irrigation plot and early plume arrival in wells L10-L12. However, the calibrated HYDRUS-3D model did reproduce the most important features of the lateral flow experiment. The model gave satisfactory levels of the peak concentrations and was able to capture arrival times in the first line of the observation wells (L5-L7).
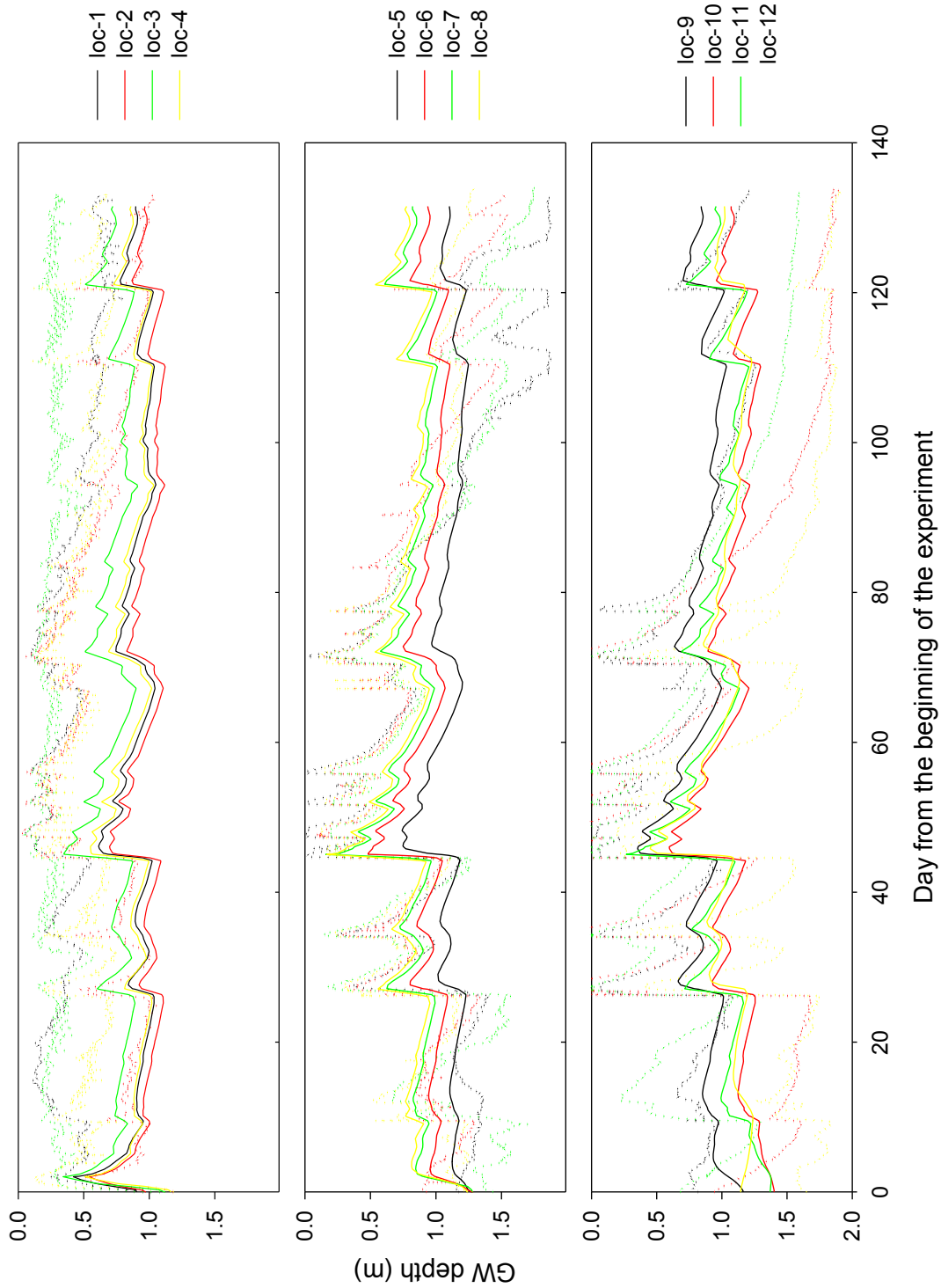
Figure 4-9   Observed (dots) and HYDRUS-3D simulated (continuous lines) groundwater depth time series.

Table 4-2   Statistics of observed and simulated subsurface Cl- transport

Well #	Peak time (day)		Peak concentration (ppm)	
	observed	simulated	observed	simulated
1	4.7	16	205.6	109
2	27.3	9	180.3	129
3	7.9	9	316.4	210
4	1.6	12	284.0	145
5	98.3	110	68.8	59.1
6	108.3	110	46.8	47.3
7	90.3	87	67.0	59.9
8	120.3	-	39.3	-
9	105.3	-	33.0	-
10	98.3	-	30.5	-
11	105.3	-	46.2	-
12	122.3	-	33.8	-
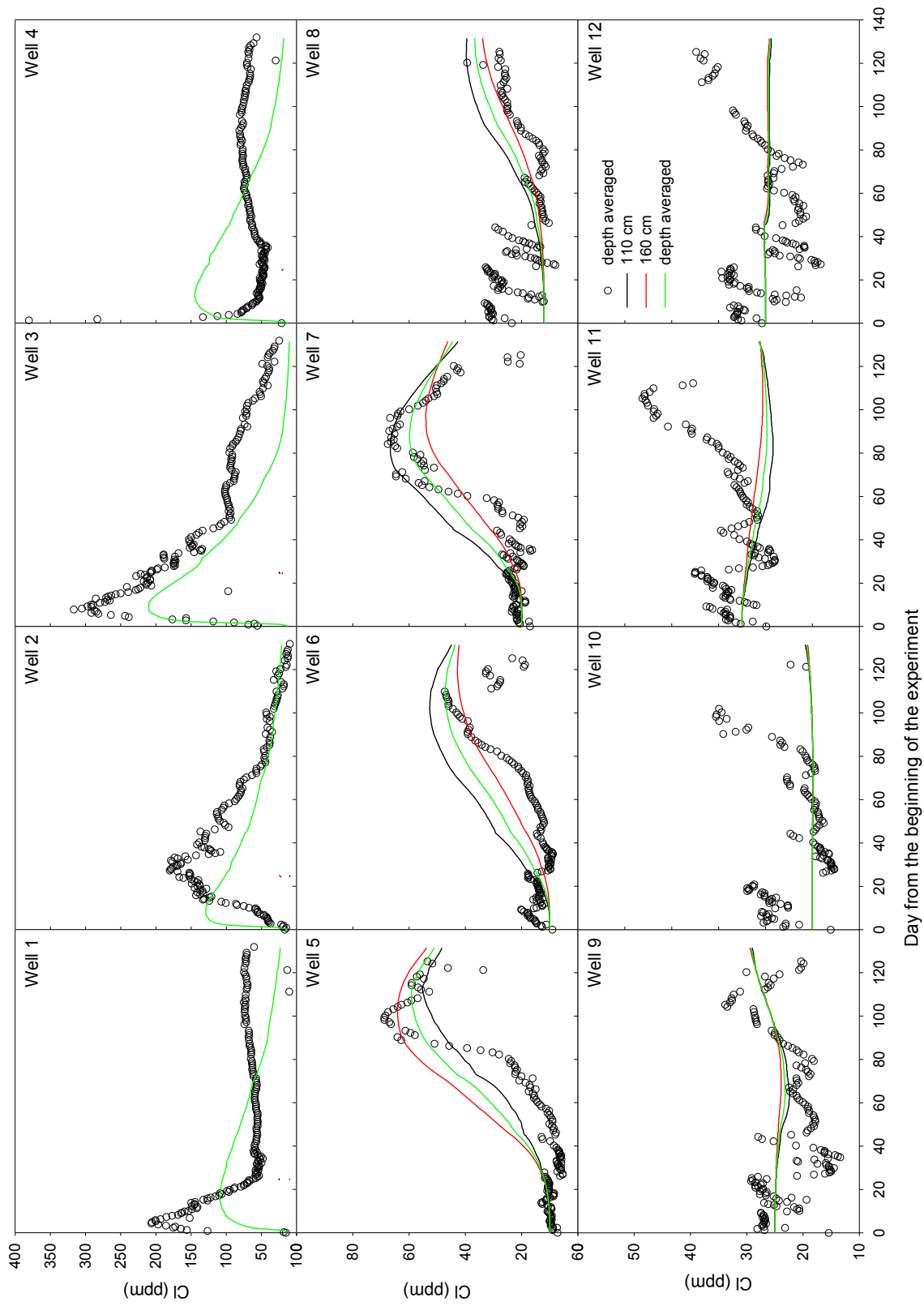
Figure 4-10 Observed (dots) and HYDRUS-3D simulated (continuous lines) Cl- concentrations in wells L1-L12.

4.2.3.2 TOUGH2 Simulations

The structure of simulated domain and the its boundaries for TOUGH2 model were set similar to those used for HYDRUS-3D simulations. The simulation domain was discretized into a 3D computational grid generated using the integral finite difference method for TOUGH2 (Fig. 4-11). The grid had an average of 28 computational layers in the vertical direction, and 1,337 columns (or grid blocks) per layer), resulting in 42,900 grid blocks (or elements) and 133,319 connections in the grid system. Fine lateral cells (1m x 1m) were implemented in the vicinity of the irrigation plot and observation wells, while coarse lateral cells (2m x 2m) were used beyond the irrigated area. The thickness of each element varied from 5 cm to 30 cm depending upon location and the soil material distribution, resulting in a varying number of vertical elements at different locations. The thickness of vertical elements in the upper 1.65-m soil layer was about 10-15 cm, and about 20 cm below this depth. The EOS9 and T2R3D modules of the TOUGH2 family were used to simulate flow and tracer transport at the OPE3 site. Based on this adopted 3-D grid, one simulation took about 6 h with a Dell PC Precision 690, Intel® Xeon® CPU 5130 @2.00 GHz.

The soil hydraulic properties, the distribution of soil materials within the domain, and the initial and boundary conditions were set similar to the calibrated HYDRUS-3D model.
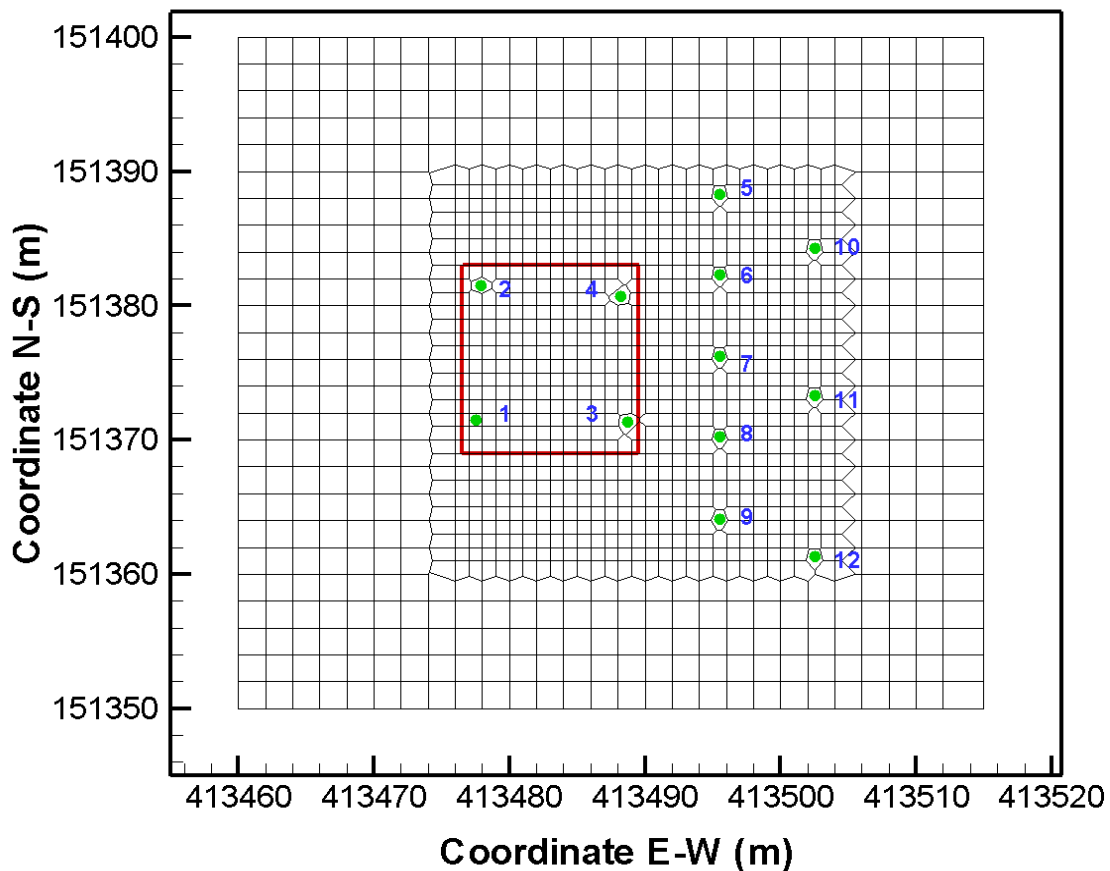
Figure 4-11 Mesh generated for the TOUGH2 simulations.

Twenty-eight observation nodes were assigned to record simulated breakthrough curves at depths corresponded to the sampling depths at the 12 locations. One observation node was set at each of four wells at the irrigation plot, while three observation nodes were placed at three depths in wells L5 through L12. Observation nodes corresponded closely to the observation depths used in the experiment.

Water flow was simulated using the EOS9 module, which is a saturated-unsaturated flow module of TOUGH2. Extensive tests with the EOS3 module, which describes coupled air and water flow, indicated instability in the results caused by changing pore air pressures. The EOS3 module was therefore replaced with the EOS9 module, which solves the Richard's equation for water flow only and does not consider the presence of an air phase. Results of the water flow simulations are shown in Fig. 4-12. Overall, the EOS9 module reproduced the dynamics of the groundwater depth slightly better than the HYDRUS-3D code. The model also appeared to be more sensitive to rainfall events and, as compared to HYDRUS-3D, produced more accurate simulations of the amplitudes of groundwater fluctuations and the receding groundwater levels observed in the summer months during the experiments. Root-mean-squared errors of groundwater levels were comparable for the two models. In wells 5 through 9, these RMSE were 0.43, 0.30, 0.40, 0.26, and 0.27 for HYDRUS 3D and 0.48,0.43, 0.36, 0.31, and 0.38 for TOUGH2 (all in m).

To simulate the tracer transport, EOS9 was coupled with the T2R3D module, which simulates radionuclide transport based on mass conservation and advection-dispersion equations as described in section 4.1. Observed BTCs and simulations with the TOUGH2 model are shown in Fig. 4-13. The model adequately reproduced the BTCs of wells L2, L3, L6 and L8. However, TOUGH2 failed to correctly predict receding Cl⁻ concentrations of wells L1 and L4 and the BTCs of wells L5 and L7. Similar to the HYDRUS-3D results, simulations of the plume did not reach the transect across wells L10-12 within the simulation time.

Overall, predictions of tracer transport was much better with the calibrated HYDRUS-3D model compared to the calibrated TOUGH2. HYDRUS-3D was more accurate in reproducing observed peak concentrations and peak times of Cl⁻ of wells L5-L8. The ARS study teamrealize that the obtained set of parameters may not have been unique and that additional simulations would be helpfull to reduce parameter uncertainty. Nevertheless, the ARS study teamaccepted the values and used them to evaluate the applicability of different model abstraction techniques to model simplification and monitoring strategy development.
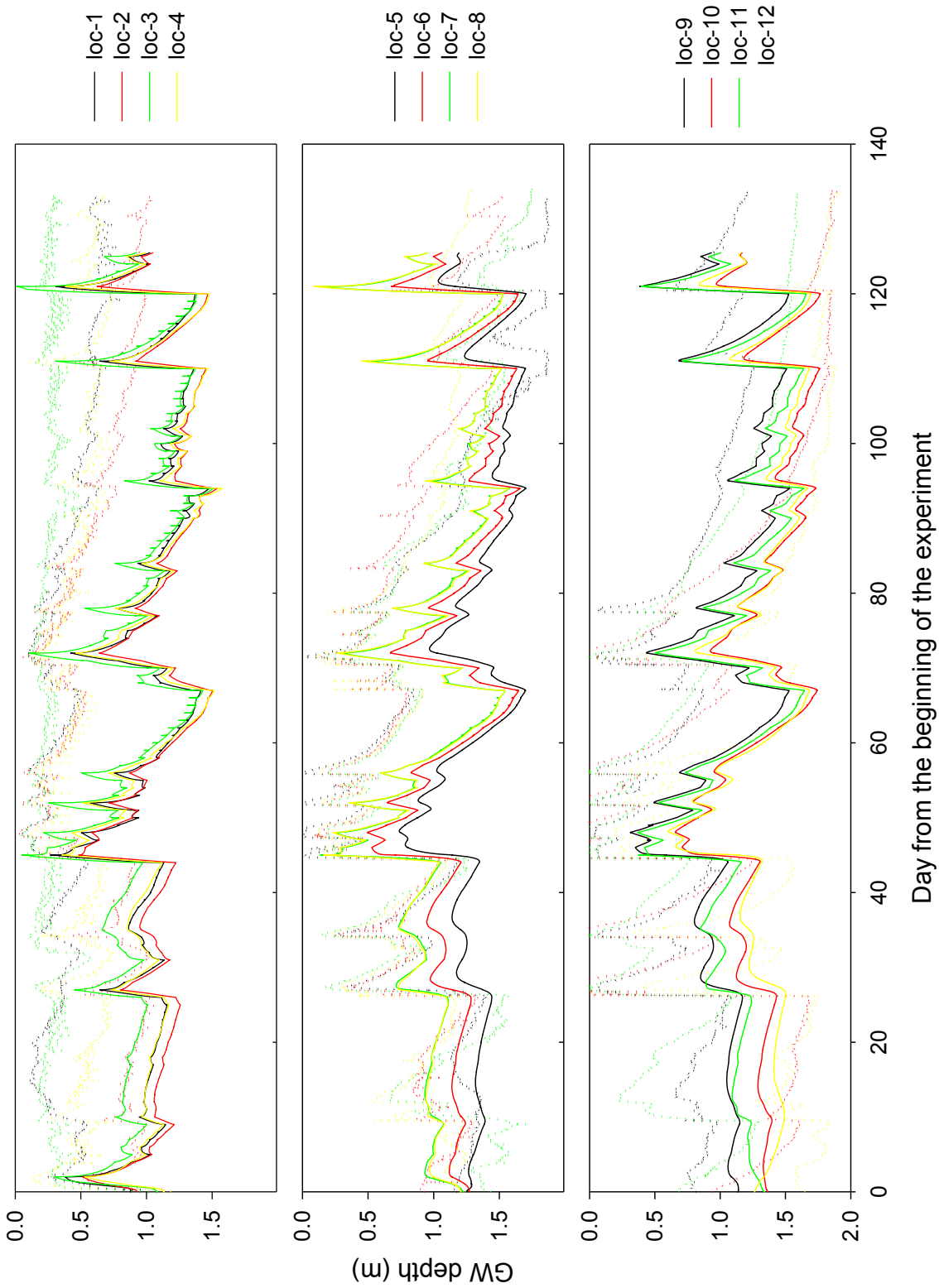
Figure 4-12 Observed (dots) and TOUGH2 simulations of groundwater depth time series.
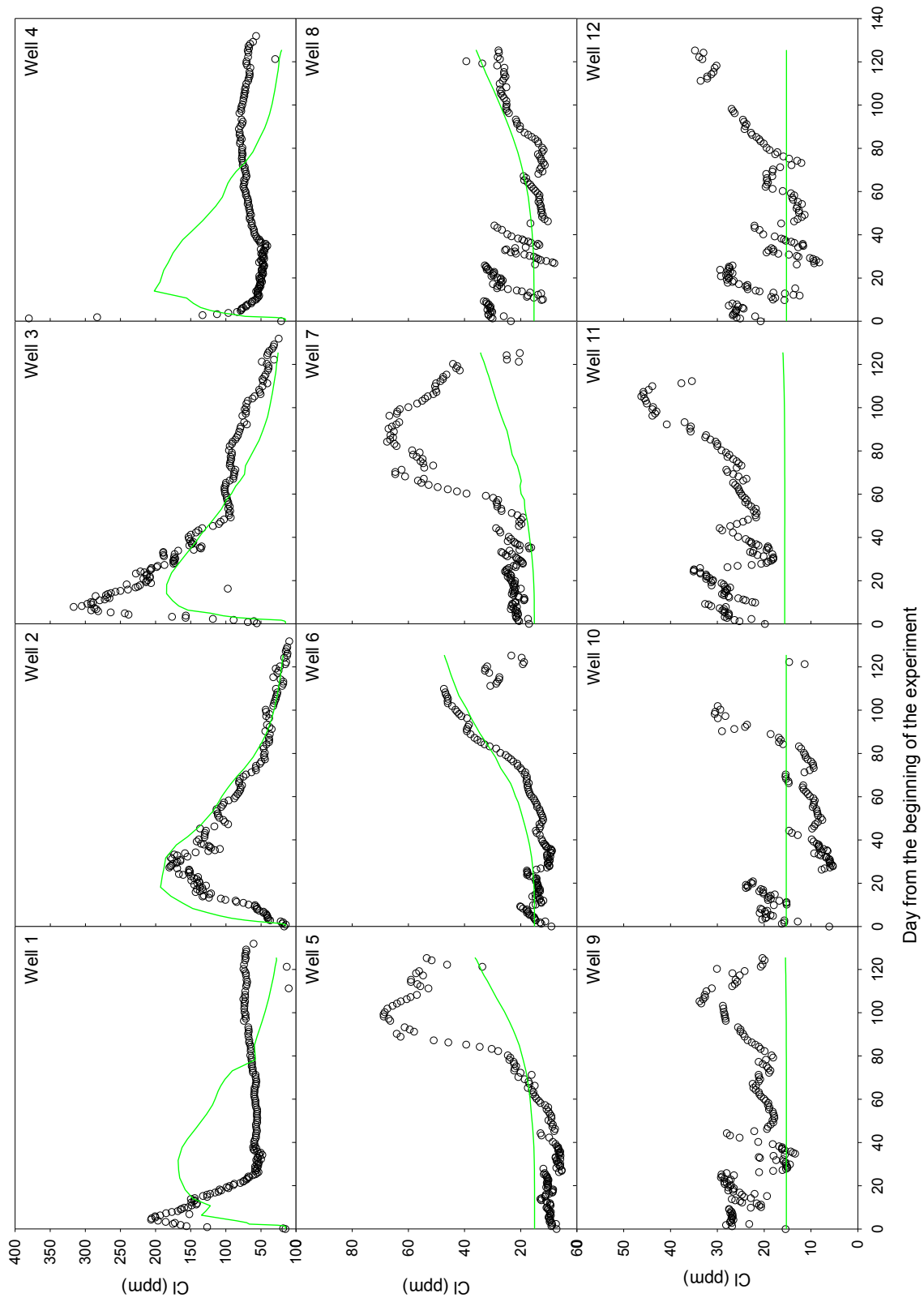
Figure 4-13 Depth-averaged observed (dots) and TOUGH2 simulated (continuous lines) Cl- concentrations of wells 1-12.

5   SYSTEMATIC APPROACH FOR SIMPLIFYING THE OPE3 SITE MODEL

5.1   Sensitivity Analysis to Direct Model Simplification

Conducting a sensitivity analysis (SA) is an efficient methodology for model simplification. The SA approach for model simplification was first proposed by Rose and Harmsen (1978), and has since proven its efficiency in various modeling fields (e.g., Confalonieri et al., 2010; Brooks et al., 2010, Do and Rothermel, 2008; Cariboni et al., 2007). The SA has a long history of use in radionuclide transport modeling, beginning with U.S. NRC-supported effort (e.g., McKay, 1995) and including research in various parts of the world (e.g., Volkova et al., 2008; Ohi et al., 2010; Iooss et al., 2008; Helton et al, 2010).

Overall, two types of SA are recognized (Saltelli et al., 2010): a local sensitivity analysis and a global sensitivity analysis. A local SA examines the local response of the output(s) by varying input parameters one at a time while holding other parameters at central values. A global SA examines the global response (averaged over the variation of all the parameters) of model output(s) by exploring a finite (or even an infinite) region. While easier to implement, a local SA can only inspect one point at a time, with the sensitivity index of a specific parameter depending on the central values of the other parameters. The approach uses a local SA technique (LSA) when the sensitivity index measures only the main effect of each input parameter on the output variance. The method does not account for interactions among parameters.

In the study, the ARS team used two indices in a LSA: a first order sensitivity coefficient (Saltelli et al., 2010) and a derivative-based measure of sensitivity (Saltelli et al., 2005).First order sensitivity coefficients estimate the variance-based first order effect

$$V_{Xi}\left(E_{X\sim i}(Y\,|\,X_i)\right) \tag{5-1}$$

of a generic factor X_i for model prediction Y written in the form:

$$Y = f(X_1, X_2, ... X_k) \tag{5-2}$$

where X_i is the i-th factor and $X_{\sim i}$ denotes the matrix of all factors except X_i, and k is the number of factors. The meaning of the inner expectation operator E in Eq.(5-1) is that the mean of Y is taken over all possible values of $X_{\sim i}$ while keeping X_i fixed. The outer variance V in Eq.(5-1) is taken over all possible values of X_i. The associated sensitivity measure (the first-order sensitivity coefficient) is:

$$Si = \frac{V_{Xi}\left(E_{X\sim i}(Y\,|\,X_i)\right)}{V(Y)} \tag{5-3}$$

where $V(Y)$ is the total variance of the model output computed as:

$$V(Y) = V_{Xi}\left(E_{X\sim i}(Y\,|\,X_i)\right) + E_{Xi}\left(V_{X\sim i}(Y\,|\,X_i)\right) \tag{5-4}$$

The first-order sensitivity coefficient Si is a normalized index varying between zero and one, as $V_{Xi}\left(E_{X\sim i}(Y\,|\,X_i)\right)$ varies between zero and $V\,(Y)$.

The second LS index used in this study is a derivative-based sensitivity index. This index is a straightforward implementation of the sensitivity concept: if the model output of interest is Y, its sensitivity to an input factor X_i is simply $\partial Y/\partial X_i$. This measure shows how sensitive the output is to a perturbation of the input. If a measure independent from the units used for Y and X_i is needed, then the normalized value of the model sensitivity can be used:

$$Di = \frac{\overline{X}_i}{\overline{Y}_i}\frac{\partial Y}{\partial X_i}$$

(5-5)

where $\overline{X}_i$ is the nominal value of factor X_i and $\overline{Y}_i$ is the value of Y when all input factors are at their nominal values.

5.1.1 Analysis of Subsurface Flow and Transport Processes

The HYDRUS-3D and TOUGH2 model calibrations based on results of the tracer experiment revealed considerable hydrologic complexity of the flow system. To understand better the behavior of the system, water and Cl⁻ fluxes were analyzed in both the saturated and unsaturated zones of the simulation domain.

Ten new observation nodes were assigned at depths of 0.4-0.7 m in addition to the 20 nodes used for the HYDRUS-3D calibration and shown in Fig. 4-2 to represent both the saturated and unsaturated parts of the soil profiles at the 12 observation locations. The HYDRUS-3D model was run with the parameters, initial conditions, and boundary conditions used for the calibration (Table 4-1). Soil water content (θ) dynamics obtained with HYDRUS-3D simulations indicated that the upper 0.5-0.7 m of the soil profiles were not saturated, while the soil was saturated below these depths during the entire simulation period (Fig. 5-1). Values of θ were generally higher at a depth of 1.1 m at locations 1, 2 and 4 than at the same depth and deeper at locations 3, and 5 through 12. Differences in θ values were caused by the material distribution in the simulation domain. The saturated water content θ_s among the soil materials was highest (0.5 cm³ cm⁻³) in SL2 that formed the soil layer at a depth of 1.1 m at locations 1, 2 and 4 (Table 4-1). Soil layers at the same depth at the other locations were formed by the SL3, L1, L2, SCL1 and SCL2 materials (Fig. 4-5), with θ_s values ranging from 0.32 to 0.38 cm³ cm⁻³ (Table 4-1).

Differences in material properties forming the simulation domain resulted in differences in simulated concentrations of the Cl⁻ tracer at locations with similar initial and boundary conditions. Within the irrigated plot (wells 1-4), the Cl⁻ tracer arrived at a depth of 0.6 m in well 3 and well 4 at about the same concentration 24 hours after the tracer application (Fig. 5-2). Tracer transport through the SL3 layer at these locations occurred in unsaturated conditions, and both the water content and the saturated hydraulic conductivity (K_s) affected Cl⁻ movement. The velocity of unsaturated vertical Cl⁻ transport was four times smaller than the K_s value of this layer. The Cl⁻ concentration in well 3 was higher than in well 4 when the tracer reached a depth of 1.1 m. Water contents at this depth were close to saturation in both wells, with Cl⁻ transport being controlled mostly by the saturated hydraulic conductivity, which was seven times higher in L2 (well 3) compared to the SL2 material (well 4) at these depths (Table 4-1). Therefore, the slower vertical Cl⁻ transport rate between depths 0.6 and 1.1 m in well 4 as compared to well 3 was attributed to the difference in the saturated hydraulic conductivities of the SL2 and L2 materials at a depth of

1.04 m at these two locations (Fig. 4-6). The Cl$^-$ breakthrough curve at a depth of 1.1 m in well 2 was similar to that of well 4. Cl$^-$ transport at this depth was controlled by the SL2 material, which had the lowest K_s value (0.1 m d^{-1}) of the different soil profiles at both locations.

Downward Cl$^-$ transport within the irrigated plot (wells 1-4) changed to lateral transport soon after the tracer approached the saturated zone with its highly conductive layers L1 and SL1. Lateral transport was relatively slow: it took from 9 to 16 days for the tracer to pass the 1.1-m depth at the irrigated plot and from 87 to 110 days to arrive at wells 5, 6 and 7 which were located at the distance of 7 m from wells 3 and 4 at the irrigated plot (Fig. 5-2). Based on the arrival time, the average flow velocity was about 0.12 m day^{-1} in the vertical direction and 0.08 m day^{-1} in the horizontal direction within and outside the irrigated plot, respectively.

The difference in flow velocities occurred due to a combination of flow conditions and hydraulic properties of the materials in the transport domain. Lateral flow was driven by the smaller (about 0.03) lateral hydraulic gradient as compared to the vertical gradient (about 1), but the saturated conditions for horizontal flow vs. the partly saturated conditions for vertical flow, and higher K_s values of the SL1 and L1 material (2.0-2.5 m day^{-1}) at locations 5-7 compared to SL2 (K_s=0.1 m day^{-1}) at locations 3 and 4, partially compensated for the difference in the hydraulic gradient. The ratio between the vertical and horizontal flow velocities was less than 3:1 for these reasons. The saturated conditions and high K_s values were not sufficient for the tracer to approach wells 10-12 within the simulation period of time. The ARS study teamdid not observe any increase in Cl$^-$ concentrations at these wells during the 132-day long simulations. Based on the flow velocity, the tracer arrival time at the transect across wells 10-12 was expected to be about 175th day after initiation of the irrigation experiment.

A decrease in Cl$^-$ concentrations with time was observed at the locations of wells 9, 11 and 12. The decrease was more pronounced at a depth of 0.6 m as compared to the deeper layers. Two processes could cause this decrease. One possible reason is that the 0.6 m depth layer was in the unsaturated zone (Fig. 5-1), with the infiltration from the top soil layer delivering precipitation with a Cl$^-$ concentration of 8 ppm, which diluted the 25 ppm Cl$^-$ concentration at this depth. Another possible reason was lateral flow from wells 3, 5 and 6 locations that may have decreased the Cl$^-$ concentration at both 0.6 m and 1.2 m depths. The initial concentration of Cl$^-$ was only about 10 ppm in the soil at these locations. The water flux from these locations hence may have had a diluting effect on the more concentrated solutions in wells 9, 11 and 12.

Lateral (q_x and q_y), vertical (q_z) and total (q) water fluxes and Cl$^-$ concentrations (C) computed with the HYDRUS-3D code for the 30 observation nodes were used to evaluate total vertical and horizontal water and tracer fluxes through 12 locations during the 132-day simulation period. The lateral (Q_x, Q_y, QC_x and QC_y) and vertical (Q_z and QC_z) components as well as the total water and Cl$^-$ fluxes in different parts of the simulation domain were computed as:
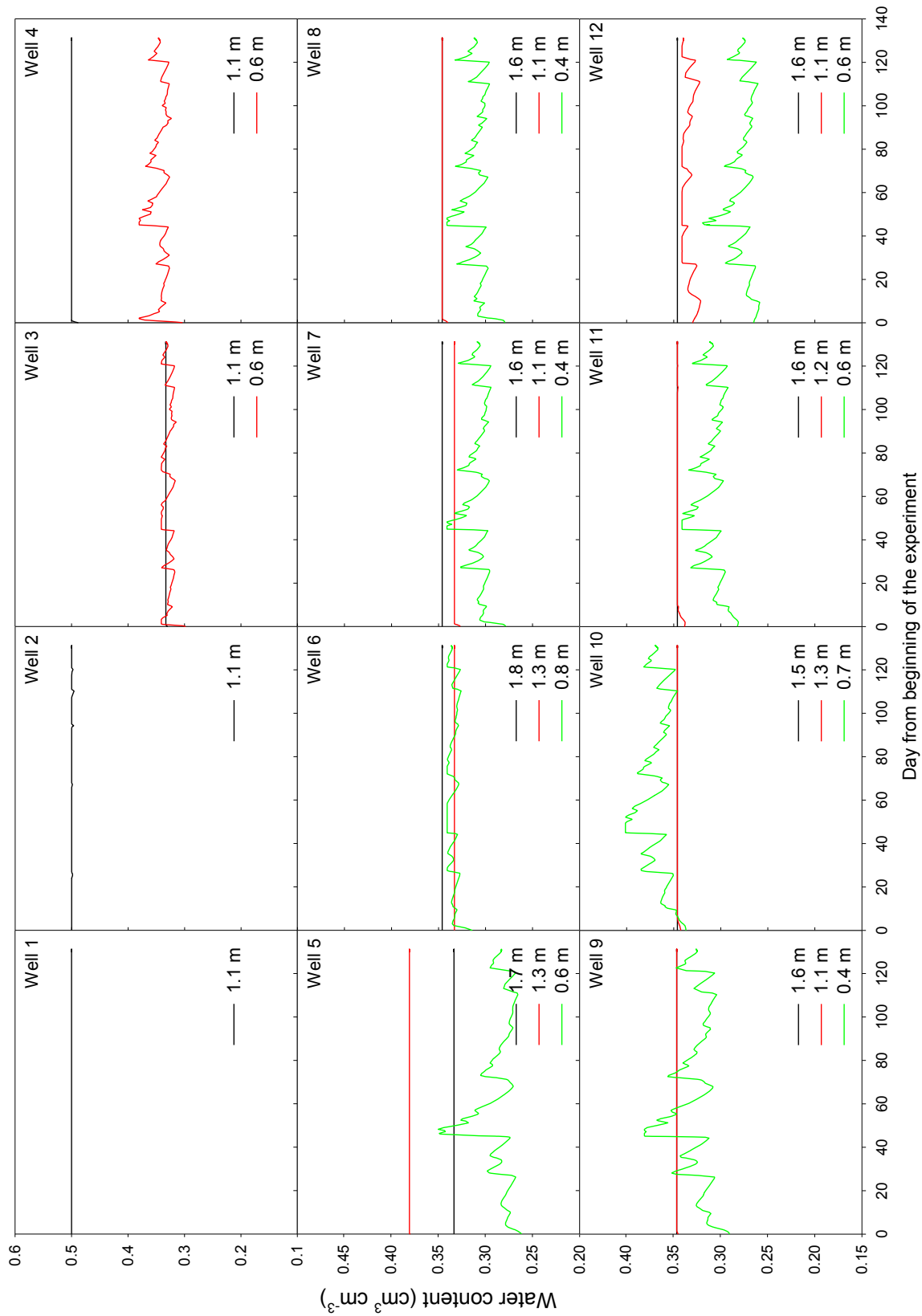
Figure 5-1   HYDRUS-3D simulated soil water contents at 12 locations. Constant values are where soil remains saturated.
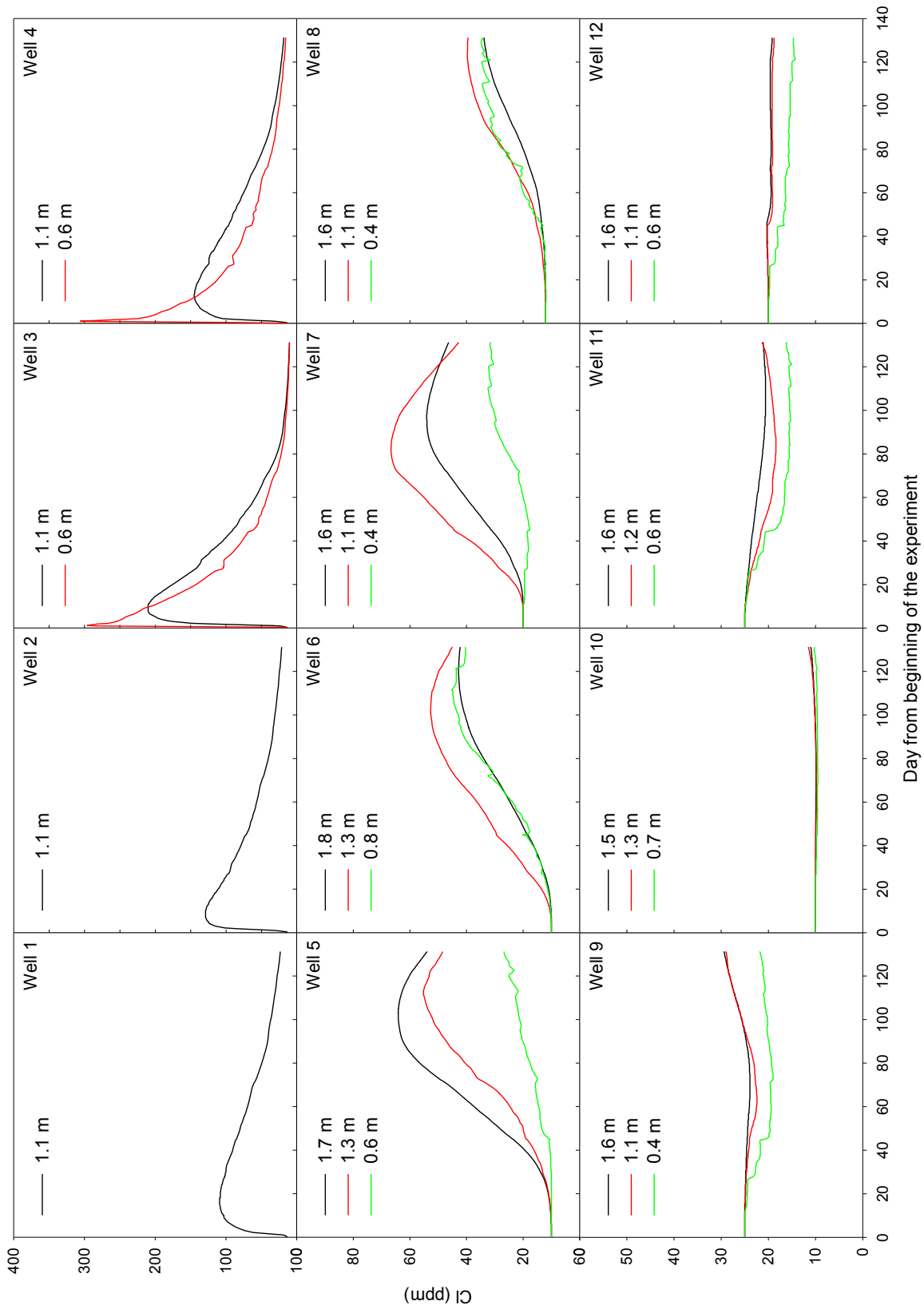
Figure 5-2   HYDRUS-3D simulated Cl- concentrations at 12 locations.

$$Q_j = \sum_{i=1}^{n} q_j \Delta t_i$$

<div align="right">(5-6)</div>

$$QC_j = \sum_{i=1}^{n} q_j C \Delta t_i$$

<div align="right">(5-7)</div>

where q_j are the computed water fluxes at the observation nodes (m day^{-1}), the index j refers to the flow direction (x,y or z), and Δt_i is the time increment in the HYDRUS-3D output (day). These fluxes can be viewed as total water and Cl$^-$ masses passing through 1 m^2 area during the simulations.

The computed water and tracer fluxes are shown in Fig. 5-3. The smallest total water fluxes (Q) within the irrigated plot were obtained at locations 1 and 2 at a depth of 1.1 m, while the largest Q values occurred at locations 3 and 4 at depths of 1.1 m and 0.6 m respectively (Fig 5-3d). Relatively small Q values at locations 1 and 2 were caused by the low hydraulic conductivity of the unsaturated SL2 and SCL1 materials in the upper 0.4-m layer of the soil profile. Saturated flow conditions and relatively high K_s values of the textures throughout the soil profile at location 3 resulted in the large total water fluxes at a depth of 1.1 m at this location (Fig 5-3d). Outside the irrigated plot, total water fluxes were generally larger in the first transect across locations 5-9 as compared to locations 10-12, except the well 9 where the total flux did not differ from the values of wells 10-12 (Fig. 5-3d). The differences in the total water fluxes observed for the two transects were caused by radial water flow from the irrigated plot. Figure 5-3 shows that the vertical component of the water flux Q_z (Fig 5-3c) was much smaller than the two horizontal components Q_x and Q_z (Fig 5-3a and Fig 5-3b) outside of the irrigation area. The contribution of the vertical component Q_z to the total flux Q was hence negligible compared with the two horizontal components. Lateral water flow eventually did approach the second transect. The Q values in wells 10 through 12 were close to their Q_x values (Fig. 5-3a and Fig. 5-3d), indicating prevalence of water flow from the irrigated plot in the x-direction towards the observation well compared to the y-direction.

The observed patterns in the total water flux translated into simulated total Cl$^-$ fluxes, QC. Within the irrigated plot maximum QC values were obtained at locations 3 and 4 as a result of high Q values (Fig. 5-3h). Beyond the irrigated plot, QC values were highest at locations 5 through 8 due to lateral flow in both the x and y directions. The ratio between the horizontal (QC_x or QC_y) and vertical (QC_z) components differed for these wells. The QC_y component for well 5 contributed more than 50% to the total Cl$^-$ flux QC, while locations 6 and 7 were more affected by the QC_x component. The reasons for this was the high heterogeneity manifested in the variability in the saturated hydraulic conductivity of the materials composing the simulation domain.

Large values of the total water flux did not necessarily lead to high tracer fluxes. The correlation coefficient between QC and Q in wells 3 through 8 was statistically insignificant (Fig. 5-4). The total Cl$^-$ fluxes in these locations were likely a combined result of the tracer fluxes from the
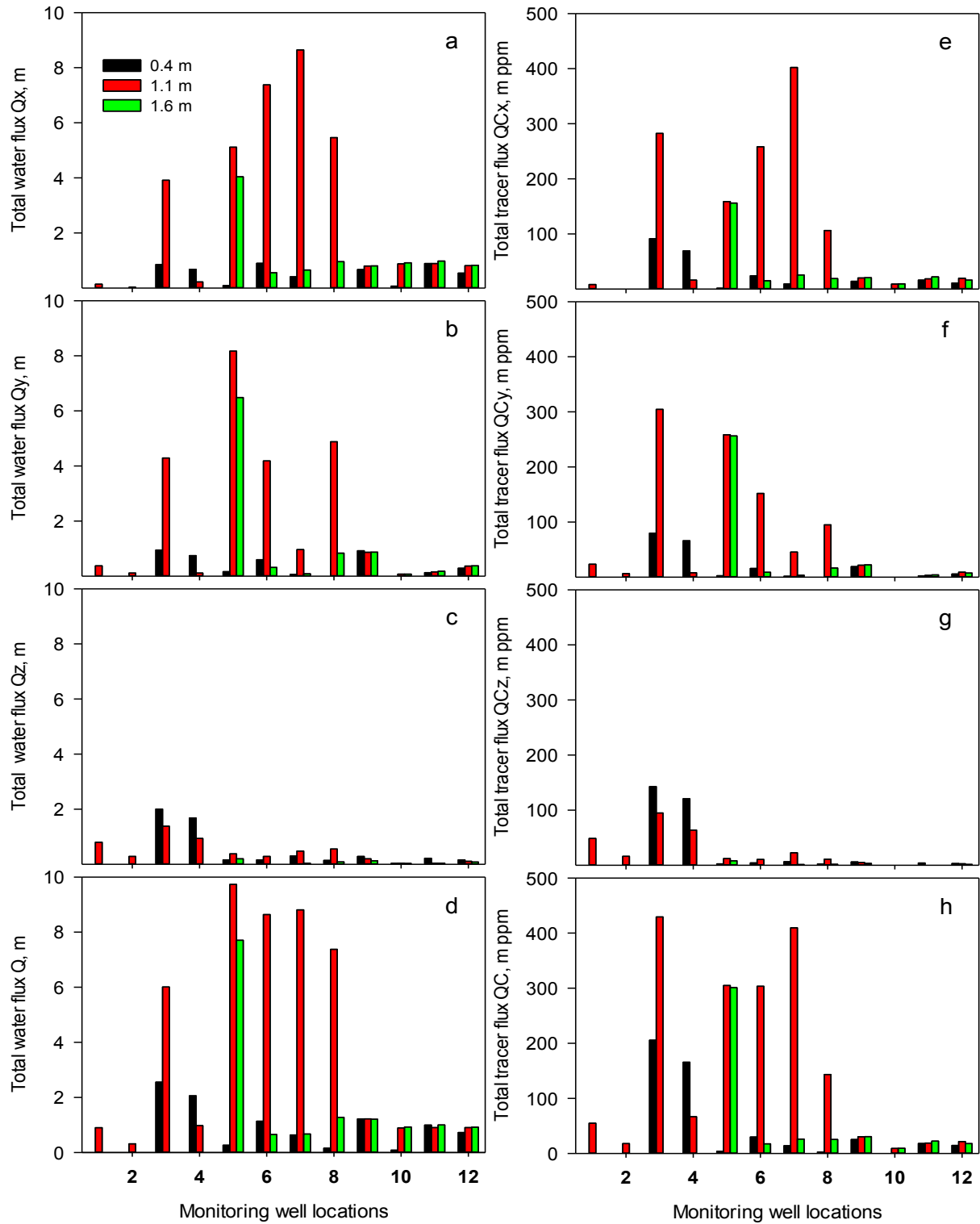
Figure 5-3   Total water (a-d) and Cl- (e-h) fluxes obtained in HYDRUS-3D simulations during 132 days at three depths in 12 monitoring locations.
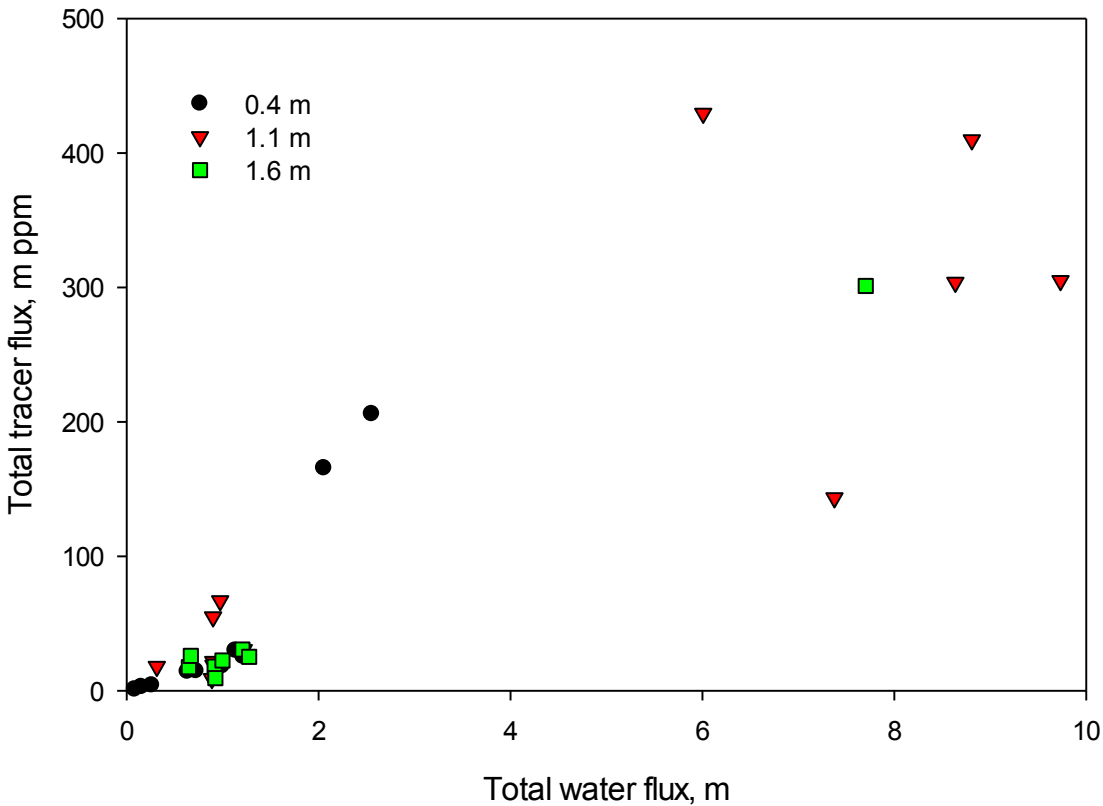
Figure 5-4   Relationships between total Cl- and water fluxes for four depths. Observation well numbers are shown above the symbols.

irrigated plot at the different depths. The considerable heterogeneity in the simulation domain caused high variability in the water and Cl- fluxes. This was the case even in the relatively small 10x10 m irrigated plot, which showed elevated QC values at depths of 0.4 and 1.1 m at location 3. The fluxes in the Northern locations (wells 5-7) were probably more affected by the tracer flux from location 3, while the Southern locations (wells 8 and 9) were affected by the flux initiated from location 4. This probably explains the relatively high *QC* values in the Northern locations as compared to those at the Southern side.

Overall, our analyses of the simulated time series of the water content and Cl- concentrations at 12 locations, along with total water and tracer fluxes and their components, revealed considerable hydrologic complexity in the simulation domain associated with spatial variability in soil properties. The presence of both saturated and unsaturated zones, fluctuating perched water tables, contrasting hydrological regimes in the irrigated and non-irrigated areas, and spatial variability in the soil properties, provided a unique opportunity to explore a range of model abstraction techniques, which will be demonstrated in the following chapters.

5.1.2 Performance Assessment Indicators

Sensitivity analyses were conducted to evaluate the applicability of different abstraction techniques to the model simplification. The simplifications refer to having distinct structural/textural units within the simulation domain and the fluxes on its boundaries. The goal of the sensitivity analyses was to simplify the model to a level where simulations still can depict key features of the modeled domain. To do this, the features should be defined explicitly. Standard Review Plan 2.4.13 (NUREG-0800) specifies pathways and travel times as the critical parameters for assessment of ground and surface water radionuclide contamination. For this reason the ARS study team used as performance indicators in their sensitivity analysis, peak concentrations (*Cpeak*) and times to reach peak concentrations (*Tpeak*). Because the peak concentration was not observed at some locations in HYDRUS-3D simulations and in most TOUGH2 simulations, the ARS study team also included entire Cl⁻ time series in the sensitivity analysis.

5.1.3 Transport Parameters

The sensitivity analysis was first used to evaluate applicability of the aggregation model abstraction technique. The ARS study team intended to reduce the number of the transport parameters by identifying materials which do not have a substantial effect on the HYDRUS simulations, and replacing them with materials having similar transport properties. To do this, the local sensitivity analysis (LSA) described in section 6.1 of this report was carried out. For the sensitivity analysis the ARS study team used the soil material distribution within the simulation domain, and the initial and boundary conditions obtained in the HYDRUS-3D calibration as described in detail in section 4.2. The local sensitivity analysis (LSA) was conducted first for the saturated hydraulic conductivity (K_s). Next, the effects of the longitudinal (α_L) and transversal (α_T) dispersivities on tracer transport were evaluated for 12 observation locations. Values of K_s, α_L and α_T of the soil materials used in the LSA were set to ±50% of their nominal (calibrated) values. Note that all parameters of soil materials SiCL2 and SiCL3 were kept constant during the LSA. The SiCL2 and SiCL3 materials served as restrictive layers at the bottom of soil profile. Parameters for these materials were not used in the LSA analyses to eliminate the effects of variations in boundary fluxes on the SA results.

As stated above, two values of each parameter were assigned (±50% of their nominal (calibrated) values) to eight soil materials (factors) resulting in 256 simulations using HYDRUS-3D and 256 simulations using TOUGH2 (2^8 simulations in a multivariate sensitivity analysis). This was done for each of the three parameters studied, K_s, α_L and α_T for a total of 256 x 2 models x 3 parameters = 1536 simulations. Results of the Cl⁻ simulations obtained for the minimum and maximum values of each factor were different for the two models. These differences were typically greater in wells 1-7 as compared to wells 8-12, and were found to increase when the Cl⁻ plume approached these wells.

The first-order sensitivity coefficients S_i were computed from the simulated Cl⁻ time series to assess the sensitivity of both models to the transport parameters for the eight soil materials. To do this, Cl⁻ transport was simulated using 256 combinations of K_s. Si values were computed subsequently for each of the eight soil materials and every output time. Specifically, for the minimum and maximum K_s values of material SL1 the ARS study team calculated the average concentration $E_{Ks\sim i}(C(z,t)\,|\,Ks_i)$ and variance $V_{Ks\sim i}(C(z,t)\,|\,Ks_i)$ at each output node for each output time t from the simulated Cl⁻ concentration time series. The index i is used here to denote the soil material. The same computations were then repeated for SL2, SL3, SCL1, SCL2, L1 and L2. This produced estimates of the variance $V_{Ks_i}\big(E_{Ks\sim i}(C(z,t)\,|\,Ks_i)\big)$ and average variance $E_{Ksi}\big(V_{Ks\sim i}(C(z,t)\,|\,Ks_i)\big)$

from the average concentrations $E_{Ks\sim i}(C(z,t)|Ks_i)$ and the variances $V_{Ks\sim i}(C(z,t)|Ks_i)$ computed for the two levels of K_s of each material. Finally, the first-order sensitivity coefficients S_i and total variance $V(C(z,t))$ were calculated using Eqs. (5-3) and (5-4), respectively.

An example of the S_i computations for the sensitivity of HYDRUS-3D simulation results to K_s of soil material SL1 is shown in Fig. 5-5. The S_i values varied in time and among locations. Generally the Si values were greater at locations where the material properties could affect the tracer transport process directly or indirectly. The direct effect of the SL1 material can be seen on the S_i time series for wells 1-5 in Fig. 5-5a. The SL1 soil material composed the soil layer at a depth of 0.62 m at locations 1 and 3, the soil layer at a depth of 1.04 m at location 2, and the soil layers at depths of 0.21, 0.41 and 0.62 m at location 5, but was not present in the soil profile at location 4 (Fig. 4-5). The maximum S_i values were for these reasons considerably greater for wells 1-3 and even for well 5 as compared to well 4, notwithstanding the cumulative tracer fluxes in well 4 being greater than in well 2 (Fig. 5-5). The indirect effect of material properties on the HYDRUS-3D sensitivity can be illustrated with Si values computed for wells 10-12. Note that the SL1 material was not presented in soil profiles at these locations. However S_i values for simulation times between 60 and 100 days were significantly greater for well 11 compared to the other two wells (Fig. 5-5a). The reason for this was the high conductive layer L1 near well 11, which at location 3 was connected with the SL1 material at a depth of 0.62 m, thus providing preferential lateral flow from location 3 towards well 11. Materials at locations 10 and 12 were not connected directly with the SL1 material. The properties of this material hence had little effect on results of HYDRUS-3D simulations at these two locations.
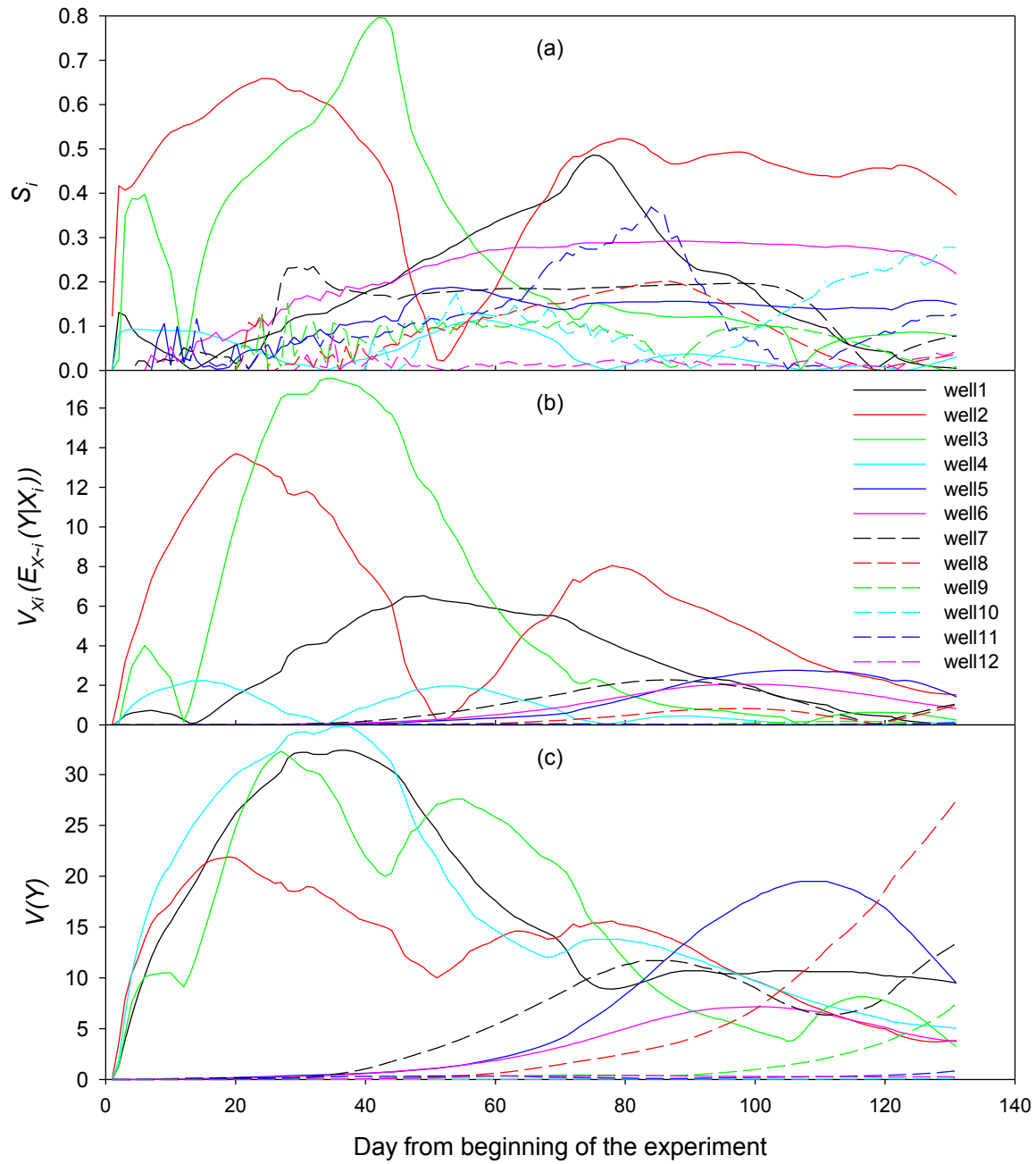
Figure 5-5   Time series of the sensitivity index Si along with its components obtained for HYDRUS-3D simulations using two levels of the saturated hydraulic conductivity of the SL1 soil material.

The sensitivity of HYDRUS-3D to K_s was different for the eight soil materials. Values of the Si index generally were larger for the SL1, SL2, SL3, L1 and SiCL1 materials, and smaller for SCL1, SCL2 and L2 (Fig. 5-6). The differences in the HYDRUS-3D sensitivities were caused by K_s values used in the LSA. The K_s values of the first group of materials were greater than K_s values of the second group, except for SL2 (Table 4-1). The high HYDRUS-3D sensitivity to the relatively small K_s value (0.3 m day^{-1}) of this material can be explained by its location. This material served as a restrictive flow layer at a depth of 0.41 m at locations 2 and 3, at a depth of 0.83 m at location 2, and depths of 1.04 m and 1.24 m at locations 2 and 4, and thus controlled downward transport of the applied tracer within the irrigated plot. The effect of this material can also be seen in wells 5-6, where SL2 was not present in the soil profile. These two wells received the dominant amount of mass of the applied tracer from the North-East corner of the irrigated plot (well 4), and therefore were affected indirectly by K_s of the SL2 material.

The sensitivity of the HYDRUS-3D modeling results to the material properties was also assessed using the derivative-based sensitivity index (D_i). D_i values were computed using Eq. (6-5) from the same time series as used for the Si computations. The parameters obtained with the HYDRUS-3D calibration and the Cl$^-$ concentrations simulated with these parameters were used in Eq. (6-5) as the nominal values of factors X_i and $\overline{Y_i}$, respectively. Computed HYDRUS-3D model time series of the Di index for K_s of the eight soil materials are shown in Fig. 5-7. Overall variability of the Di index was greater compared with S_i, with D_i values over time changing signs from positive to negative and vice versa. This occurred due to acceleration of the tracer flux with an increase in K_s, resulted in a shift of the Cl$^-$ breakthrough curves along the time axis such that smaller Cl$^-$ concentrations at $K_{s,min}$ corresponded to higher concentrations at $K_{s,max}$ in the rising limb, and vice versa in the declining limb of the Cl$^-$ breakthrough curves.

The calculated values of D_i indices confirmed the high sensitivity of HYDRUS-3D to K_s values of materials SL1, SL2, SL3, L1 and SiCL1 obtained in the analysis of S_i values. The sensitivity to the materials was generally higher for soil types that directly or indirectly affected tracer transport. However, unlike the S_i index, D_i values changed signs from positive to negative, when an increase in the factor produced a decrease in the Cl$^-$ concentration. Some discrepancy existed between the data in Fig. 5-6 and Fig. 10a, and data in Fig. 5-7. Both indices (S_i and D_i) indicated a high sensitivity of the HYDRUS-3D calculations to K_s values of SL1 for well 2 and well 3, and a low sensitivity for well 4. However, unlike the S_i index, D_i values did not indicate high model sensitivity for well 1.

The first-order sensitivity index (Si) is presumably a more informative measure of model sensitivity compared to the derivative-based sensitivity index (Di) since the term $V_{Xi}\left(E_{X\sim i}(Y\mid X_i)\right)$ in Eq. (6-3) accounts for the combined effects of X_i and all other factors. Unlike S_i, the D_i index estimates the effect of X_i factor when all other factors are fixed at their nominal values. Therefore, the D_i value for factor X_i may depend upon the nominal values of other factors. Nevertheless, Si can also mislead when the model sensitivity to all factors is low. In this case small values of the denominator in Eq. (6-3) may result in non-zero value of the Si index. Figures 5-5b and 5-5c illustrate this case. The variance of K_s of material SL1 did not affect results of HYDRUS-3D simulations. The Cl$^-$ concentrations in well 10 did not change during the simulation time (Fig. 5-2), thus leading to relatively small values of V_{Xi} for this parameter (Fig. 5-5b). However results of HYDRUS-3D simulations were about the same when K_s values of all materials varied, resulting in small values of $V(Y)$ (Fig.5-5c). Division of a small V_{Xi} value by a small $V(Y)$ gave Si a value about 0.3 on day 130 of the HYDRUS-3D simulation (Fig. 5-5a), which indicated a moderate sensitivity of the model to K_s of material SL1.

The S_i time series computed with the HYDRUS-3D and TOUGH2 models were used to assess the sensitivity of both models to K_s, and the longitudinal (D_L) and transverse (D_T) dispersivities of the eight soil materials. To reduce the amount of information and simplify the data analysis, the S_i time series were characterized by the mean, $\overline{S}i$, and the standard deviation, $SDSi$, of the S_i index over the simulation time. The relationships between $\sigma(\overline{S}i)$ and $\overline{S}i$ are shown in Fig. 5-8. Generally $SDSi$ increased with increase in Si value for all materials. Based on data in Fig. 5-8 HYDRUS-3D model was more sensitive to K_s and $\square_L$ than TOUGH2 model (Fig. 5-8a,b,d,e). The sensitivity to $\square_\square$ was about the same in both models (Fig. 5-8c,f). The smallest values of Si and $SDSi$ were obtained for SCL2 material in all locations indicating the low sensitivity of the HYDRUS-3D and TOUGH2 models to the transport properties in this material within the range of studied parameters. Therefore, based on Si index, the SCL2 material can be considered as a candidate for the model simplification.

The LSA was performed above based on the tracer concentrations representing the whole simulation period of time. Two performance indicators: simulated peak concentrations ($Cpeak$) and time of the peak concentrations ($Tpeak$) at the tracer time series will be considered here as an alternative to the entire Cl⁻ series. To evaluate the applicability of $Cpeak$ and $Tpeak$ performance indicators the ARS study team analyzed the Cl⁻ time series computed in the LSA of the HYDRUS-3D model. New indicators were not applicable for the data from wells 8-12 simulated with both models and for wells 5-7 simulated with the TOUGH2 model due to absence of peaks on the Cl⁻ curves. Therefore only the HYDRUS-3D results for wells 1-7 are used hereafter. The LSA was conducted using Si and Di indices computed both for the $Cpeak$ and

$Tpeak$ data. The nominal values of factors X_i and $\overline{Y_i}$ for D_i index were taken from the HYDRUS-3D calibration data, similarly to the previous LSA. Results were presented as relationships between the sensitivity indices computed for $Cpeak$ and $Tpeak$ data in wells 1-7. General increase in $Si(Cpeak)$ with increase in $S_i(Tpeak)$ with the slope close to one was observed in the simulations indicating approximately equal model sensitivity to both performance variables (Fig. 5-9a,b,c). Depending on location, the model sensitivity to the soil materials differed, however the values of $Si(Cpeak)$ and $Si(Tpeak)$ were consistently smaller in SCL2 material for K_s, $\square_L$ and $\square_\square$ than in other soil materials. The Di values indicated high sensitivity of the HYDRUS-3D model to K_s, moderate sensitivity to $\square_\square$ and low sensitivity to $\square_L$ in soil materials (Fig. 5-9d,e,f). Unlike Si index no relationship was found between $D_i(C_{peak})$ and $D_i(Tp_{eak})$. The values of D_i for K_s were positive and negative indicating both an increase and decrease in C_{peak} and T_{peak} associated with increasing value of K_s. The $D_i(C_{peak})$ values were mostly negative, but $D_i(T_{peak})$ were close to zero for the α_T parameter reflecting a decrease in the peak concentration with an increase in parameter of transversal dispersivity, while the peak time remained almost unchanged, which indicates a flattening of Cl⁻ BTC. Based both on S_i and D_i indices computed for peak concentrations and peak times, material SCL2 can be selected for model simplification.

Overall, introducing the peak concentration (C_{peak}) and the time to peak concentration (T_{peak}) as performance indicators for the LSA allowed us to reduce significantly the input information needed for the sensitivity analysis and to define the soil material that can be used for the model simplification example. One limitation of the new performance indicators is the need to have peaks in the simulated BTCs.
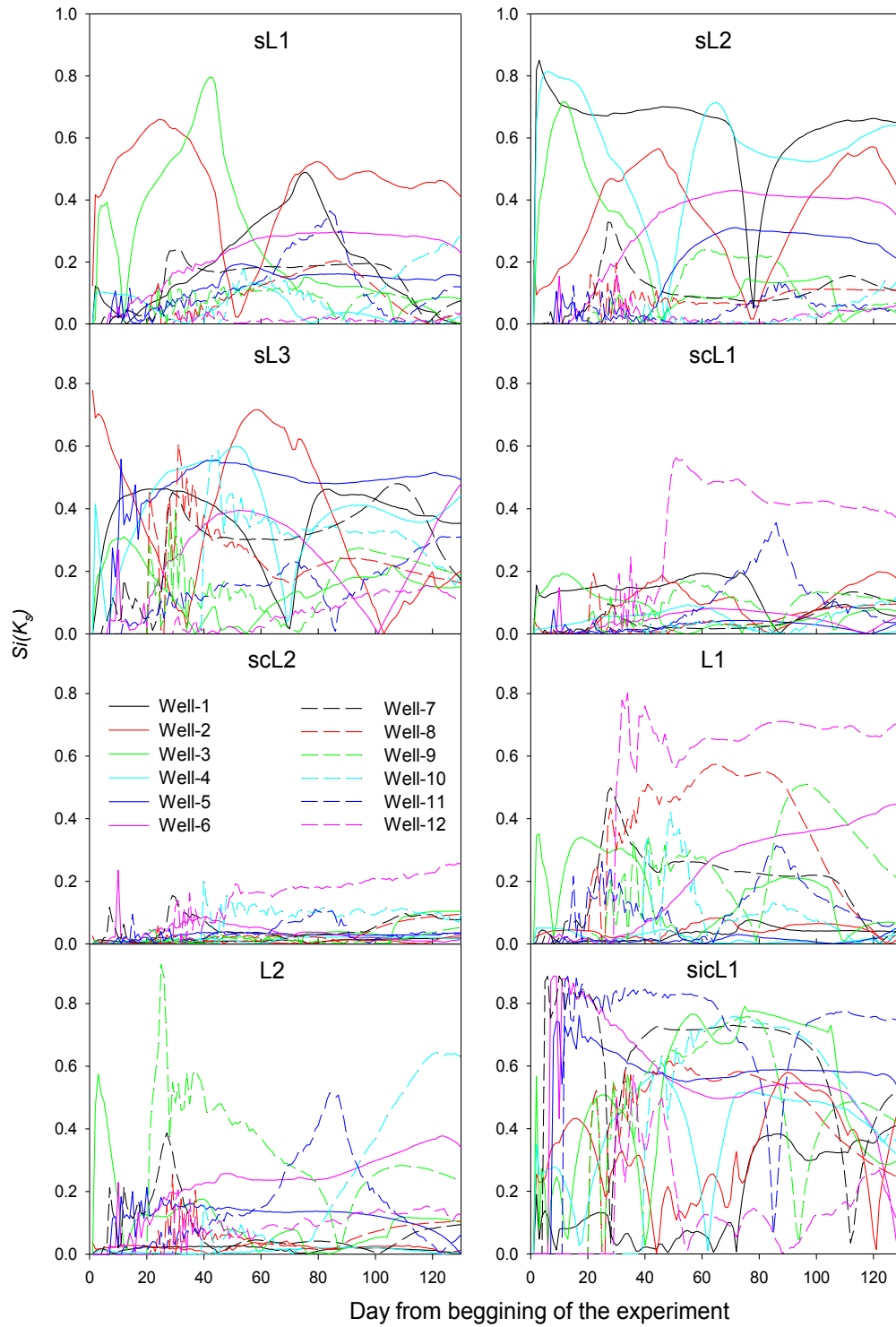
Figure 5-6 Time series of the sensitivity index Si obtained in HYDRUS-3D simulations with two levels of the saturated hydraulic conductivity for eight soil materials.
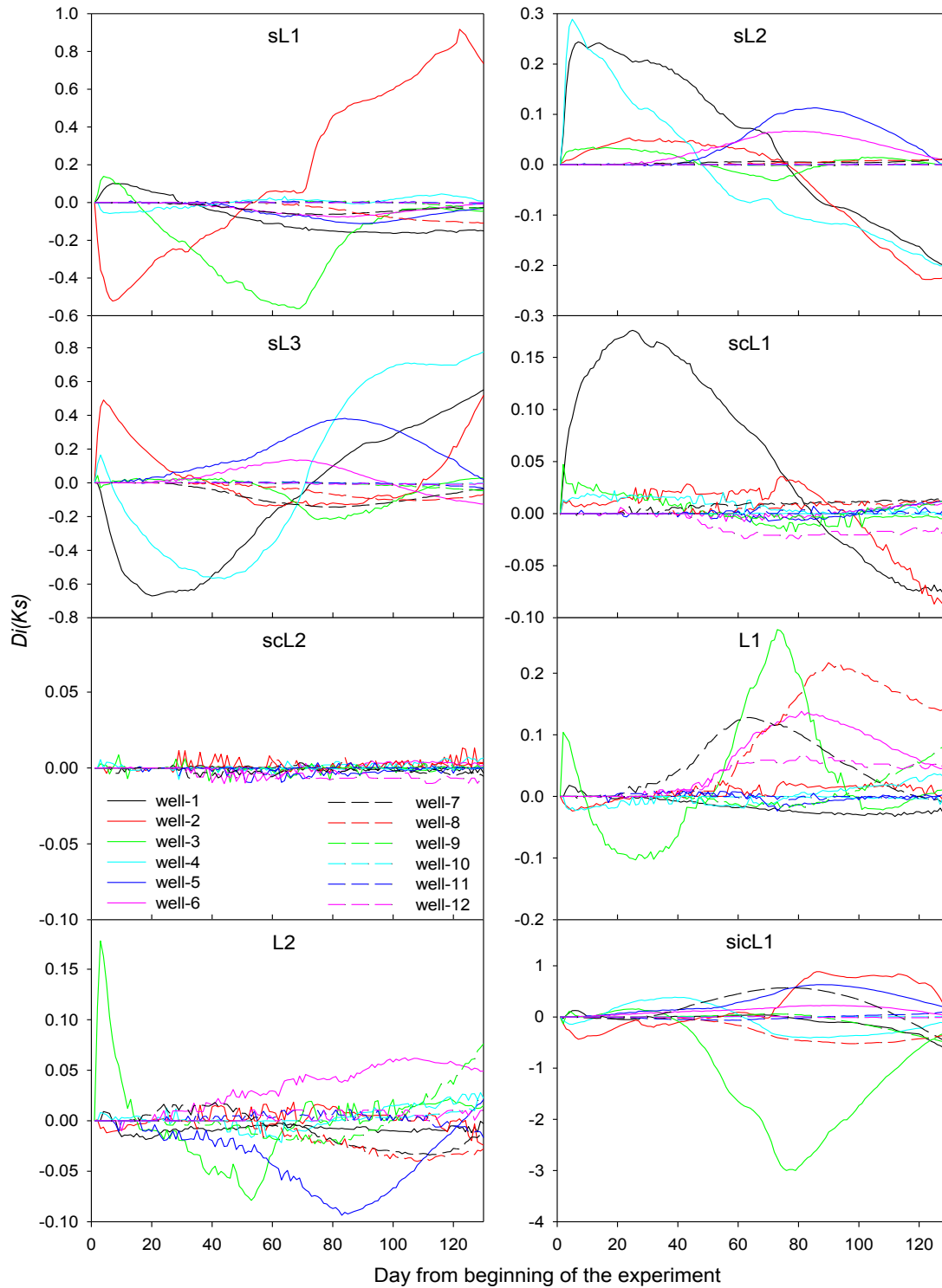
Figure 5-7   Time series of the sensitivity index Di obtained in HYDRUS-3D simulations with two levels of the saturated hydraulic conductivity for eight soil materials.
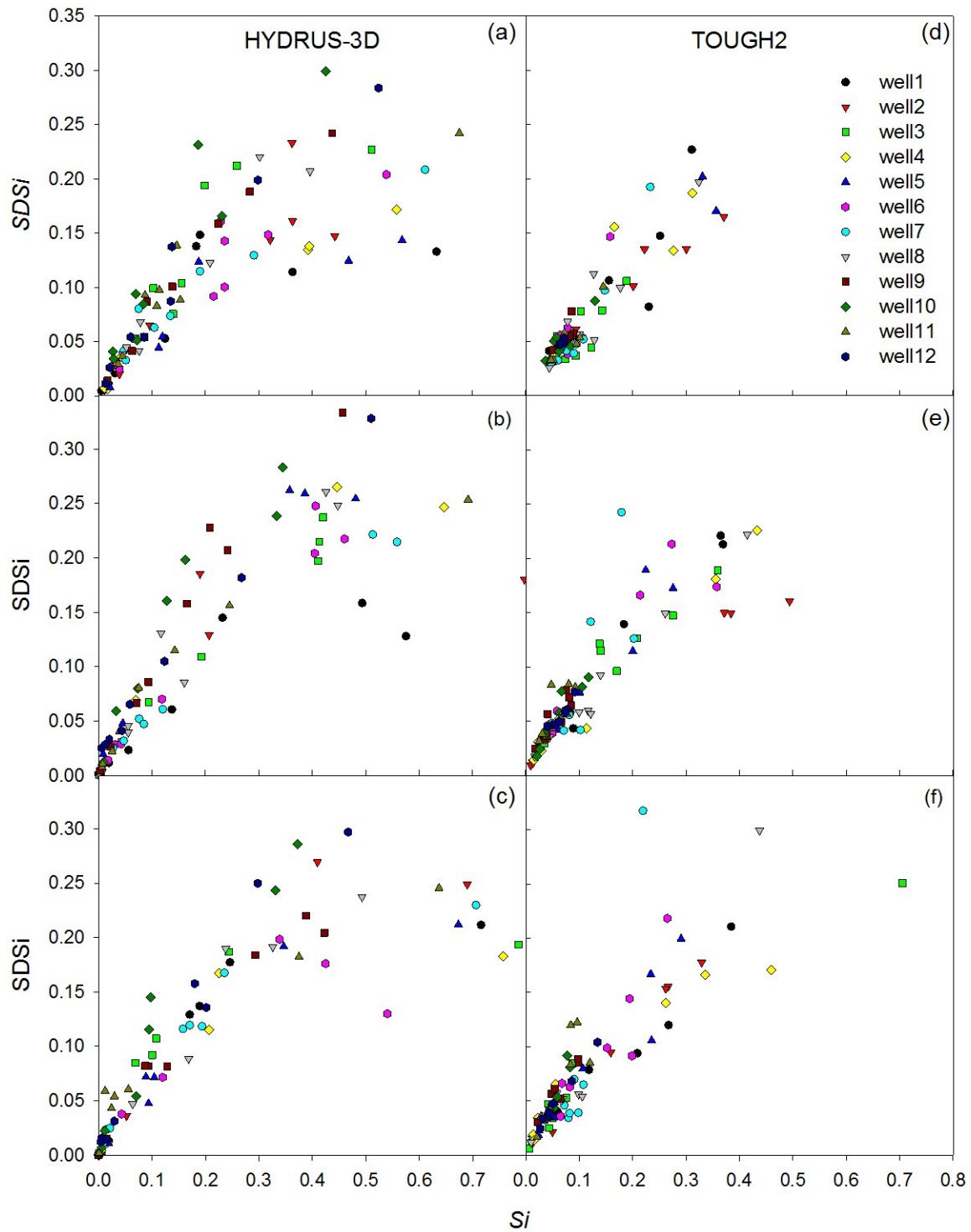
Figure 5-8   Sensitivity of the HYDRUS-3D and TOUGH2 models to the hydraulic conductivity (a, d), longitudinal (b, e) and transversal (c, f) dispersivity assessed for eight soil materials using time-averaged Si indices and their standard deviations SDSi.
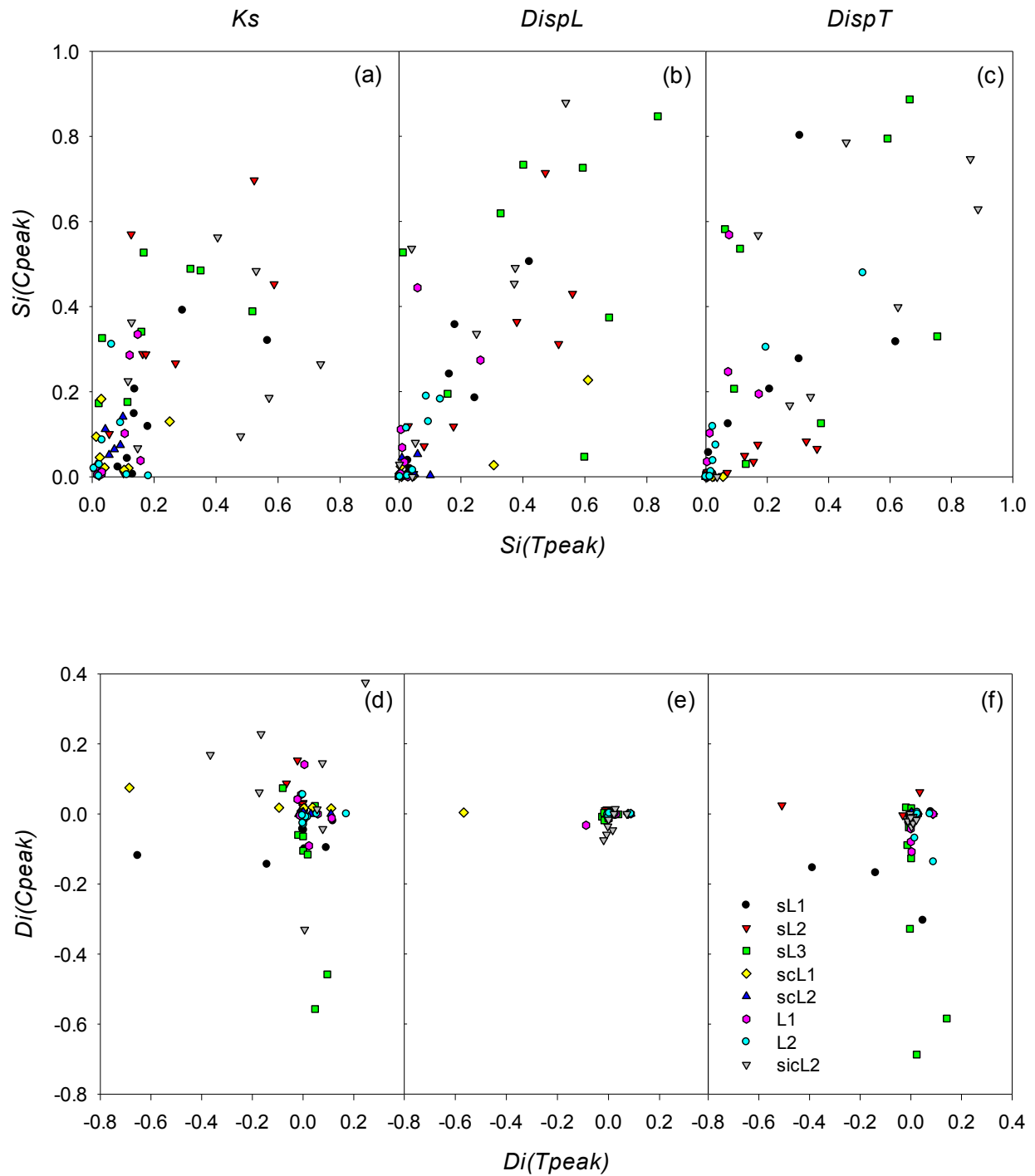
Figure 5-9 Sensitivity of the HYDRUS-3D model to the hydraulic conductivity Ks (a,d), longitudinal dispersivity αL (b,e) and transversal dispersivity αT (c,f) assessed at seven observation wells using Si and Di indices for simulated peak concentrations C_{peak} and times T_{peak}.

5.1.4 Subsurface Textural Units

The spatial boundary between soil materials was also subjected to the LSA. The materials in the simulation domain were vertically distributed within the soil profiles based on soil texture measured at 10 depths at 12 locations as described in section 4.2.1 of this report. The lateral boundaries between different textural materials were equally spaced between neighboring wells. Since no information existed about the actual locations of the various lateral material boundaries, it seemed worthwhile to assess the effect of material delineation on results of the tracer simulations. To do this, the lateral boundary of each soil material was extended or truncated for about 1 m. Only one material boundary out of nine was changed in each model run, resulting in a total of 18 runs for the HYDRUS-3D and TOUGH2 models. The model sensitivities to the material distributions were assessed using the derivative-based sensitivity index D_i. The first-order sensitivity coefficient S_i was inapplicable for this LSA since boundary locations were not independent variables for the different materials. The nominal values of factors X_i and $\overline{Y}_i$ in Eq. 5-5 were taken from the HYDRUS-3D and TOUGH2 simulations using the calibrated parameters.

The values of the sensitivity index D_i for both models varied greatly in time (Fig. 5-10 and Fig. 5-11). Modification of the material boundary caused at many locations an increase following by a decrease in the Cl⁻ concentration, and vice versa, with time, resulting in the D_i sign changing from positive to negative values and back. For some reason these changes were not seen in the TOUGH2 simulations, which produced D_i values which were either consistently positive or negative during the entire simulation (Fig. 5-11). Common for the two models was the low sensitivity to modification in the SCL2 material boundary. A possible explanation for this is the location of SCL2 in the simulation domain. The SCL2 material was primarily located at depths of 0.21 m and 0.41 m at locations 6 and 7. The ARS study team previously showed (section 5.1.1) that the tracer was transported from the irrigated plot towards wells 5-8 predominantly laterally in the upper layer of the saturated zone, while vertical water fluxes beyond the application area were insignificant. The SCL2 material was located far from the preferential fluxes and therefore did not affect their velocities and directions..

5.1.5 Boundary Conditions

The ARS study team previously examined the properties of the soil as a possible option for model simplification. Another possible way for model abstraction is simplification of the flow conditions on the boundaries of the simulation domain. The ARS study team used thus far daily-averaged precipitation and evapotranspiration rates for the top boundary of the entire, and daily-averaged irrigation rates less runoff in addition to the atmospheric boundary conditions within the irrigation plot. A zero flux boundary condition was imposed along the bottom of the domain. Spatially variable but constant in time pressure head profiles were specified along the lateral boundaries based on groundwater monitoring data obtained in the summers of 2006-2007. To evaluate the role of the boundary conditions in the model and to identify variables for model simplification, the LSA was applied separately to the top and lateral boundaries of the domain. For simplicity the ARS study team used the one-factor-at-a-time approach, thus excluding the combined effect of several factors on the LSA results. The D_i index was used as a measure of model sensitivity.
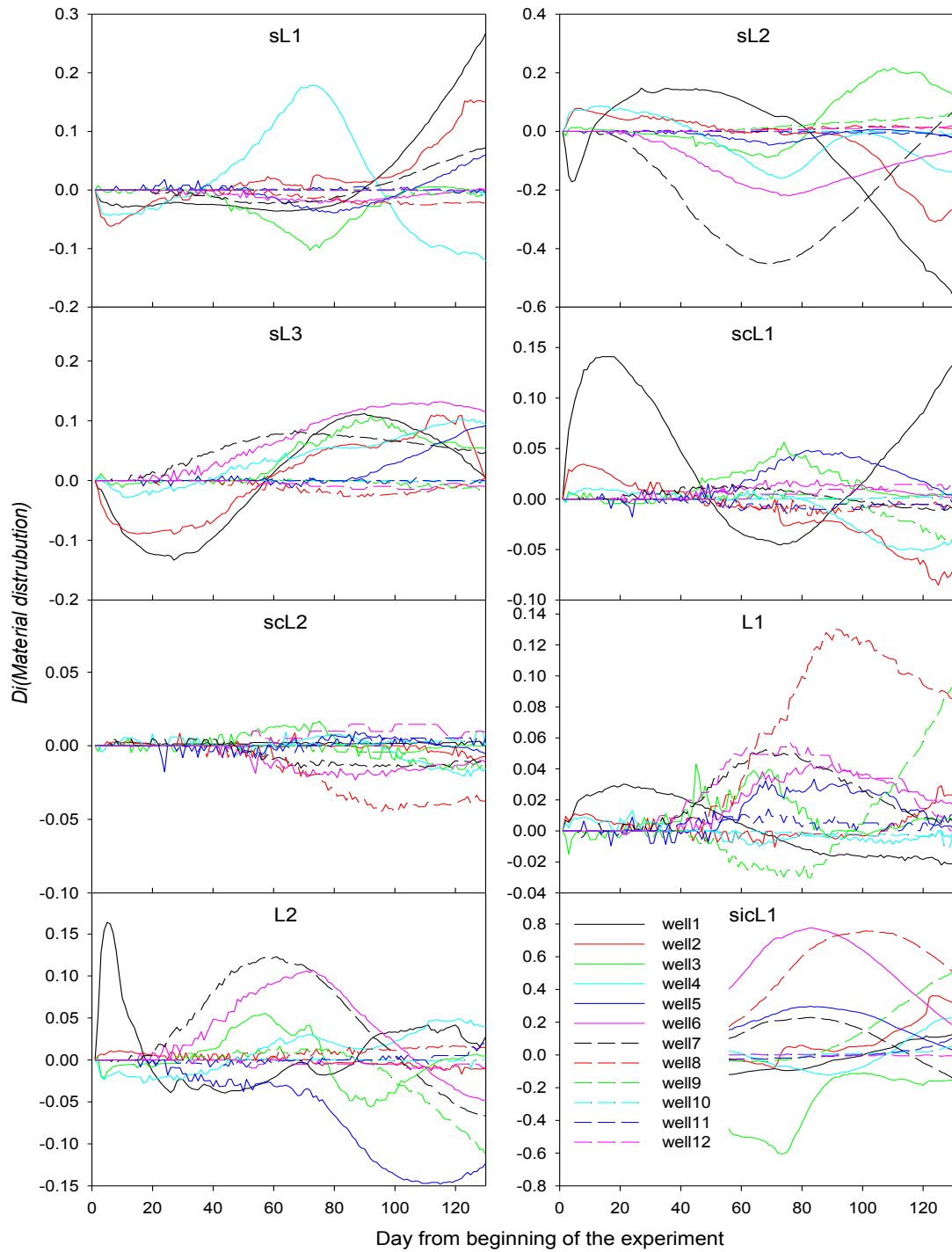
Figure 5-10 Temporal variation of the sensitivity index Di obtained in HYDRUS-3D simulations with extended and truncated material boundaries.
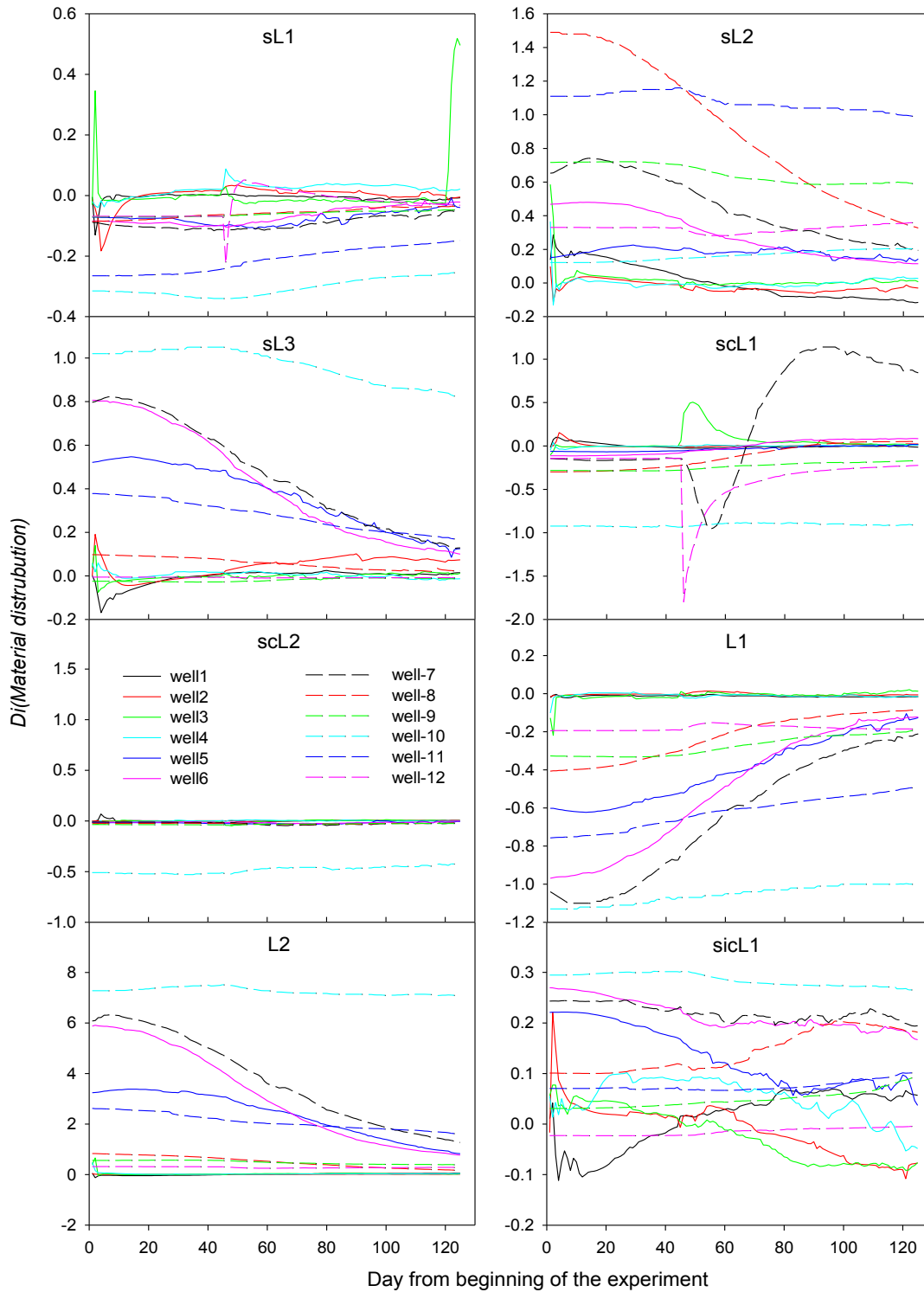
Figure 5-11 Temporal variation of the sensitivity index Di obtained in the TOUGH2 simulations with extended and truncated material boundaries.

The LSA was first applied to the runoff data. Runoff was experimentally estimated as 28% of irrigation and precipitation in our previous study (Pachepsky et al., 2011). For the LSA the ARS study team used 20% and 40% of irrigation and precipitation as estimates for runoff in both HYDRUS-3D and TOUGH2. Soil properties and the other boundary conditions were set similar to the HYDRUS-3D calibration simulation. The changes in runoff values dramatically affected the simulated Cl⁻ BTCs. A decrease in runoff resulted in a general increase in Cl⁻ concentrations and a decrease in the tracer arrival times at wells 1-9. An increase in runoff had the opposite effect on the simulated Cl⁻ BTCs. Minor changes in Cl⁻ concentrations were observed in well 11 and well 12. These changes were less pronounced for the TOUGH2 simulations, particularly for wells 5-8. The difference between results of the HYDRUS-3D and TOUGH2 simulations was mostly a general underestimation of the measured Cl⁻ concentrations in wells 5-7, which originated from omitting any calibration of TOUGH2. In spite of differences in their prediction accuracy, both models were sensitive to the runoff changes. The D_i indices ranged from -2 to 1 and from -0.5 to 0.6 for HYDRUS-3D and TOUGH2, respectively (Fig. 5-12a,b), thus indicating high model sensitivity to inflow along the upper boundary of the domain. The LSA was not conducted with precipitation and irrigation since the results would be essentially the same. This is because the flux in both models was set as irrigation and precipitation less runoff.

Next, the LSA was applied to the evapotranspiration data. Evapotranspiration ranged from 0 to 2.3 mm with a variance of 63% during the observation period. The goal of the LSA was to replace daily-averaged ET values with those averaged over the entire simulation period (i.e. $\overline{ET}$). The HYDRUS-3D and TOUGH2 simulations were run with daily- and long-term averaged and $\overline{ET}$ values using parameters and boundary conditions obtained with the HYDRUS-3D calibration. The Cl⁻ breakthrough curves simulated with differently averaged ET were very close to each (data not further shown). Values of the sensitivity index D_i were of the same order of magnitude in both models, except for well 2 in HYDRUS-3D simulations (Fig. 5-12c,d). For some reason the D_i values for this well decreased in time, likely indicating the increase in Cl⁻ concentrations at this location when daily-averaged ET were replaced with the longer-term $\overline{ET}$ values. Overall, small values of D_i obtained in the simulations allowed model simplification using averaging evapotranspiration values over the full 132-day simulation period.

Finally, the LSA was used also to estimate the effect of groundwater depth on the simulation results. The depth to groundwater as obtained with the HYDRUS-3D calibration was increased and decreased by 0.10 m. Both HYDRUS-3D and TOUGH2 were run with the modified lateral boundary conditions and parameter sets resulting from the HYDRUS-3D calibration. The changes in the groundwater depth affected the Cl⁻ BTCs differently at 12 locations using HYDRUS-3D. Within the irrigation plot the changes affected mostly the shapes of the BTCs. Simulated BTCs flattened when the depth to groundwater increased. This increase caused an increase in the Cl⁻ concentrations of wells 5-8, most likely because of an increase in the hydraulic gradients controlling lateral flow from the irrigation plot towards the observation wells. Changes in Cl⁻ were not significant in wells 10-12 since an increase in the gradient was not sufficient to deliver the tracer to these wells within the simulation period. Results of the TOUGH2 simulations revealed similar tendencies, except for smaller changes in the Cl⁻ concentrations of wells 5-8 as compared to HYDRUS-3D simulations. The D_i indices computed for both models are shown in Fig. 5-12e,f. HYDRUS-3D was generally more sensitive to changes in the groundwater depth compared to TOUGH2. The D_i indices ranged from -1.3 to 0.5 and from -0.4 to 0.4 for HYDRUS-3D and TOUGH2, respectively. In spite of different sensitivities, the effect of the boundary conditions on the simulation results was significant for both models. This indicates that groundwater depth could not be subjected to model simplification.

5.1.6 Sampling Frequency

The impact of different data collection frequencies on the modeling results was studied also using the HYDRUS-3D and TOUGH2 models. The sampling frequency in the field experiment ranged from twice a day during the first 37 days to once a day afterwards. Over 4,500 water samples were taken from 3 depths at 12 locations and analyzed for Cl⁻ content during the 132-day experiment. The intensive sampling scheme provided detailed information about tracer transport within and beyond the irrigated plot at different depths. The tracer was found to be transported preferentially beyond the irrigated plot in the upper layer of the saturated zone due to relatively high hydraulic conductivity of the soil material composing this layer (Pachepsky et al., 2011). Unfortunately, the adopted intensive sampling scheme is impractical for most application. Our study hence aimed at reducing the sampling frequency without significantly affecting model performance.

Alternative Cl⁻ sampling intervals were selected within the range from 1 to 10 days, with 1 day increments. Since sampling always started at the beginning of the experiment, the number of samplings ranged from 131 to 13. Model performance was evaluated using the root mean squared error (RMSE) for both models, as well as errors in predicting peak Cl⁻ concentrations (ΔC_{peak}) and peak times (ΔT_{peak}) using the HYDRUS-3D software. Both models were run using the HYDRUS-3D calibration parameters. Values of RMSE obtained with the simulations using the calibrated parameters (the zero sampling time intervals in Fig. 5-13) were compared with RMSEs computed for the reduced number of samplings (i.e. 131 and 13 sampling dates were used for the RMSE computations with 1-day and 10-day frequencies, respectively). The computed RMSE values did not change noticeably with an increase in the sampling time interval for wells 5-13. Changes were noticeable only for wells 1 and 4, which were located within the irrigated plot. The RMSE value decreased about two times for well 4 when the sampling interval became greater than 2 days. This occurred because the highest Cl⁻ concentrations in this well were observed within the first 3 days since the beginning of the experiment, with the model failing to reproduce the experimental data. Similar results were obtained from an analysis of errors in predicting the Cl⁻ peak concentrations and peak times using the HYDRUS-3D model.

Noticeable changes in the errors occurred only for well 4 after the sampling time interval exceeded 2 and 4 days for ΔC_{peak} and ΔT_{peak}, respectively. The discrepancy between these two time intervals for this well was caused by a fast decline in the Cl⁻ BTC during the first 37 days of the experiment, followed by a slow increase in the Cl⁻ concentration during the next 60 days, with a second peak time occurring at 96th day from beginning of the experiment. The increase in the sampling time interval caused a decrease in the peak concentration until this concentration became smaller than the second concentration peak. The sampling time interval was 5 days when the second peak became greater than the first peak (Fig. 5-13).

Overall a sampling interval of 10 days was sufficient for collecting data in wells located beyond the irrigated plot. A 2-day sampling interval was more appropriate for wells within the irrigated plot where the higher water fluxes caused more rapid transport of the applied tracer.
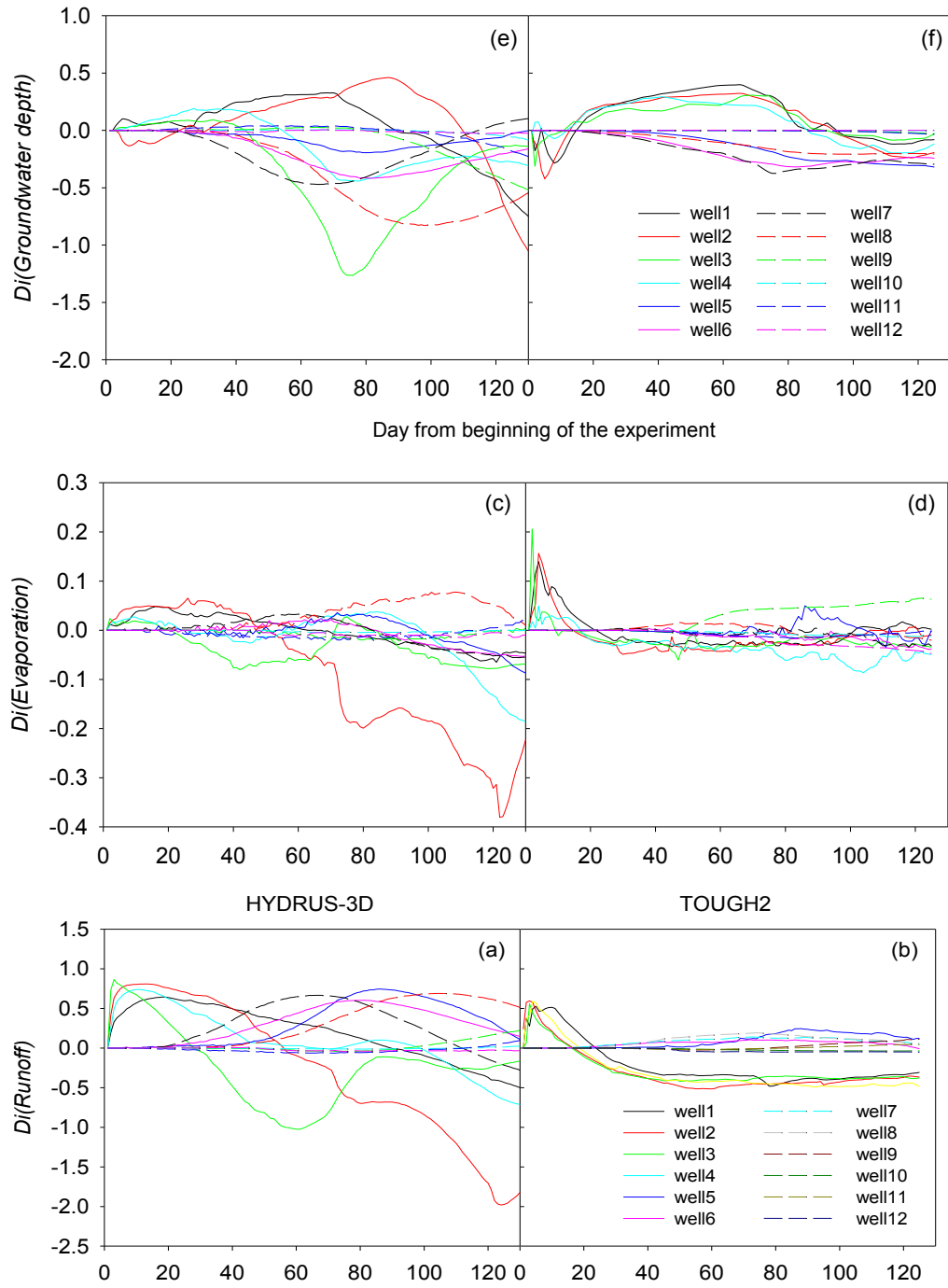
Figure 5-12 Temporal variation of the sensitivity index Di obtained in the LSA of the HYDRUS-3D and TOUGH2 models to runoff (a,b), evaporation (c,d) and groundwater depth.
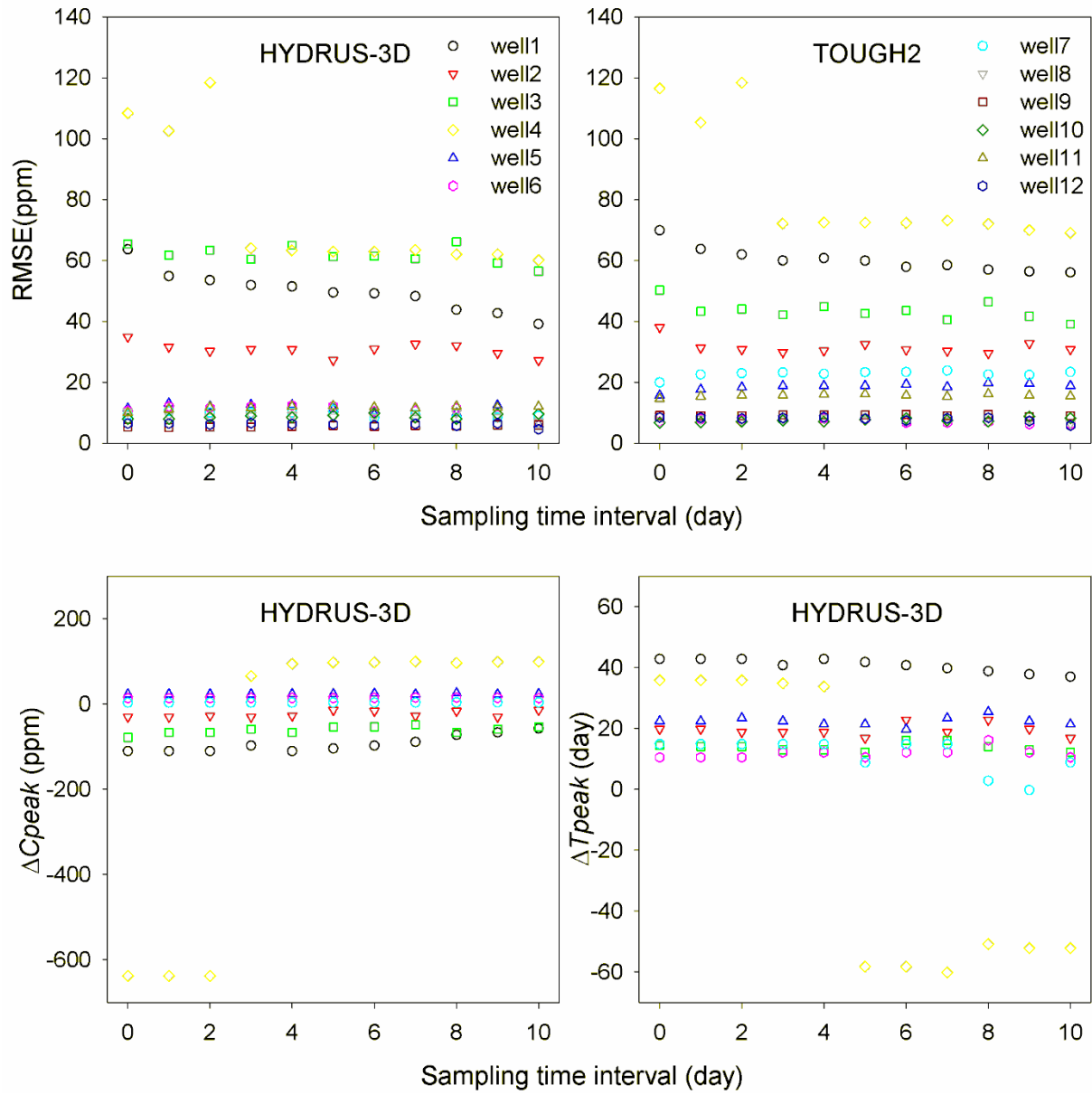
Figure 5-13 Errors of predicting Cl- concentrations (RMSE), peak Cl- concentrations (ΔC*peak*) and peak times (ΔT*peak*) obtained in the HYDRUS-3D and TOUGH2 simulations for measurements taken at different time intervals.

5.1.7 Soil Material Elimination

The LSA local sensitivity analysis revealed different model sensitivities to the 8 soil materials that represented the soil properties of the simulation domain. The very small values of the sensitivity indices D_i and S_i for material SCL2 suggest that the simulation domain perhaps could be simplified by using fewer soil materials, and consequently fewer transport parameters for the HYDRUS-3D and TOUGH2 simulations.

Three soil materials (SL3, SCL1, L1) were considered as possible replacements for SCL2. These three materials were selected since they were the nearest neighbors among the textural classes as obtained with the cluster analysis in Task II. The SL3, SCL1, L1 alternative soil materials (Fig. 5-14) had D_L and D_T parameter values similar to SCL2, except for SL3 whose D_L value was ten times smaller. The K_s values of the alternative materials were from 5 to 25 times higher than the K_s of SCL2 (Table 4-1). Three new simulations were carried out with the HYDRUS-3D and TOUGH2 models using the same initial and boundary conditions as for the calibration simulations, but with the modified parameters for SCL2.

The BTCs computed using HYDRUS-3D were found to be very close to the original BTCs obtained for SCL2. The main difference was for the simulation using SL3 for well 2, which showed a BTC that was much closer to the observed Cl⁻ curve. Unlike results of HYDRUS-3D simulations, the BTCs computed with TOUGH2 differed from the original curve of wells 1-8. The maximum and minimum differences were obtained with simulation using the SCL1 and L1soil materials, respectively (Fig. 5-15).

In view of these findings, the transport domain could now be simplified based solely on results of the HYDRUS-3D and TOUGH2 simulations using the original (SCL2) and alternative materials (SL3, SCL1, and L1). Both SCL1and L1 could replace the SCL2 material when using the HYDRUS-3D model, while L1 material was the best alternative to SCL2 for TOUGH2. The observed Cl⁻ data could provide additional information for model simplification. This information may include RMSE values of the simulated Cl⁻ concentrations and prediction errors in the peak concentrations and the times to peak concentrations. Based on the RMSEs (Fig. 5-15) all three materials could be selected for simplification of both models, provided the Cl⁻ concentrations of wells 1 and well 2 are not a concern. The ARS study team showed earlier in section 5.1.1 that the cumulative tracer fluxes did not affect Cl⁻ transport toward the other wells. However, if the concentrations in those wells are important for decision making, then SL3 for HYDRUS-3D and SL3 or L1 for TOUGH2 would be the best alternatives to the SCL2 soil material.

The same conclusions can be derived for the HYDRUS-3D code using other performance indicators, specifically the peak concentration (C_{peak}) and the time of the peak concentration (T_{peak}) of the tracer time series. Figure 5-16 illustrates the accuracy of HYDRUS-3D model in predicting *Cpeak* and *Tpeak* in wells 1-7. The errors for both performance indicators were smaller for the SL3 material, thus confirming the results of the RMSE analysis.
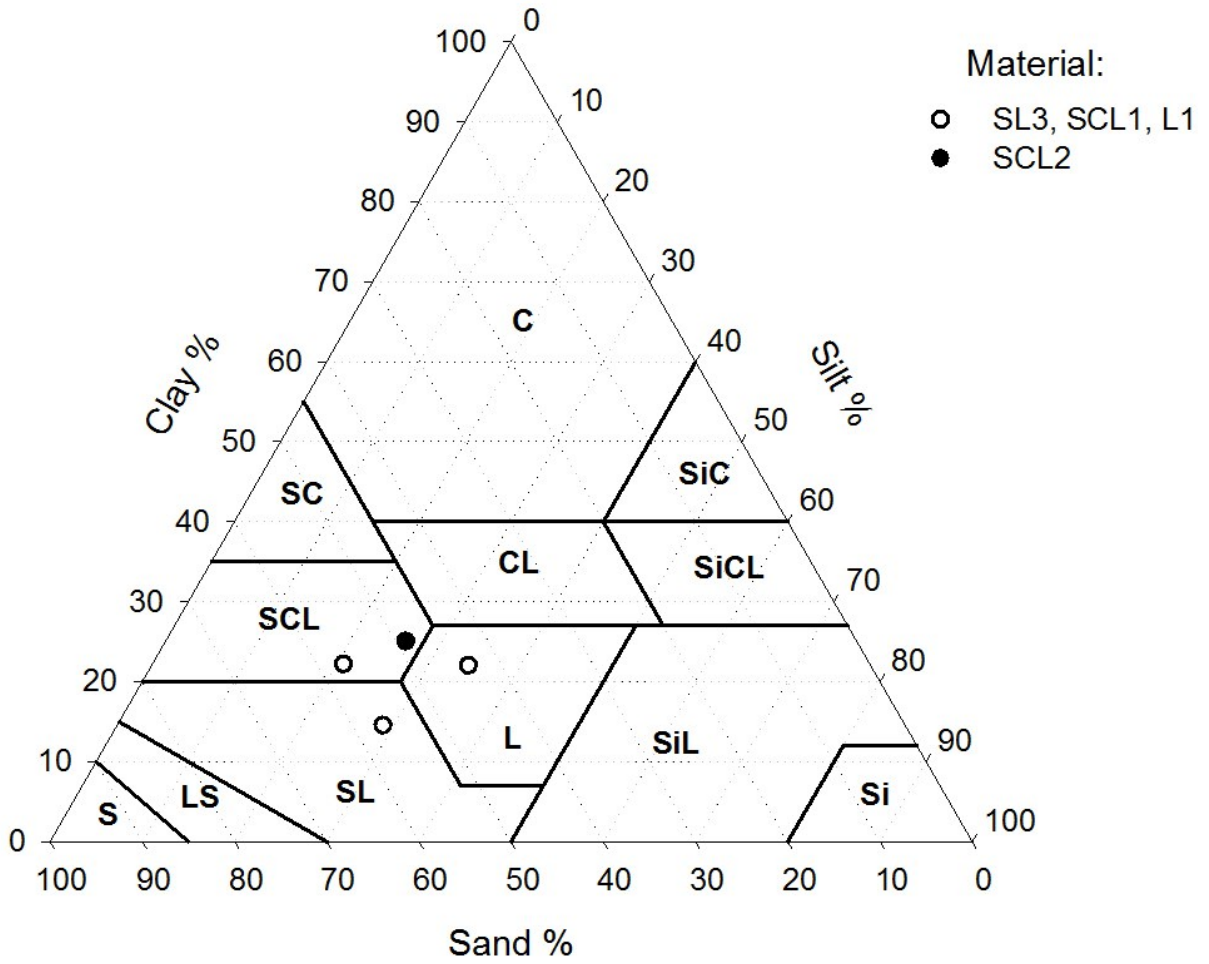
Figure 5-14 Alternative soil texture for the HYDRUS-3D model simplification.
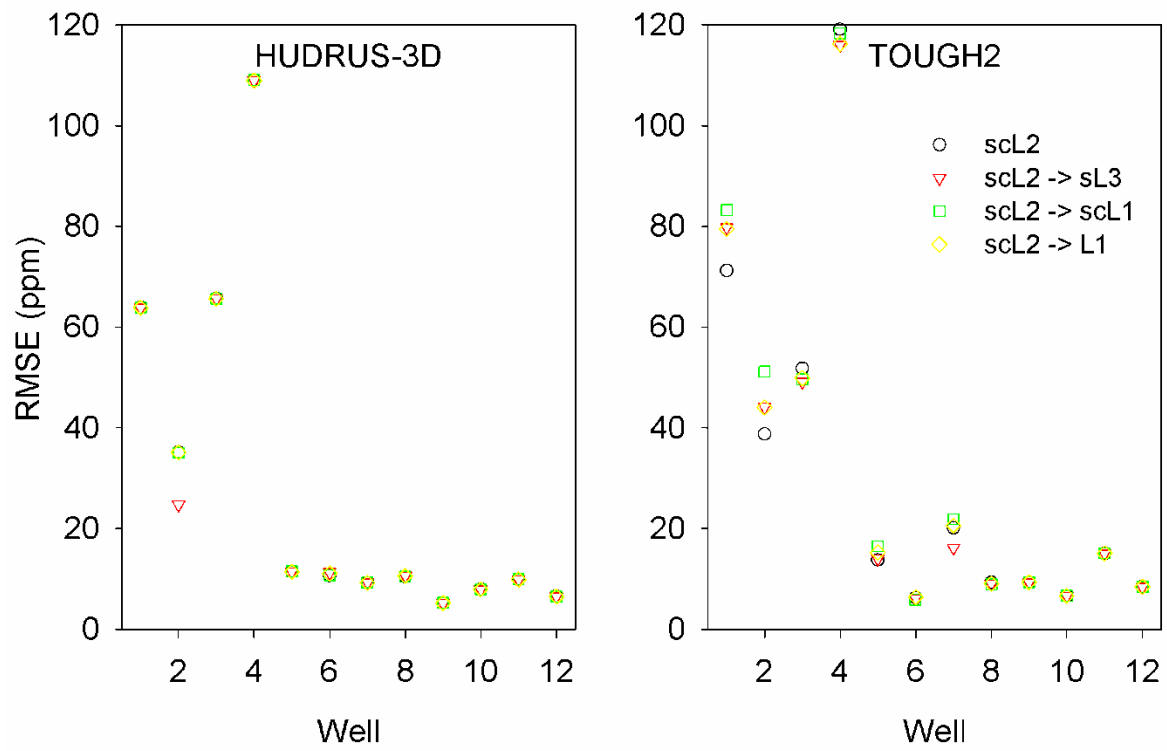
Figure 5-15 Root-mean-square errors (RMSE) of predicting Cl- concentrations obtained with two models using the original and modified properties for SCL2 soil material.
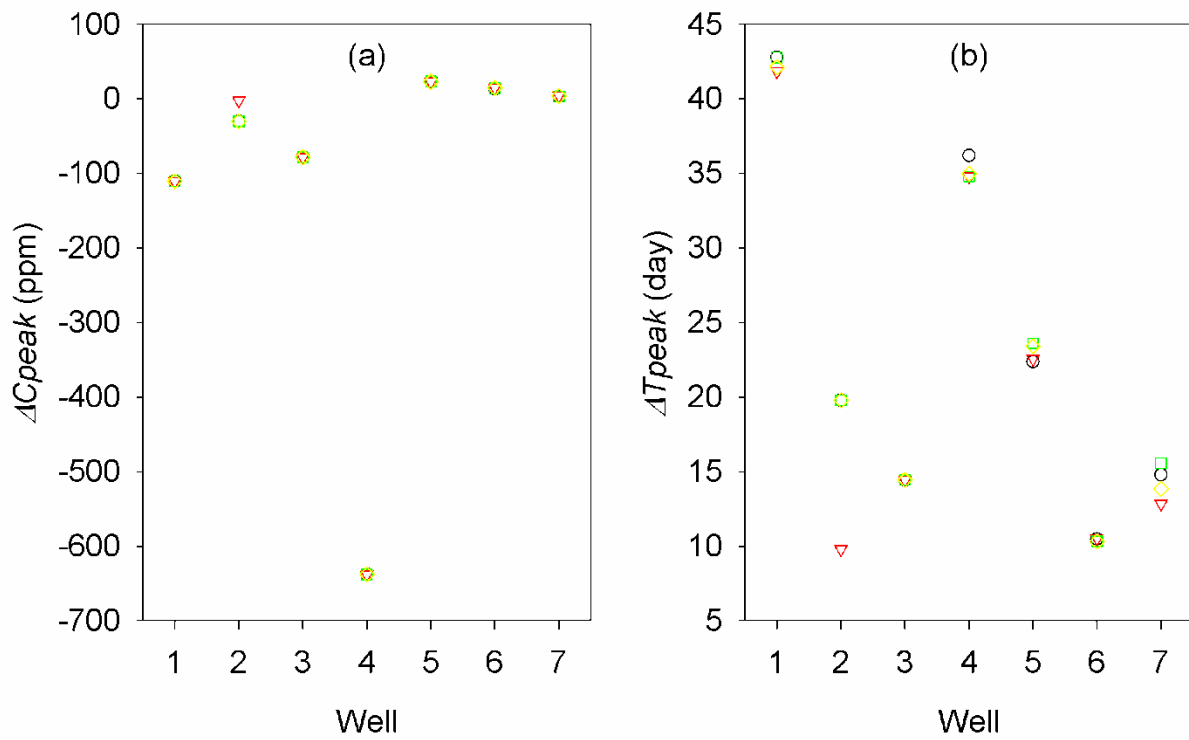
Figure 5-16 Errors of peak Cl- concentrations (ΔCpeak) and peak times (ΔTpeak) obtained with the HYDRUS-3D using the original and modified properties for SCL2 soil material.

5.2 Model Abstraction Applications

The model abstraction methodology, developed in this study, is comprised of the following steps: (i) site conceptualization, (ii) analysis of hydrological conditions, (iii) specifying accidental release scenarios, (iv) selecting performance indicators, (v) simulating contaminant transport for the assumed scenarios, and (vi) model simplification based on the performance indicators. The adopted abstraction techniques may differ and be specific for licensing NRC sites, depending upon hydrological complexity of a particular site. For the OPE3 site in Beltsville, MD, the ARS study team evaluated the applicability of three model abstraction techniques: (i) parameter abstraction, (ii) profile aggregation, and (iii) ignoring the unsaturated zone. To implement these three abstraction techniques and make the results transferable to other NRC sites the ARS study team assumed that:

- The calibrated HYDRUS-3D model adequately reproduced subsurface contaminant transport in the studied area.
- The hydrologic complexity of field B at the OPE3 site is commensurate with the complexity of NRC sites.
- Variability in the ground water depths as measured in the observation wells during the monitoring period (2006-2010) is similar to the variability at NRC sites.
- The irrigated area at the experimental site represents a decommissioning or licensing NRC site; accidental release is equally probable at any point within the site area.
- The creek and observation wells at the eastern boundary of the OPE3 site represent protected ground and surface water resources in the vicinity of the NRC site.
- Weather conditions at the decommissioning or licensing site are adequately represented by the 60-year weather data collected by the Beltsville weather station.

The ARS study team realized that some these assumptions may not be applicable to all NRC sites, but believe that the proposed methodology can serve as a general guidance to simplify models predicting subsurface contaminant transport from a particular release point.

5.2.1 Geometry of the Simulation Domain and Soil Material Distribution

NRC regulation 10 CFR 100.20(3) specifies that factors important to hydrological radionuclide transport (such as soil, sediment, and rock characteristics, adsorption and retention coefficients, ground water velocities, and distance to the nearest surface body of water) must be obtained from on-site measurements (61 FR 65176, Dec. 11, 1996). These onsite measurements provide a basis for creating a conceptual site model that evaluates and provides a qualitative representation of the important features, events, and processes of the groundwater flow and transport system (Interim). In this study all available information obtained from soil surveys, groundwater monitoring and modeling was integrated into the conceptual model of the decommissioning or licensing NRC site.

A new simulation domain that mimics a NRC site was defined within the boundaries of the GPR survey conducted in 2006. The domain included the tracer application plot, 44 old groundwater observation wells installed in 2000, and 12 new sampling wells installed in 2006 (Fig.5-17). The domain extended laterally for 190 m in the x (east-west) direction and for 118 m in the y (north-south) direction, bordered by a first-order distributary in the eastern part. The thickness (z direction) of the domain varied from 2.60 m in the east to 8.73 m in the west and was set so that it included a flow restrictive layer along the bottom boundary. This difference in the thickness was caused by a slope of the soil surface. The coordinate system in the HYDRUS-3D

simulation assumed the origin to be at the bottom of the domain (the *z*-axis was assumed to be positive upward).

The domain was subdivided vertically into ten hydrological layers based on data from the GPR survey conducted in 2006 (Fig. 3-31). The material in each layer was distributed based on soil texture measured in logs during installation of the observation wells. Soil properties within the irrigated plot were assigned similar to those used for the HYDRUS-3D calibration, while soil textures measured in 44 logs during well installation in 2000 (Table 3-2) were used outside that area. In all 12 texture classes/subclasses were needed to represent the soil materials within the transport domain (Fig 5-18). These classes included: Sand (S), Loamy sand (lS), two Sandy Loam classes (SL1, SL2), Sandy Clay Loam (SCL1), Clay Loam (CL), two Loam classes (L1 and L2), three Silt Clay Loam classes (SiCL1, SiCL2, SiCL3), and Silt Loam (SiL). Boundaries between different soil materials of each computational layer were set at approximately equal distances from the observation wells where the materials were measured. The soil below a depth of 4.03 m was considered to represent a flow restrictive layer, and was represented by the SiCL3 texture which has a very low value of the saturated hydraulic conductivity (K_s=0.001 m day^{-1}). Subdividing the soil hydraulic properties into 12 groups was expected to reflect hydrological complexity of the simulation domain, and to provide flexibility for model abstraction.

Soil hydraulic parameters for the 12 materials are shown in Table 5-1. Within the irrigated area these properties were set equal to the parameters obtained in the HYDRUS-3D calibration. Beyond this area the ARS study team used ROSETTA estimated parameters of the van-Genuchten water-retention equation and K_s values estimated using the PTFs of Rawls et al. (1998). The dispersivity parameters α_L and α_T of all materials were set to 1.0 m and 0.2, respectively.

The unstructured (triangular prisms) finite element mesh consisted of 11,770 nodes and 20,380 3D elements. The vertical size of the finite element prisms varied depending upon location, with the elements following the soil surface and the surface of the lower sub-layer. The vertical size of the finite elements for the upper 2-m soil layer ranged from 0.2 to 0.3 m, while the size below this depth was about 0.3 m and 2-m at the eastern and western boundaries of the domain, respectively (Fig. 3-19a). The horizontal mesh size increased from about 1.3 m within the irrigation plot where the fastest fluxes were observed, to 5 m outside of that area. Using this geometry and associated finite element mesh, a 10-year long simulation generally took between 6 and 8 h with a Computech Intel Core™ i7 CPU 930 @2.80 GHz computer. Further refinement of the finite element mesh required longer simulation times. A total of 100 observation nodes (Fig. 5-19b) spaced 10-m apart were used to record the simulated breakthrough curves at depths of approximately 2.3 m. Although the maximum lateral fluxes of the tracer during the sensitivity analysis of HYDRUS-3D were observed to occur at a depth of 1.1 m, the ARS study team increased the output depth since the 10-year simulations showed that the water table decreased to below 2 m during periods of low precipitation, which caused the long-term tracer fluxes to become larger at a depth of 2.3 m as compared to the 1.1 m depth.
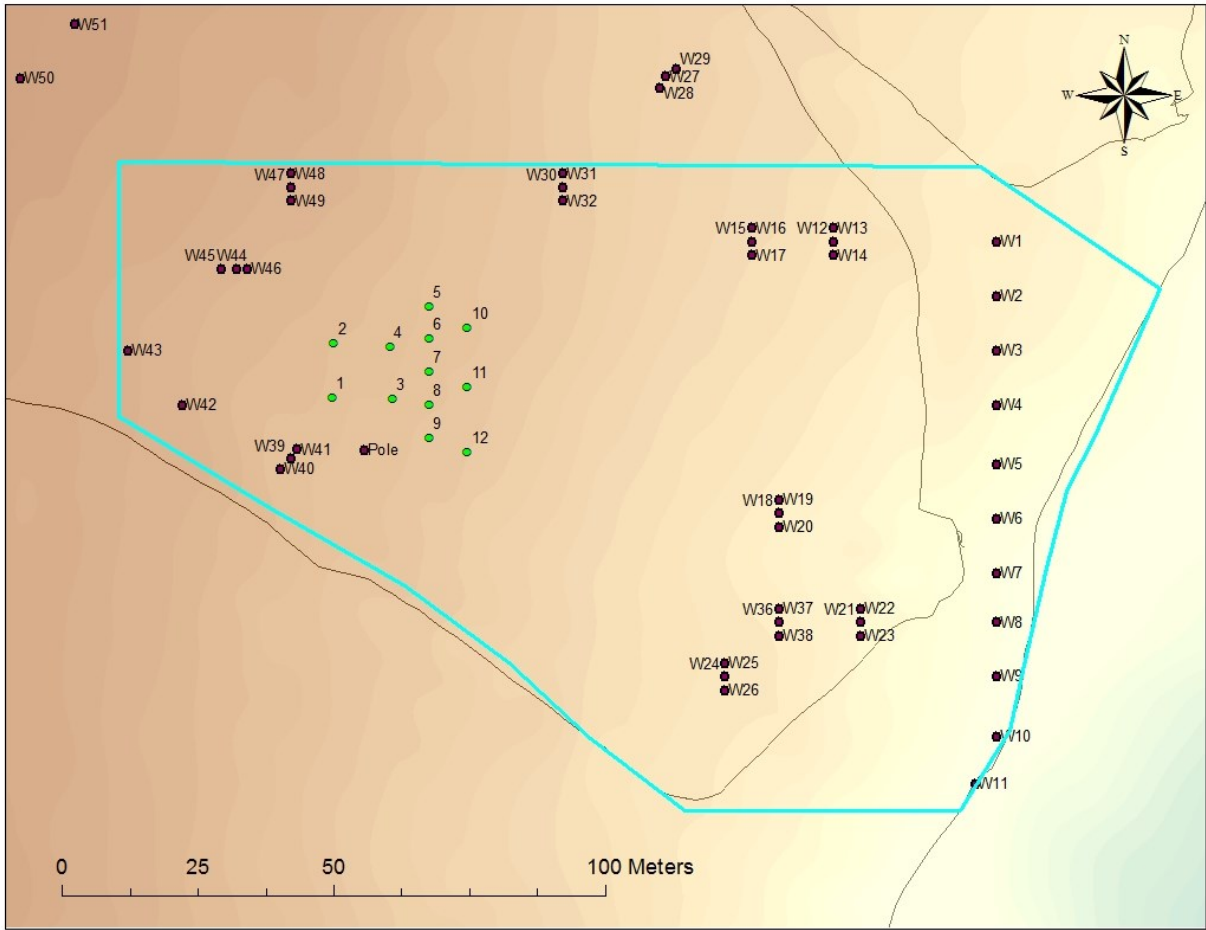
Figure 5-17   Boundary of new simulation domain with old (brown circles) and new (green circles) groundwater wells.
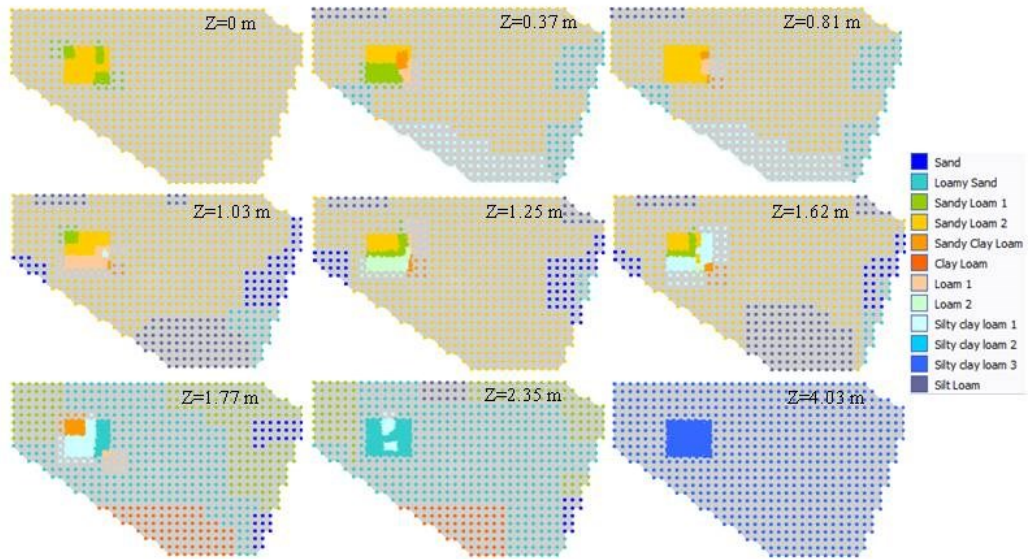
Figure 5-18 Material distribution for the HYDRUS-3D mesh.

Table 5-1 Hydraulics conductivity values obtained by calibration, estimation with PTFs from Rawls et al. (1998) and scaling abstractions A1 through A6.

Soil texture	L_s, m		0.05			0.2		
	b		0.3	0.5	0.7	0.3	0.5	0.7
	Saturated hydraulic conductivity, K_s, m d^{-1}							
	Calibrated	Rawls et al. 1998	A1	A2	A3	A4	A5	A6
SL1	2.50	1.339	5.331	13.390	33.634	3.517	6.695	12.745
SL2	0.10	0.307	1.222	3.070	7.711	0.806	1.535	2.922
SL3	2.50	1.339	5.331	13.390	33.634	3.517	6.695	12.745
SCL1	0.45	0.067	0.267	0.670	1.683	0.176	0.335	0.638
SCL2	0.11	0.017	0.068	0.170	0.427	0.045	0.085	0.162
L1	2.00	0.149	0.593	1.490	3.743	0.391	0.745	1.418
L2	0.70	0.149	0.593	1.490	3.743	0.391	0.745	1.418
SiCL1	2.00	0.118	0.470	1.180	2.964	0.310	0.590	1.123
SiCL2	0.15	0.118	0.470	1.180	2.964	0.310	0.590	1.123
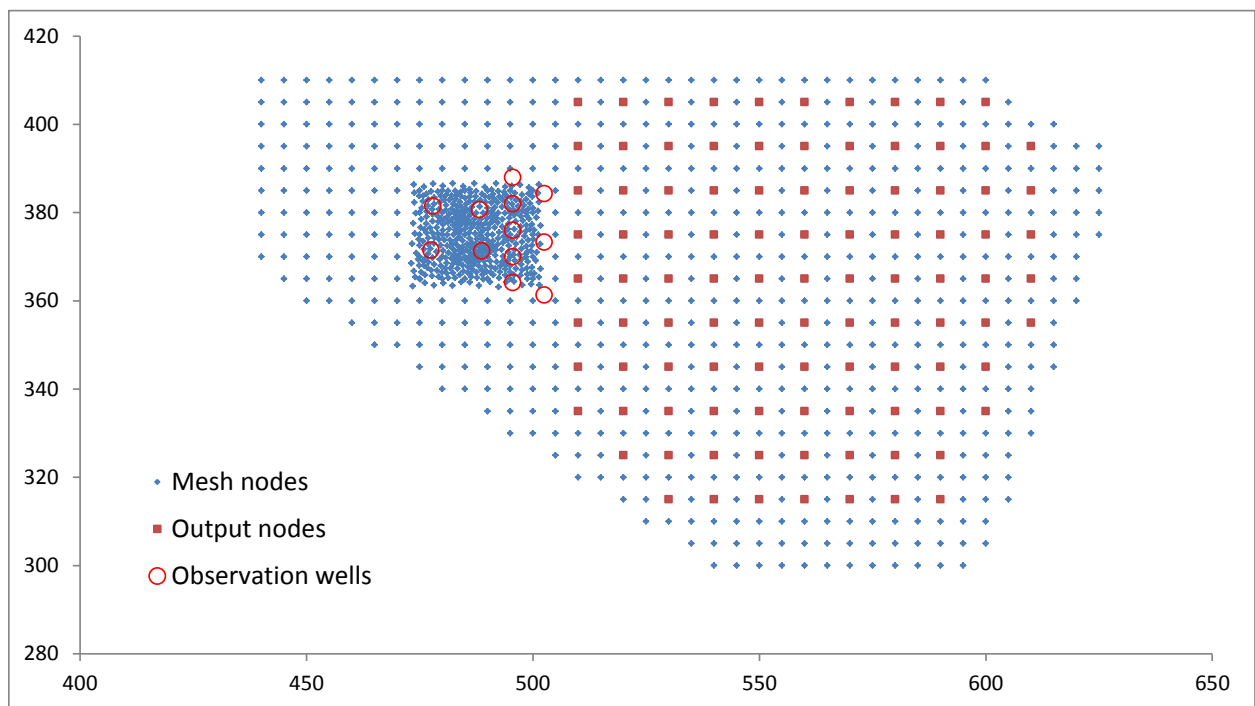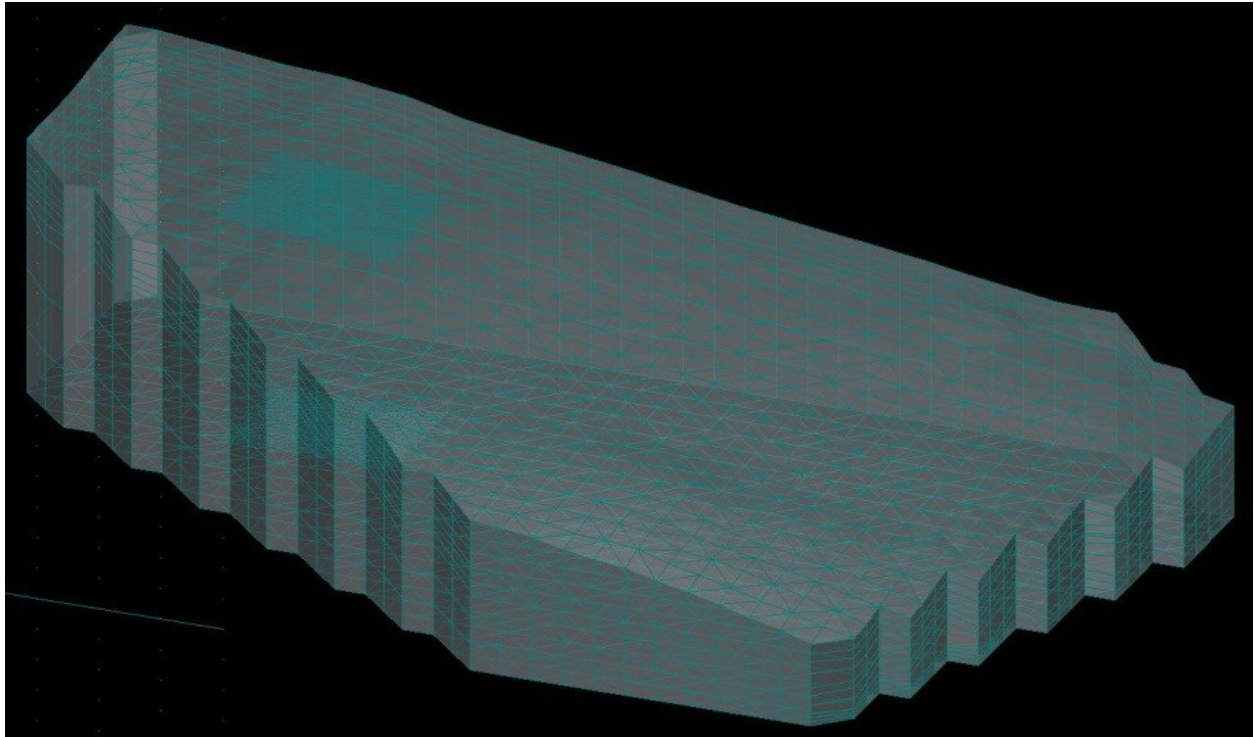
Figure 5-19   Mesh generated for HYDRUS-3D simulations (a) with coordinates for 100 output locations (b).
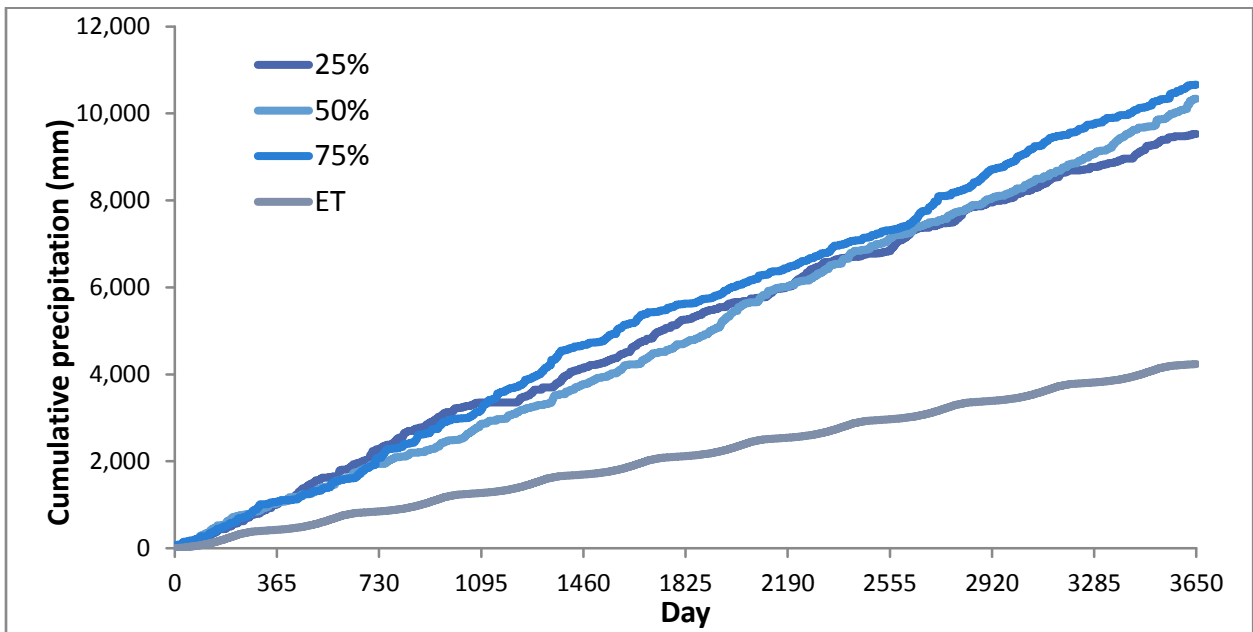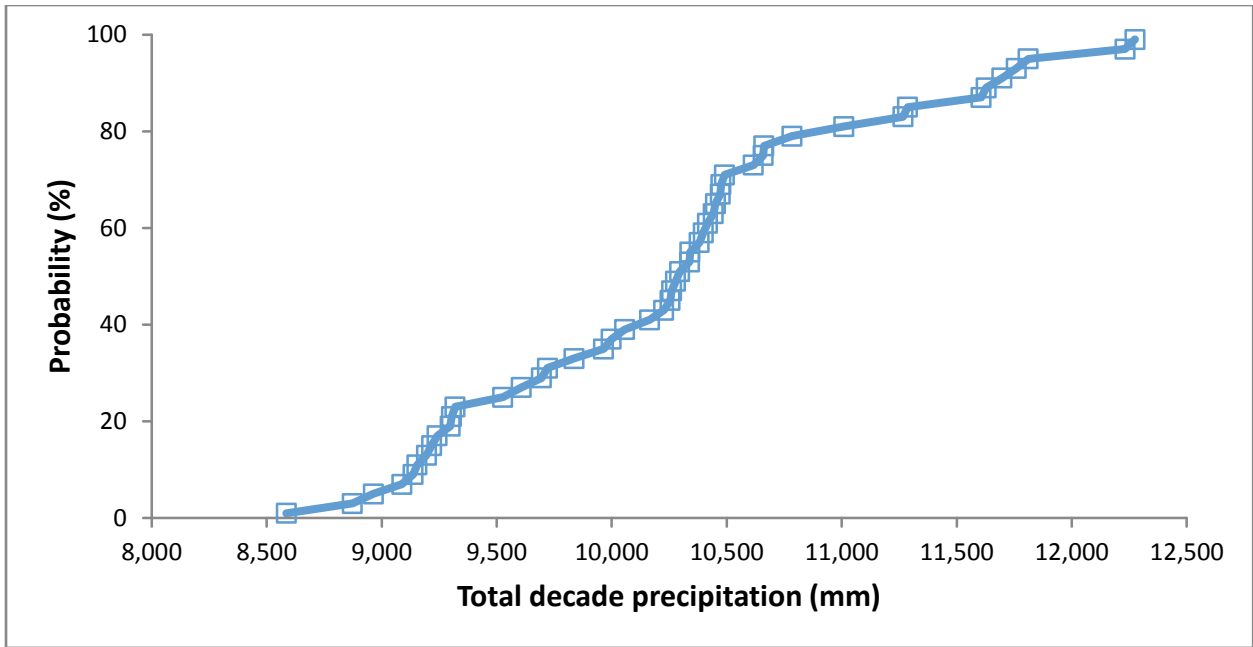
Figure 5-20 Probability distribution of the decade precipitation (a) and cumulative precipitation for decades with 25%, 50% and 75% probabilities (b).

5.2.2 Initial and Boundary Conditions

In this part of our study the ARS study team examined the subsurface transport of the tracer leaked from the irrigated area. The assumed initial spatial distribution of Cl^- was found to affect the breakthrough curves calculated using HYDRUS-3D. The ARS study team observed that relatively high initial Cl^- concentrations of wells 9, 11 and 12 (Fig. 5-2) decreased partly due to inflow of low concentrated Cl^- solution from locations wells 5 and 3. To eliminate the influence of the initial concentration on the tracer distribution, and to better follow the subsurface migration of the applied chemical, the initial tracer concentration was set equal to zero across the simulation domain. A third-type (Cauchy) boundary condition was used along the soil surface to simulate solute transport. The tracer concentrations of precipitation and groundwater along the lateral boundaries were set to zero.

As initial condition for the water flow simulations the ARS study team used equilibrium pressure head distributions. These distributions were assigned to four groundwater depth scenarios defined later in section 5.2.3. Ten-day averaged atmospheric boundary conditions were used at the soil surface. Evapotranspiration rates were computed using the Penman-Monteith method as documented by FAO (Allen et al., 1998) for weather data of year 2006 and repeated for each simulation year. Based on the flow velocities obtained in section 5.1.1, a 10-year simulation period was considered to be sufficient for the tracer to arrive at the eastern boundary of the simulation domain. The ARS study team discarded extra days in leap years, which allowed us to fix the total simulation period at 3650 days. Three precipitation scenarios were considered in the 10-year HYDRUS-3D simulations. To generate these scenarios, the precipitation data collected in Beltsville from 1949 to 2006 were summed for each 10 consecutive years and ranked in an increasing order (Fig. 5-20a). Ten-year precipitation rates at probability levels of 25%, 50% and 75% were subsequently selected for HYDRUS-3D simulations. The selected decades were 1992-2001, 1984-1993 and 1976-1985, with total precipitation rates of 9527 mm, 10,339 mm and 10,659 mm for the 25%, 50% and 75% probabilities, respectively. Since the interdecadal distributions of precipitation were relatively uniform (Fig 5-20b), the ARS study team did not expect that temporal variability in the upper boundary conditions would affect results of the 10-year HYDRUS-3D calculations. Daily variability in rainfall values affected the numerical stability of HYDRUS-3D simulations, which could be avoided by using 10-day averaged precipitation values in our simulations. Total decadal evapotranspiration was 4237 mm, less than half of the precipitation. A zero flux boundary condition was imposed along the bottom of the domain.

5.2.3 Simulation Scenarios

To model the accidental release of a tracer from a single holding tank outside of a containment structure, the ARS study team considered four scenarios. In the first two scenarios the release points were set on the soil surface at locations of wells 3 and 4 within the irrigated plot. The accidental release of a 1000 ppm spill was assumed to occur at rate of 0.2 m day^{-1} through a 1-m^2 area within the first 10 days of the simulation. This translates to a total of 2 kg of tracer being released as a pulse with 2 m^3 of water. The release points of the other two scenarios were in groundwater at depth of 2.3 m. The tracer fluxed for the third and fourth scenarios hence occurred in the saturated zone. The ARS study team expected much faster solute transport rates when released in the saturated zone as compared to the partly unsaturated conditions considered for the first two scenarios.

Four different scenarios were developed also for the lateral boundary conditions. These conditions were set as equilibrium pressure head profiles for spatially variable, but constant in time, groundwater depths. The main reason for assuming a constant in time boundary condition

was the considerable complexity of the shape at the lateral boundary of the simulation domain. Whereas technically HYDRUS-3D allows the implementation of time-variable pressure head boundary conditions, the transition between the physical boundaries of the simulation domain and temporally changing variable groundwater levels creates a problem of conjunction and often results in abrupt changes in the boundary conditions which hamper rapid convergence of the slow that numerical solution of the Richards equation. Another reason for using different boundary scenarios is uncertainty due to limited information about the spatial and temporal variability of groundwater levels. Long-term hydrological data are not available at many newly designed NRC sites. For these cases, indirect methods are recommended for determining annual average conditions (Interim SRP 2.412 and 2.4.13). Because of uncertainty in the indirect methods, it is necessary in these cases to conduct a consequence analysis assuming conservative hydrologic conditions and parameter estimates. The groundwater depths measured in the wells at field B in 2006-2007 (Fig. 3-11) and 2007-2008 (Fig. 3-27) provided general information about the spatial and temporal variability of the water table in the flow domain. The ARS study team realized that the 10-year variability in groundwater levels may differ from the observed values. For this reason the ARS study team generated several scenarios with different slopes and depths to groundwater for HYDRUS-3D simulations. These were: (1) a slope in x direction ($SlopeX$) of 1.6%, a slope in the y direction ($SlopeY$) of 1%, and a groundwater level (GWD) of 5.5 m relative to the bottom of the simulation domain in Fig. 5-19, (2) the same slopes with GWD=6.0 m; (3) $SlopeX$=1.6%, $SlopeY$=0.4%, and GWD=5.5 m; and (4) $SlopeX$=1.6%, $SlopeY$=0.4%, GWD=6.0 m. Note that GWD is the level above the zero vertical coordinate, which served as the bottom boundary of the flow domain. These GWD levels corresponded to groundwater depths varying from 3.5 m to 1.5 m for GWD=5.5 m, and from 3.0 m to 1 m for GWD=6.0 m in the western and eastern parts of the domain, respectively. These levels of the groundwater were also used for the initial pressure head distributions in the domain as was described in section 5.2.2.

Overall, 48 different scenarios were simulated with HYDRUS-3D, comprising the four release scenarios, four groundwater scenarios and three precipitation scenarios. These scenarios provided input for the model abstraction process described in the following sections.

5.2.4 Performance Indicators

The LSA was performed using tracer concentrations representing the entire simulation period. In practice different performance indicators can be used for decision making based on the model results. The American National Standard for evaluation of subsurface radionuclide transport at commercial nuclear power plants (ANSI/ANS-2.17-2010) recommends using radionuclide concentrations (either volume- or flux averaged), mass fluxes, travel times and predicted doses as performance indicators. For these reasons the ARS study team used in this study simulated tracer concentrations and fluxes to characterize chemical transport from the assumed accidental release locations. To characterize the direction and velocity of contaminant migration the ARS study team used the simulated peak concentration (C_{peak}) and times to peak concentration (T_{peak}) at the observation nodes and along transects at different distances from the release points. The spatial distributions of T_{peak} were used to estimate the velocities and directions of the contaminant plume for all simulation scenarios.

To evaluate the accuracy of each model abstraction technique, several efficiency criteria were used in this study. Model efficiency criteria were defined by Beven (2001) as mathematical measures of how well a model simulation fits observations. In hydrological modeling, efficiency criteria are typically based on the ratio between the prediction error and variability in the observations. In spite of apparent simplicity, there is no general agreement on how to use the

error term and the observation variability. The simplest way to compute model efficiency was proposed by Nash and Sutcliffe (1970). They introduced the efficiency criterion *E* as the deviation from one of the absolute squared differences between the predicted and observed values, normalized by the variance of the observed values:

$$E_{NS} = 1 - \frac{\sum (O_i - P_i)^2}{\sum (O_i - \bar{O})^2} \qquad (5\text{-}8)$$

where O_i and P_i are the measured and predicted values. Values of the efficiency index *E* are within the range from 0 to -∞ and increase in absolute value when the deviations between predicted and measured values increase. A deficiency of this criterion is in its sensitivity to large measured or predicted values, since squared values are used in both the numerator and denominator.

Another efficiency criterion is the index of agreement *d* proposed by Willmot (1981). This criterion is defined as the ratio between the mean square error and the potential error as follows

$$d_W = 1 - \frac{\sum (O_i - P_i)^2}{\sum (|P_i - \bar{O}| + |O_i - \bar{O}|)^2} \qquad (5\text{-}9)$$

A third model efficiency index used in this study to evaluate model abstraction accuracy is the squared correlation coefficient:

$$r^2 = \left(\frac{\sum (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum (O_i - \bar{O})^2} \sqrt{\sum (P_i - \bar{P})^2}} \right)^2 \qquad (5\text{-}10)$$

The power of the above efficiency indices for hydrological model assessment were examined by Krause et al. (2005). These authors showed that selection and use of specific efficiency criteria, as well as interpretation of the results, can be a challenge since each criterion places different emphasis on different types of simulated and observed results. The ARS study team believes that the use of multiple criteria for the abstracted models will reduce uncertainty in the result interpretation. Therefore all three criteria were applied to the performance indicators discussed earlier in this section.

To assess the performance of the selected indicators 48 HYDRUS-3D runs were carried out for the scenarios defined in section 5.2.3 using parameters obtained in the HYDRUS-3D calibration. Required hydraulic parameters for the new materials beyond the calibration domain were estimated using the ROSETTA software (Table 5-1). These simulations will be referred to as hereafter as "calibrated" to distinguish them from the abstraction simulations. Results of the simulations for the three weather, two source and two groundwater scenarios are shown in Figs. 5-21 through Fig. 5-23. The ARS study team selected these 12 cases as the most representative of the 48 HYDRUS-3D simulations. Results showed that the shape of the contaminant plume as represented by the peak concentration was influenced most by the location of the release point. Overall, the release from locations 1 and 3 produced considerably wider plumes with larger tracer concentrations as compared to locations 2 and 4 (only locations 1 and 2 are shown in Fig. 5-21). Weather scenarios and the slope of groundwater had minor effects on the plume shape and the concentrations. Tracer concentrations generally decreased exponentially with distance from the release point (Fig. 5-21). This decrease was more pronounced for releases from locations 2 and 4 rather than from locations 1 and 3, and more pronounced for a groundwater level of 5.5 m as compared to the 6.0-m scenario.

The location of the release points also influenced the tracer peak concentration travel time, Tpeak. Spatial patterns of Tpeak were different for release locations 1 and 2, and were affected by groundwater depth scenario for location 1 (Fig. 5-22). The shapes of Tpeak distributions create an illusion of having preferential tracer pathways in x-direction from the source at locations 1 and location 2 for the 5.5-m groundwater level scenario. However, these pathways significantly differ from zones with high Cpeak values shown in Fig. 5-21 and no correlation was found between the simulated Tpeak and Cpeak values. Differences between the simulated Tpeak and Cpeak spatial patterns can be explained by flow diversity. Due to high heterogeneity of the simulation domain, the tracer arrived in many instances at the nodes within a short period of time, but in small concentrations. Soon after arrival, the tracer may become diluted by inflow from the upslope area and hence may not contribute further to plume propagation. The highly concentrated flux at other locations was moving slower than the less concentrated flux, which caused the Tpeak values to become smaller in the main flow path as to the branches. The Tpeak and Cpeak values were derived from the tracer breakthrough curves at each observation node and did not characterize the total mass of the transferred contaminant. Since the spatial pattern of mass is the only true measure of the migrating contaminant plume, the ARS study team supplemented the Tpeak and Cpeak analysis with the analysis of the contaminant mass distribution.

Figure 5-21 Spatial distributions of log10(C$_{peak}$[ppm]) performance indicator obtained in HYDRUS-3D simulations with the calibrated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.

Standard HYDRUS-3D simulation results include time series of the tracer concentration $C(t)$ and nodal water fluxes $q(t)$. The ARS study team used these values to compute the tracer fluxes QC at the 100 output nodes according to Eq. 5-7. The flux values, shown in Fig. 5-23, represent the total tracer masses passing through a 1 m2 area during the simulations. The QC patterns closely resemble the Tpeak and especially the Cpeak patterns. The total tracer mass generally decreased more or less exponentially within the first 30-m travel distance from the release point. This shows that only a fraction of the total tracer mass passed through the observation wells and that this fraction decreased with distance due to the plume spreading beyond the monitoring network. Beyond the 30-m travel distance the total tracer mass remained somewhat constant. Similar to the Cpeak patterns, the QC spatial distribution was wider for the release from locations 1 and 3 compared to locations 2 and 4, with considerably more mass being transported from locations 1 and 3 (Figs. 5-21 and 5-23). The major effects of source location and groundwater level on tracer transport, and the relatively minor effects of groundwater slope and the weather scenario, can be seen in Fig. 5-23. Different release locations produced different spatial distributions of the performance indicators, with the distributions being further altered by the flow conditions associated with the groundwater level.

Overall, both release location and lateral boundary condition were major factors affecting the contaminant transport process in our study. The largest tracer concentrations and fluxes were observed when the tracer was released at locations 1 and 3 into shallow groundwater. Among the performance indicators, the total tracer flux QC was the most accurate, and the times to peak concentration Tpeak the least accurate, for assessing contaminant transport in this study. The ARS study team realized that the results may be site- and scenario-specific, and that further research may be needed to investigate the power of the performance indicators for different contaminant transport condition.
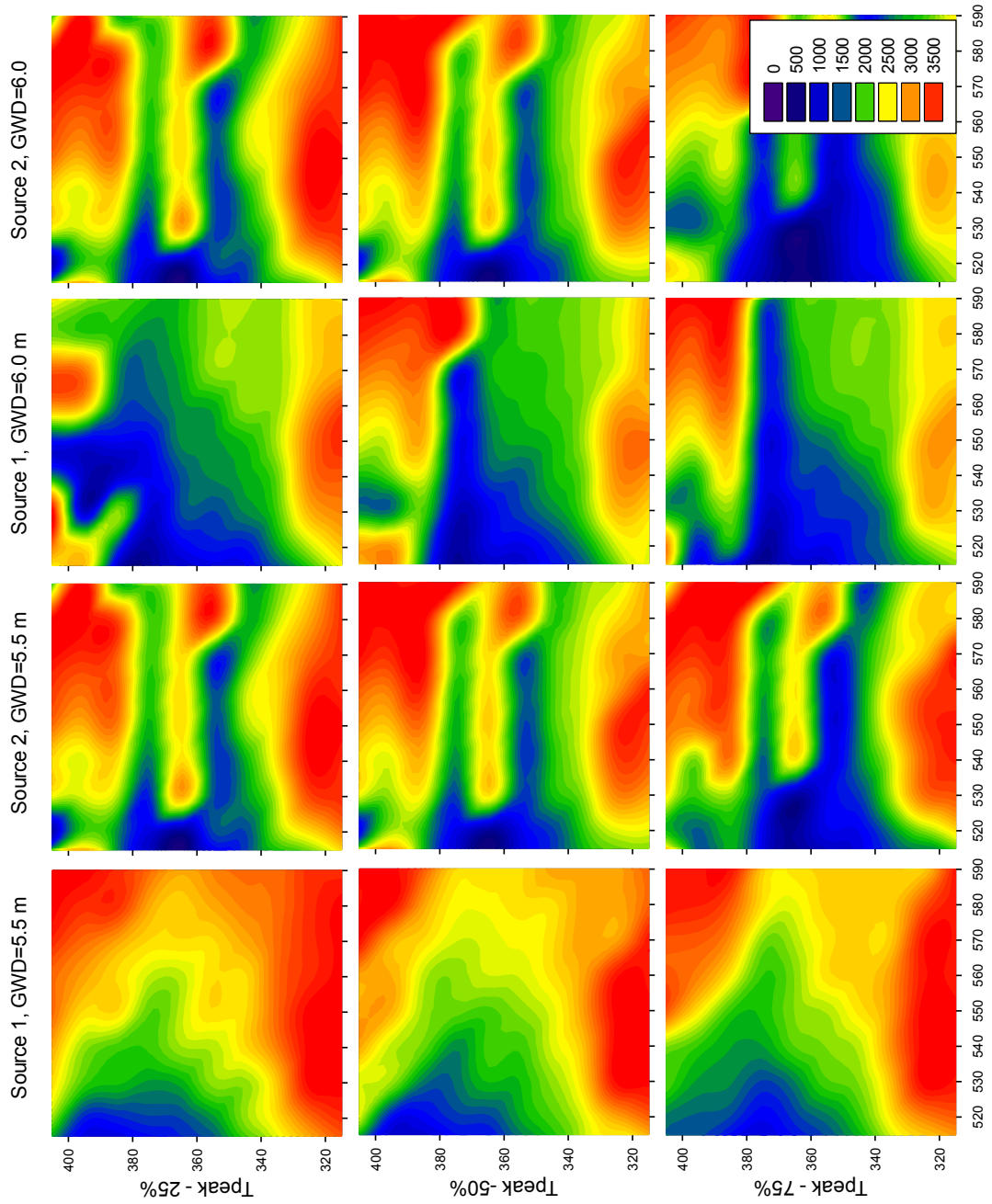
Figure 5-22 Spatial distributions of T $_{peak}$ [day] performance indicator obtained in HYDRUS-3D simulations with the calibrated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
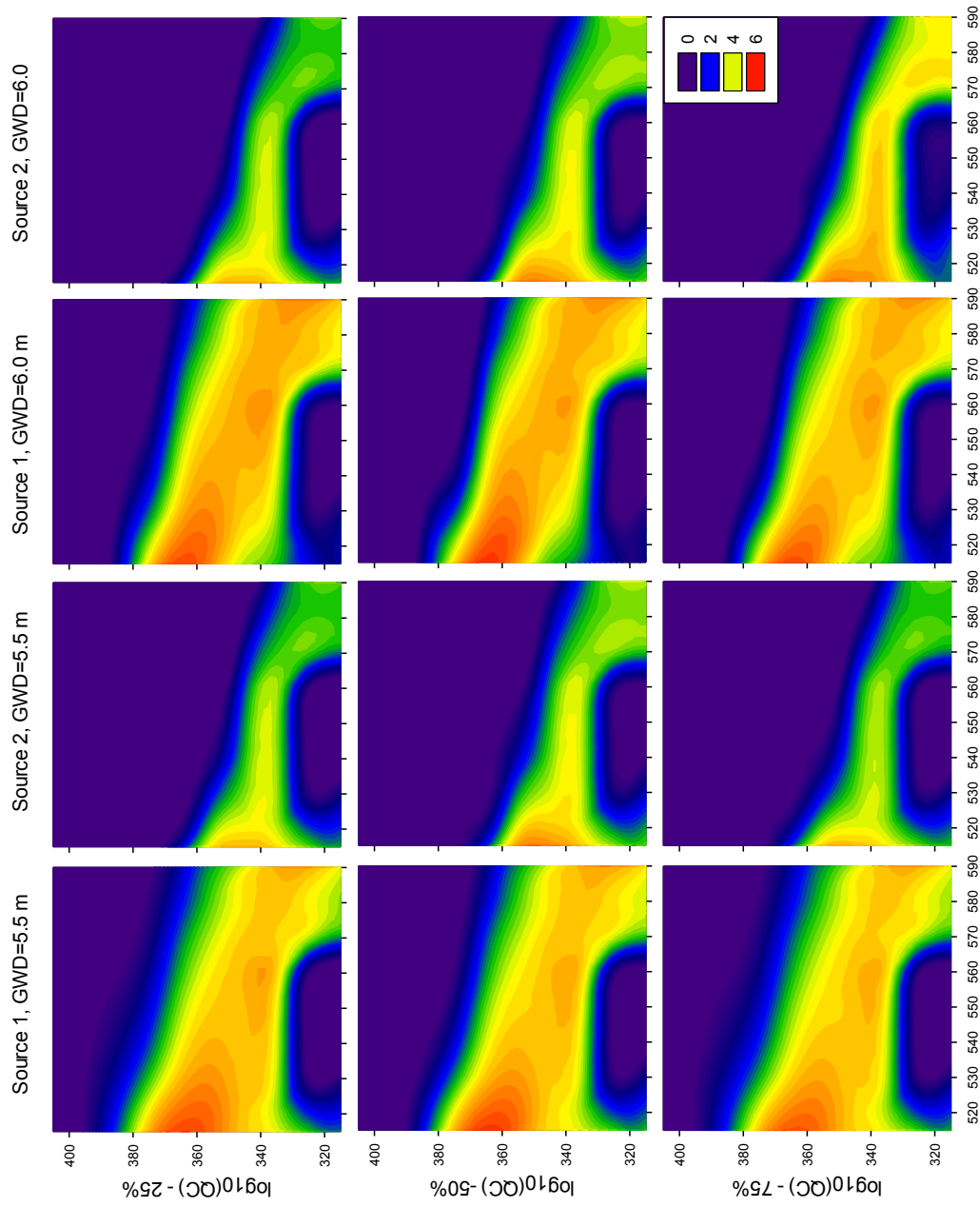
Figure 5-23 Spatial distributions of log10(QC [ppm m day-1]) performance indicator obtained in HYDRUS-3D simulations with the calibrated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.

5.2.5 Abstraction of Parameter Determination: Using Pedotransfer Functions

The first model abstraction technique implemented in this study was abstraction of the soil hydraulic parameters. The purpose of this abstraction was to skip the time-consuming and computationally demanding model calibration process, but instead to work with estimated hydraulic parameters. The ARS study team showed earlier in section 4.2 that the calibrated parameters differed from the pedotransfer function (PTF) derived estimates. However, the obtained results could be scale and scenario-specific. The HYDRUS-3D calibration was based on application of the Cl⁻ tracer to a relatively large plot (10x10m), followed by measurements in the vicinity (7 m) of the application area. A steady-state infiltration regime was artificially created and maintained within the application area. Water fluxes in reality are rarely at steady-state due to irregular rainfall events and groundwater fluctuations such that steady-state regimes may persist only for short periods of time. The ARS study team assumed that the model will be less sensitive to the hydraulic properties and boundary conditions if the release occurs over a small area (1 m^2), and if the tracer is transported over a long distance with unsteady water flow. The ARS study team expected to see a scale and time effect on propagation of the contaminant plume caused by tracer dilution and dispersion.

To test the parameter abstraction technique, 48 HYDRUS-3D runs were carried out with scenarios defined in section 5.2.3, water retention parameters used in the calibrated model, and saturated hydraulic conductivities of the entire flow domain estimated using the PTFs of Rawls et al. (1998). The estimated K_s values were 4.37, 2.95, 0.31, 1.34, 0.07, 0.02, 0.15, 0.15, 0.12, 0.12, 0.001 and 0.35 m day^{-1} for S, LS, SL1, SL2, SCL, CL, L1, L2, SiCL1, SiCL2, SiCL3 and SiL, respectively. The K_s value of the SiCL3 material was not altered in this abstraction since SiCL3 served as a flow restrictive layer at the bottom of the profile. Thus only the saturated hydraulic conductivity was abstracted in this section, while soil water retention parameters were kept unchanged in all simulations. The performance indicators C_{peak}, T_{peak}, and QC were derived from the simulated concentration time series at 100 output nodes and along 11 transects passing through the observation nodes and crossing the simulation domain in the y-direction.

Iso-contours of the peak concentration, the time to peak concentration and the total solute flux over the 3650-day simulation for the three weather, two groundwater and two source location scenarios are shown in Fig. 5-24 - Fig. 5-26. Substituting the calibrated K_s values with values determined using the PTF estimates of Rawls et al. (1998) affected the spatial distributions of C_{peak}, T_{peak} and QC in the transport domain. Values of C_{peak} for the abstracted model were in general larger than those for the calibrated model, with the shapes of the plume being more elongated in the x-direction. These changes can be attributed to the almost threefold increase in K_s of the loamy sand material at depths of 1.77 m and 2.35 m (Fig.5-18), which must have facilitated tracer transport. Increased K_s values noticeably reduced the tracer arrival times (Fig. 5-22 and Fig. 5-25). T_{peak} values computed were from 1.1 to 1.9 times smaller for the abstracted model compared to the calibrated model. The flow acceleration was not proportional to the increase in K_s, most likely due to a change in the pressure head gradient between the release points and the lateral boundaries of the simulation domain. The pressure heads were set constant in time along the lateral boundaries, while in the central part of the domain the pressure heads changed over time and were controlled mostly by the K_s values of the materials above the flow restrictive layer. Small values of K_s in the calibrated model created conditions for temporary increases in the groundwater level during periods with high precipitation, leading to larger hydraulic gradients between the central and outlying parts of the domain. Since the flux is a product of the conductivity and pressure gradient, an increase in K_s hence is partly compensated for by a lower pressure gradient in the abstracted model.

Figure 5-24 Spatial distributions of log10(C$_{peak}$[ppm]) performance indicator obtained in HYDRUS-3D simulations with the abstracted Ks parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.

Figure 5-25 Spatial distributions of T$_{peak}$[day] performance indicator obtained in HYDRUS-3D simulations with the abstracted Ks parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.

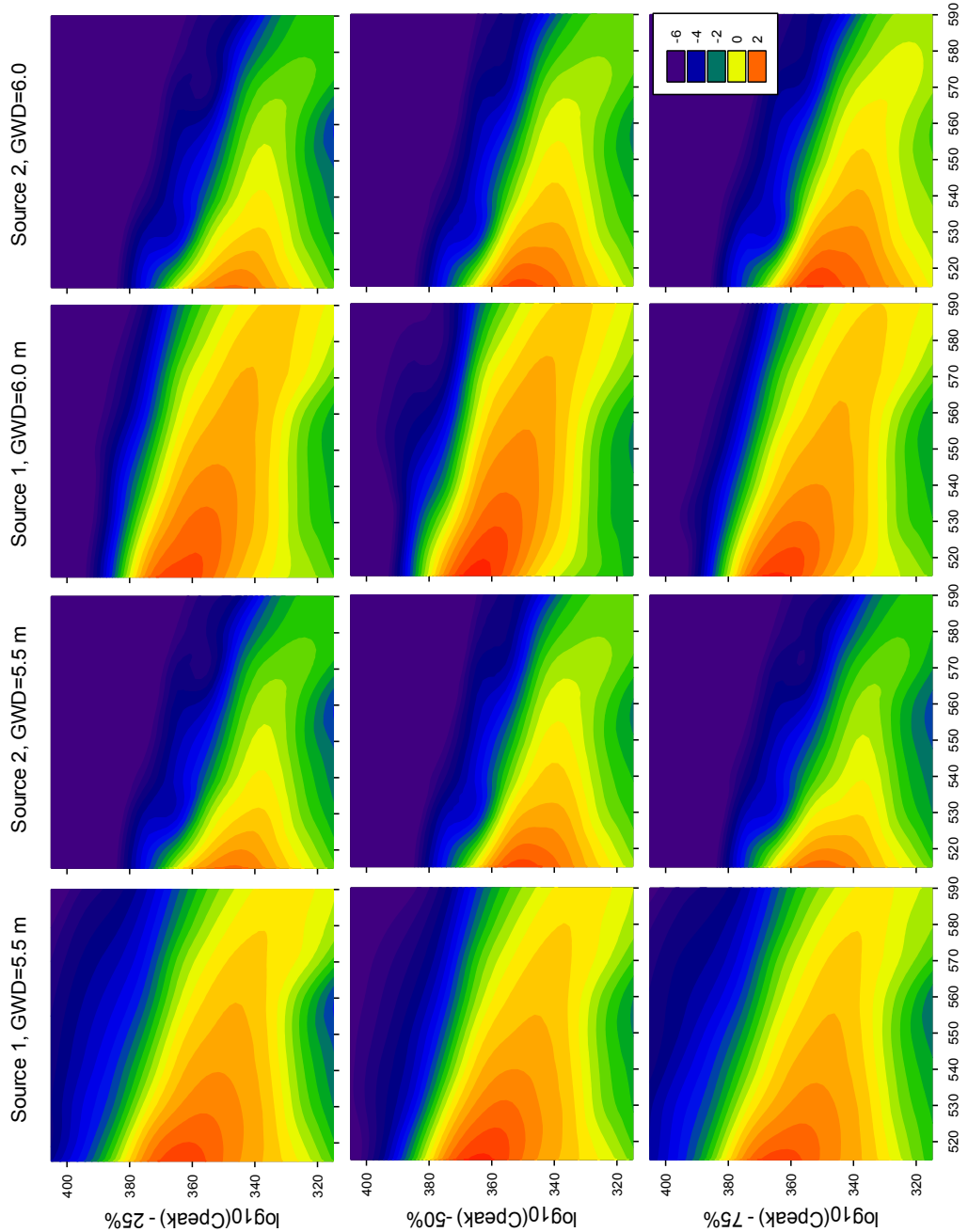**Figure 5-26 Spatial distributions of log10(QC [ppm m day-1]) performance indicator obtained in HYDRUS-3D simulations with the abstracted Ks parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.**

Alteration of the soil hydraulic properties affected the sensitivity of the model to the upper boundary conditions. The spatial patterns of the performance indicators were different for years with 25%, 50% and 75% precipitation probabilities. The C_{peak} and QC spatial distributions were narrower for the 50% probability in the source 1 release scenarios compared to 25% and 75% probability weather scenarios (Fig. 5-24 and Fig. 5-26). A possible reason for this is the inter-decadal precipitation distribution (Fig. 5-20).

Larger C_{peak} values for the abstracted model translated into the larger QC values (Fig. 5-23 and Fig. 5-26). The QC values computed for the 11 transects were from 2 to 10 times larger for the simulations with abstracted parameters than with the calibrated model. The differences were found to increase with distance from the source of the release. The changes were relatively dramatic due to the strong nonlinearity of the advection-dispersion equation used to describe subsurface contaminant transport in this study.

Values of the performance indicators obtained with HYDRUS-3D using the calibrated and abstracted K_s values were used to assess the applicability of the parameter abstraction technique. The ARS study teamcomputed the Nash-Sutliffe and Willmot efficiency criteria and the coefficients of determination using Eqs. (5-8) – (5-10). The performance indicators C_{peak}, T_{peak} and QC as obtained with the calibrated and abstracted parameters served as measured (O) and predicted (P) values in these equations. Values of the efficiency criteria computed for 100 output nodes are shown in Fig. 5-27 a,c,e. Generally, values of the efficiency indices were highest for C_{peak} and lowest for T_{peak} and/or QC, indicating better performance of the abstracted model in predicting peak concentrations as compared to peak times and total fluxes. The E values of QC computed for all scenarios, and of T_{peak} for most scenarios were negative (Fig. 5-27a), which indicates that the mean value of the performance indicator would be a better predictor than the simulated values. Values of the index d and coefficients of determination for C_{peak} ranged from 0.74 to 0.99 and from 0.39 to 0.97, respectively (Fig.5-27 c,e), reflecting a potentially good agreement between the abstracted and calibrated models. High values of the computed indices do not always provide reliable estimate of model accuracy since they are not sensitive to systematic model over- and under-predictions. To test the abstraction results for the presence of systematic errors, the coefficients of linear regression $P = a + bO$ were computed. The computed values of the intercept (a) and slope (b) ranged from -0.03 to 0.69 and from 0.67 to 1.79, respectively indicating the presence of systematic errors and poor correspondence between the abstracted and calibrated models. Systematic errors were also obtained for predicted T_{peak} and QC values, indicating failure of the abstracted model to reproduce the calibrated model.

Values of the efficiency criteria of the performance indicators computed for the transects differed from those computed for the 100 observation nodes. Generally better agreement was found between the calibrated and abstracted simulations when the C_{peak} and QC performance indicators were used for assessing the accuracy of the abstracted model as compared to the T_{peak} indicators. The E values computed for all scenarios for $Tpeak$ and approximately half of the scenarios for the QC indicators were again negative (Fig. 5-27b), indicating poor performance of the abstracted model. The large negative E values computed for QC at the observation nodes became smaller along the transects, while small negative E values computed for $Tpeak$ at the nodes became larger when computed for the transects (Fig. 5-27a and Fig.5-27b). This occurred due to bias of the Nash-Sutcliffe criterion towards large values. The tracer did not arrive at some locations or could arrive only in small amounts at distant nodes 5 to 10 years after the release event. Values of T_{peak} for those locations were set to 3650 days or were close to this number. In the abstracted model the total number of nodes with large T_{peak} values was less as compared
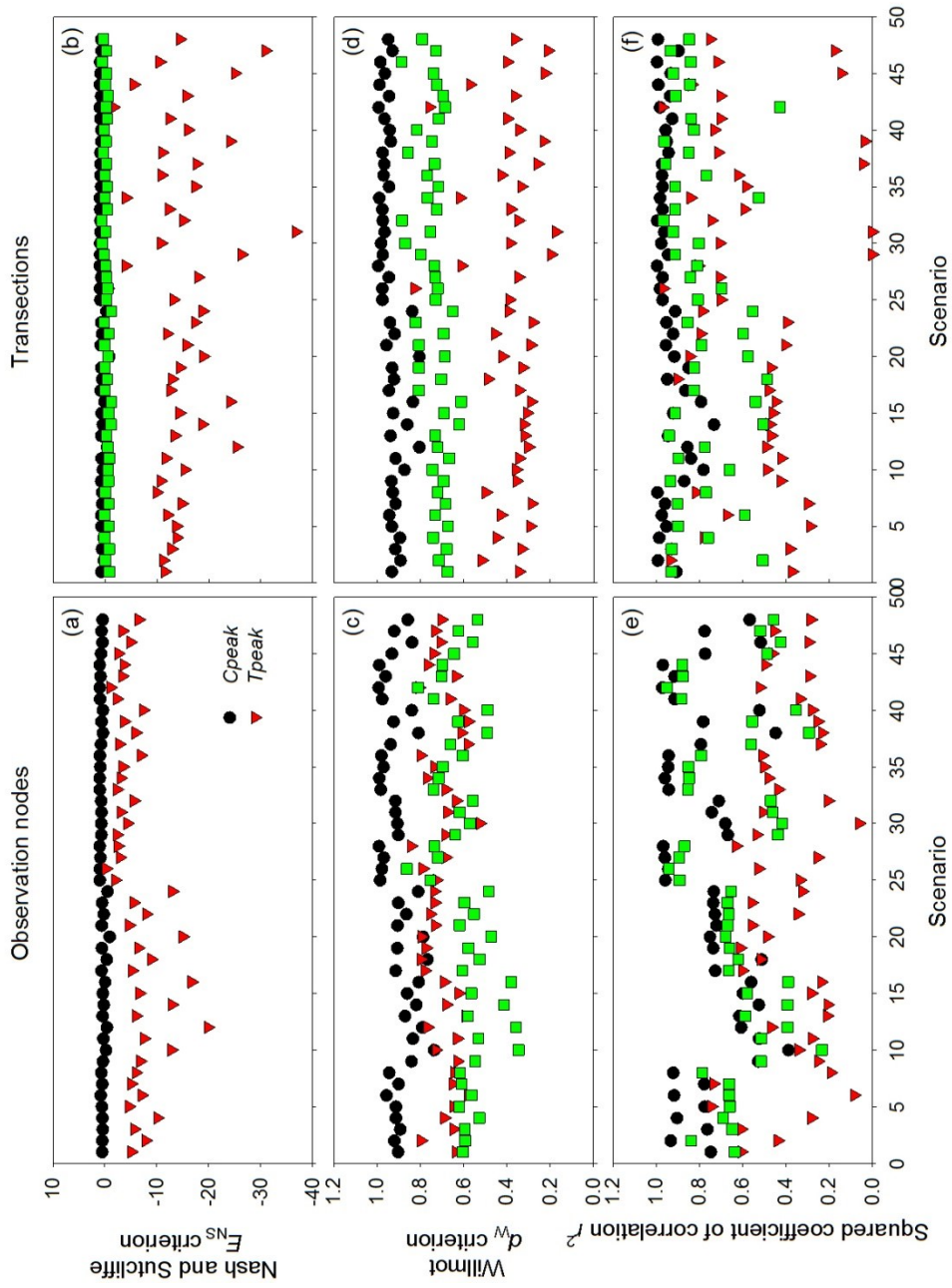
Figure 5-27 Efficiency indices of the parameter determination abstraction.

to the calibrated model (Figs. 5-22 and 5-25). The difference (P-O) in the numerator of Eq. (5-2) was therefore significantly larger than the difference (O-) in the denominator, thus producing high negative E values. The T_{peak} values for the transects were significantly smaller than at the observation nodes for both the calibrated and abstracted models since the tracer had passed through all 11 transects within the 10-year simulation period. Smaller T_{peak} values produced smaller values of the E efficiency criterion shown in Fig. 5-27b.

Similar to results obtained for the 100 observation nodes, values of the Willmot efficiency criteria and coefficients of determination for C_{peak} were reasonably high along all transects, thus suggesting that the abstracted model adequately reproduced the original model (Fig. 5-27b,d,f). For instance, d values ranged from 0.805 to 0.995 and values of R^2 were statistically significant ($P<0.05$) and ranged from 0.733 to 0.996. To confirm these statistics the ARS study team tested the results for the presence of systematic errors. Computed slope and intercept values varied from 0.76 to 1.86 and from 0.23 to 2.79, respectively, indicating the presence of a systematic error and poor performance of the abstracted model.

Overall, the three performance indicators used in this study pointed to inappropriateness of the model parameter abstraction that would use PTF without the appropriate attention to scale effects. The high sensitivity of the HYDRUS-3D modeling results to the saturated hydraulic conductivity resulted in significant changes in flow velocities and concentrations, as well as in total mass fluxes of the tracer transported from the release points assuming different weather, groundwater and release scenarios. These changes were traced using the C_{peak} , T_{peak} and QC performance indicators and assessed using the Nash-Sutcliffe and Willmot efficiency criteria and the coefficients of determination for 100 output nodes and 11 transects. The T_{peak} and QC performance indicators were better discriminators between the calibrated and abstracted models. Both efficiency criteria and the coefficients of determination failed to discriminate the models in terms of the C_{peak} performance indicator due to insensitivity of these statistics to the systematic error.

5.2.6 Abstraction of Profile Aggregation

The second model abstraction technique implemented in this study was profile aggregation. A motivation for this abstraction was dominance of the SL2 material in the simulation domain. To assess the effect of soil profile aggregation on the simulation results, the ARS study team converted the heterogeneous profile of the upper sub-layer into an equivalent homogeneous medium having SL2 soil texture. Soil properties below the flow restrictive layer were kept unchanged. The transport parameters for both layers were taken from the calibrated HYDRUS-3D model. A total of 48 scenarios were run again as described in section 5.2.5. Selected simulation results are presented in Figs. 5-28 - 5-30

Profile aggregation was found to have a major effect on tracer transport. Introducing a homogeneous SL2 profile produced changes in the spatial distribution of the performance indicators. The tracer flux turned approximately 30° to the South, while the shape of the C_{peak} distribution changed from infuser- to diffuser-like. This caused the narrowest part of the tracer plume to occur in the vicinity of the release point (Fig. 5-28). The abstraction also affected concentrations and flow velocities. Generally, the peak concentrations were higher and the travel times towards the output nodes were shorter for the abstracted model as compared to the calibrated model (Fig. 5-28 and Fig. 5-29). The plume spread even faster than in simulations with parameter abstraction discussed in the previous section. The main reason for this was the relatively high saturated hydraulic conductivity (K_s=2.5 m day^{-1}) of material SL2, which provided conditions for fast flow in the entire simulation domain.

Based on visual comparisons of the spatial patterns of the three performance indicators obtained with the abstracted (Fig. 5-28 - Fig. 5-30) and calibrated (Fig. 5-21 - Fig. 5-23) models, one should not expect good agreement between the two models. Indeed the values of the efficiency criteria for profile aggregation (Fig. 5-31) were smaller than those computed for parameter abstraction (Fig. 5-27). Tests for the presence of systematic errors are therefore, not really needed to conclude inadequacy of the model assuming a more homogeneous profile as compared to the calibrated heterogeneous profile model.

5.2.7 Abstraction by Ignoring Unsaturated Zone

The third abstraction technique implemented in this study involved ignoring the unsaturated zone. To do this, the tracer release sources were moved from the soil surface to the saturated zone at 2.1 m depth. In general terms, unsaturated flow in the absence of preferential transport should retard water and chemical fluxes due to the nonlinear decrease in the hydraulic conductivity with decreasing of water content, and the associated long residence time in the vadose zone. Therefore, the ARS study team expected more rapid transport for the abstracted model compared to the original model. The 48 scenarios were run with parameters obtained in the HYDRUS-3D calibration. Results of the simulations are shown in Fig. 5-32 – Fig. 5-35. Surprisingly, ignoring the unsaturated zone did not lead to noticeable changes in the performance indicators. Values of C_{peak}, T_{peak} and QC obtained while ignoring of the unsaturated zone varied within the same ranges as the performance indicators in the original model. The plume shapes in Fig. 5-32 - Fig. 5-34 closely resembled the shapes obtained in simulations with the calibrated parameters (Fig. 5-21 - Fig. 5-23). The similarities between two sets of simulations are quite remarkable, and were confirmed by the efficiency criteria. Values of the Nash and Sutcliffe efficiency coefficients ranged from 0.948 to 0.992, from 0.612 to 0.999, and from 0.956 to 0.994 for the peak concentration, time to peak concentration, and the total flux, respectively (Fig. 5-35). No negative values were obtained for this criterion. The Willmot efficiency coefficients and coefficients of determination were both close to one, which also indicates good agreement between the abstracted and original models.

Values of the efficiency indices for both the observation nodes and the transects were greater for C_{peak} and QC as compared to T_{peak} for several scenarios. For this reason the ARS study team tested all performance indicators for the presence or absence of systematic errors. The tests showed the presence of systematic errors only for the T_{peak} indicator at probability level $P<0.01$. This led us to conclude that there was a delay in the tracer transport in some scenarios and that abstraction by ignoring the unsaturated zone is applicable if the key performance indicators are the peak concentration and the total contaminant flux passing through a transect or observation well. Model simplification by ignoring the unsaturated zone seemed possible in our study because of the presence of perched water and relatively shallow water tables. However, eliminating the unsaturated zone should be done with caution at NRC sites having thicker unsaturated zones.
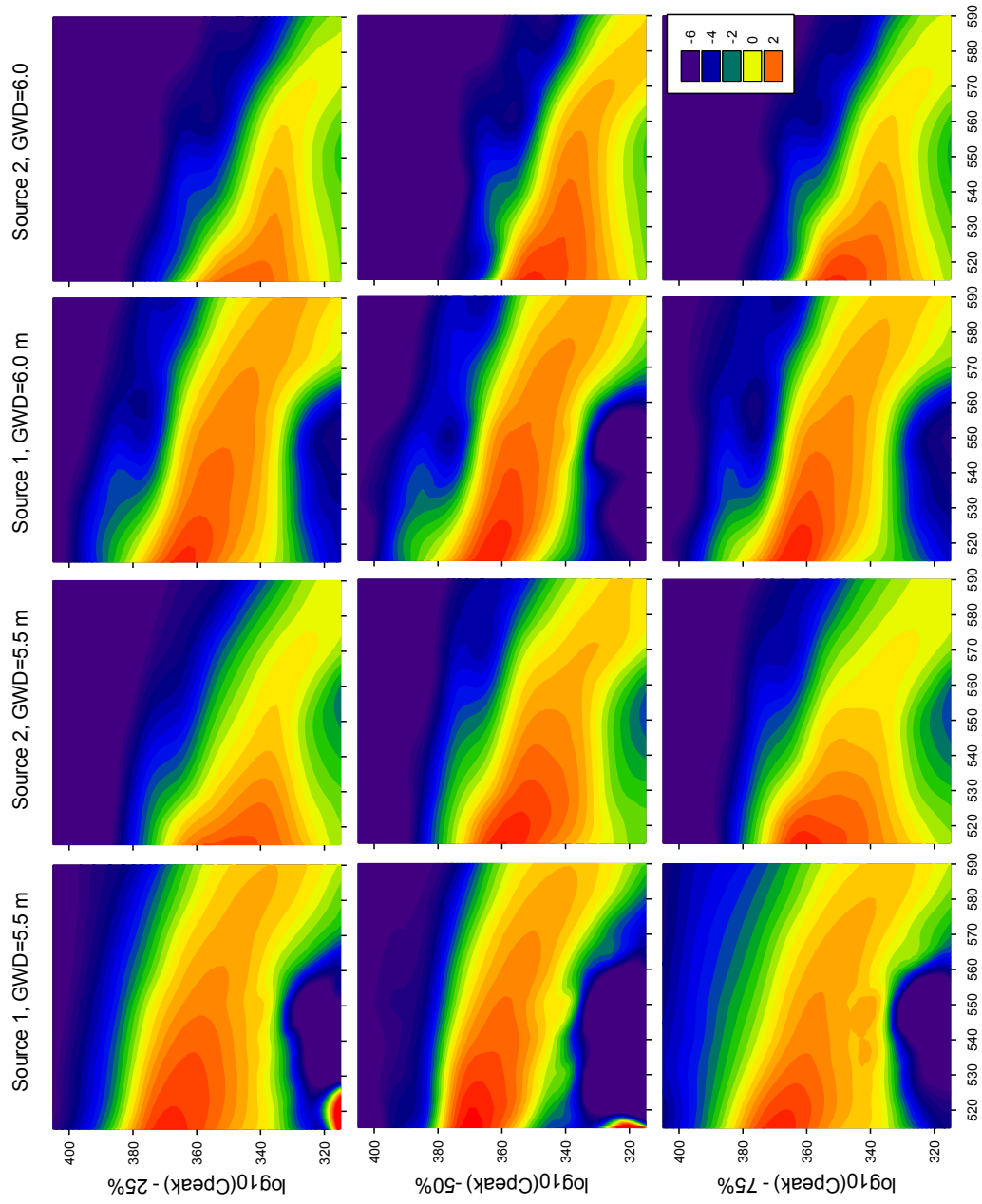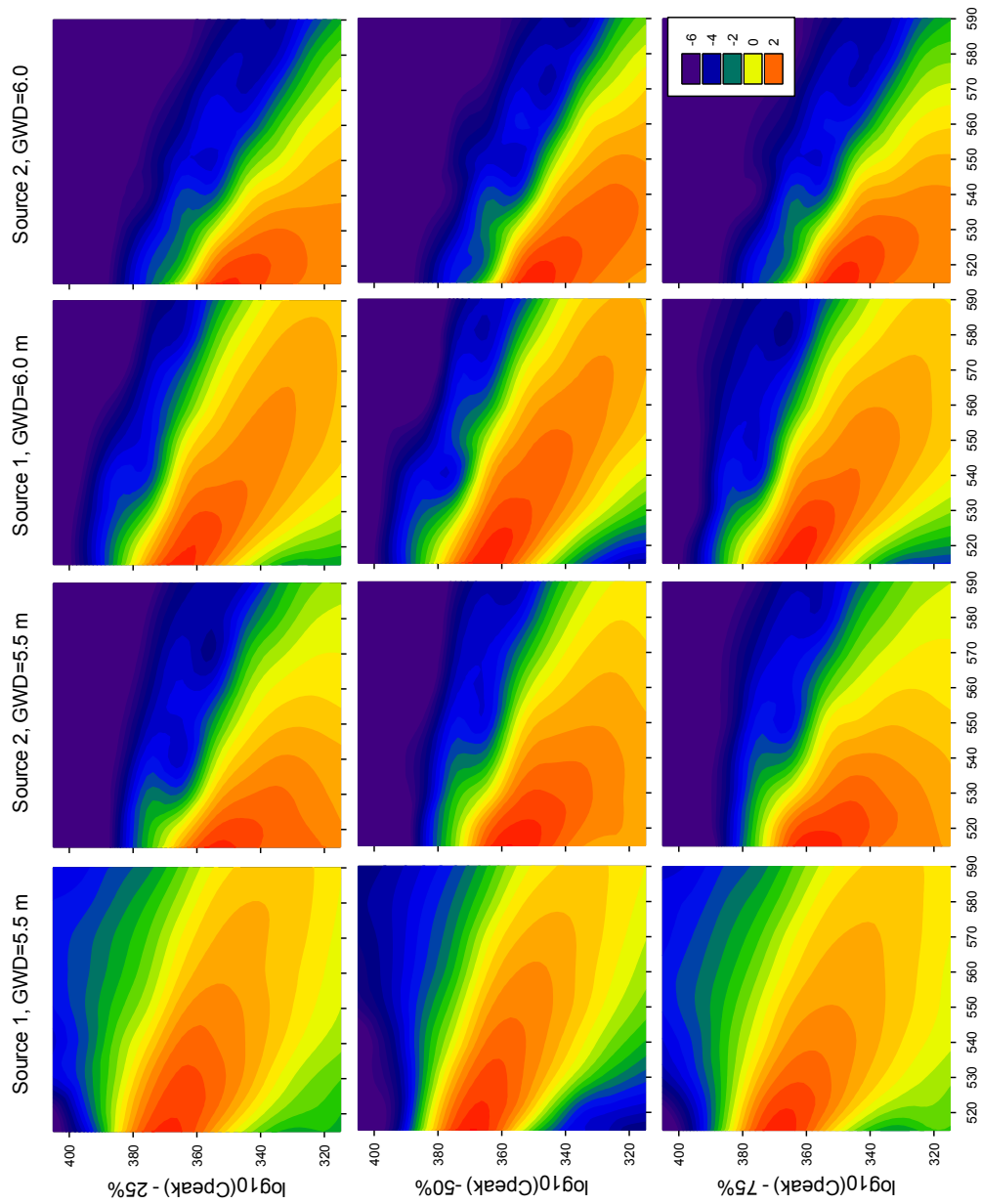
Figure 5-28 Spatial distributions of log10(C_{peak}[ppm]) performance indicator obtained in HYDRUS-3D simulations with profile aggregated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
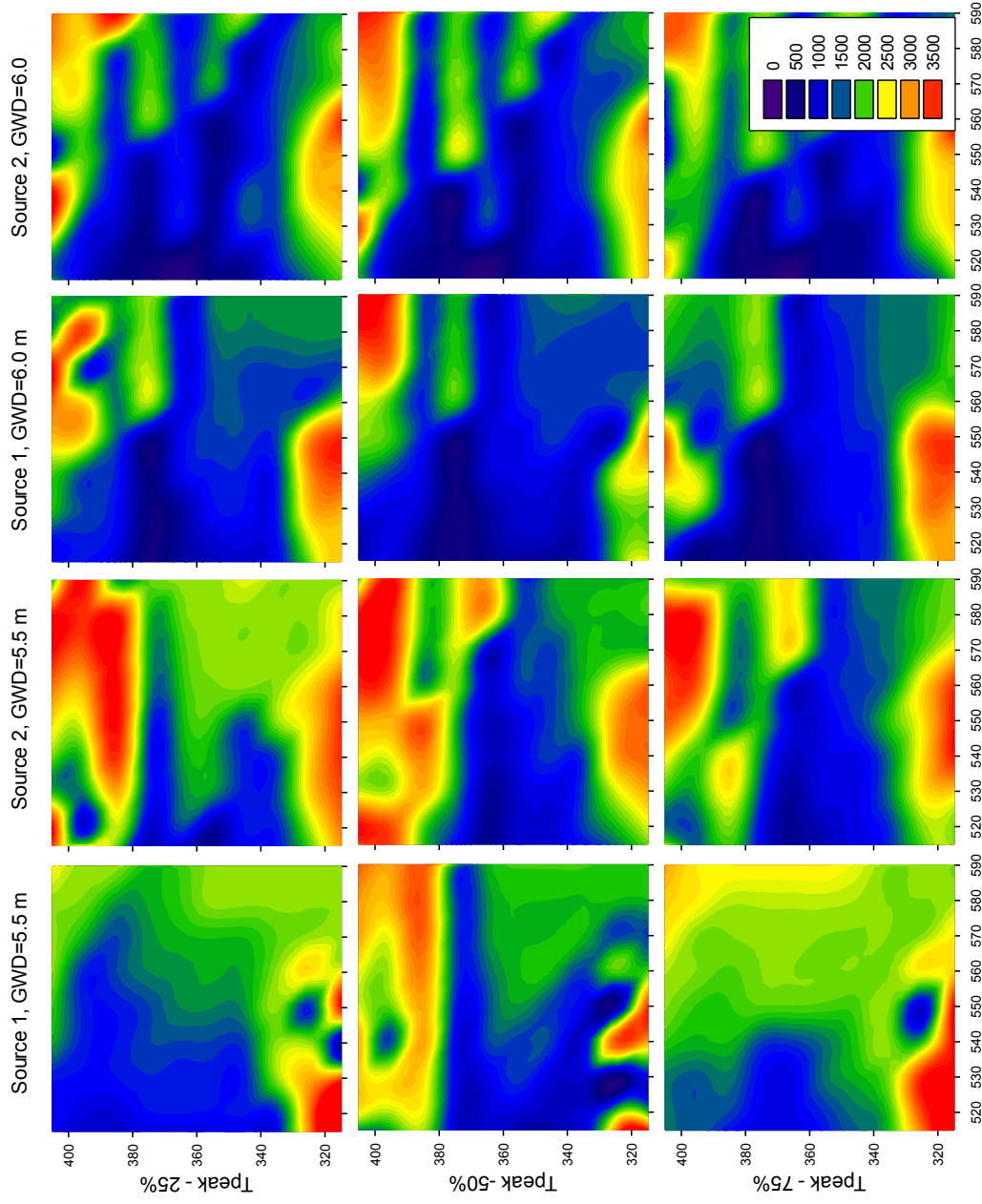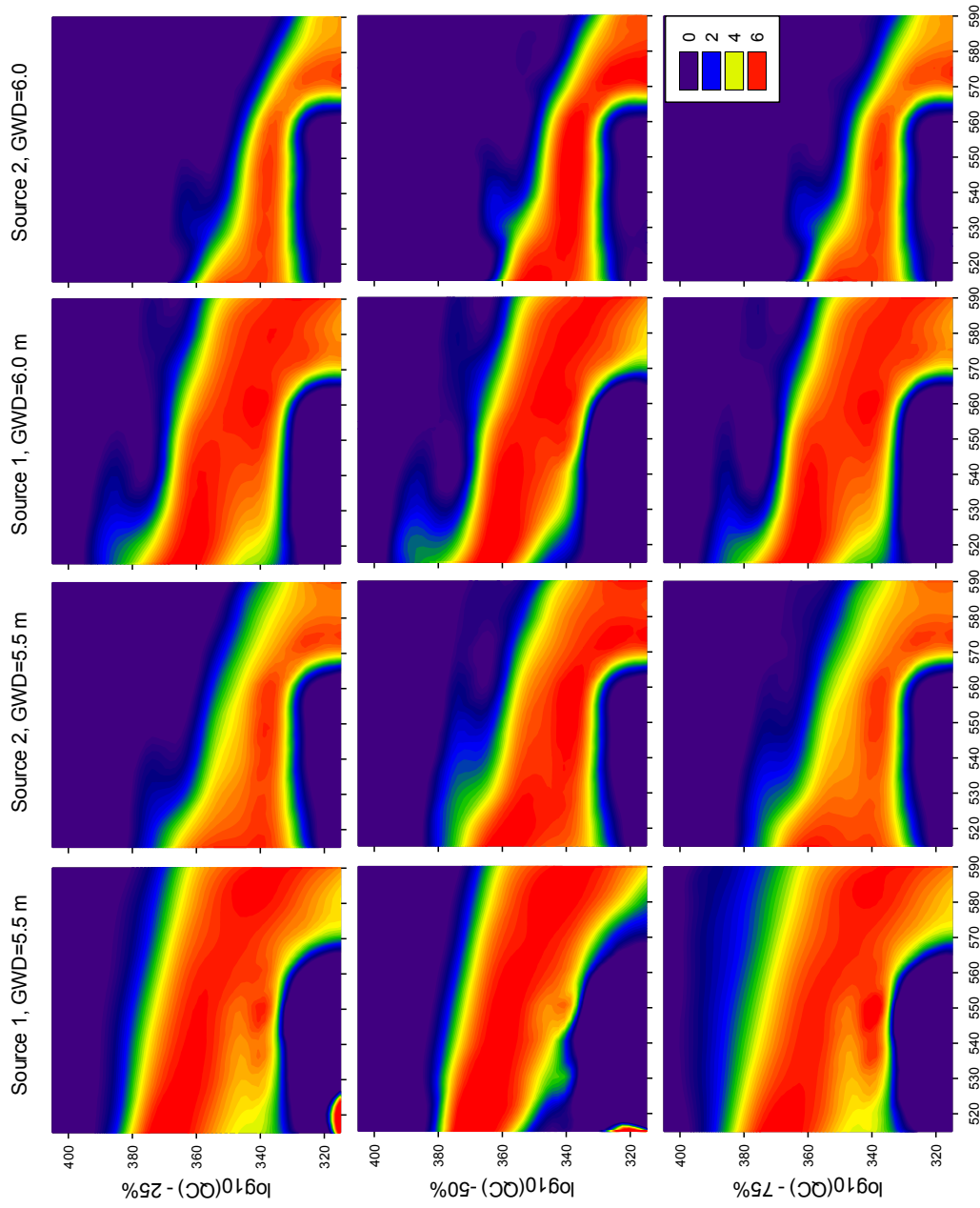
Figure 5-29 **Spatial distributions of T*peak*[day] performance indicator obtained in HYDRUS-3D simulations with profile aggregated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.**
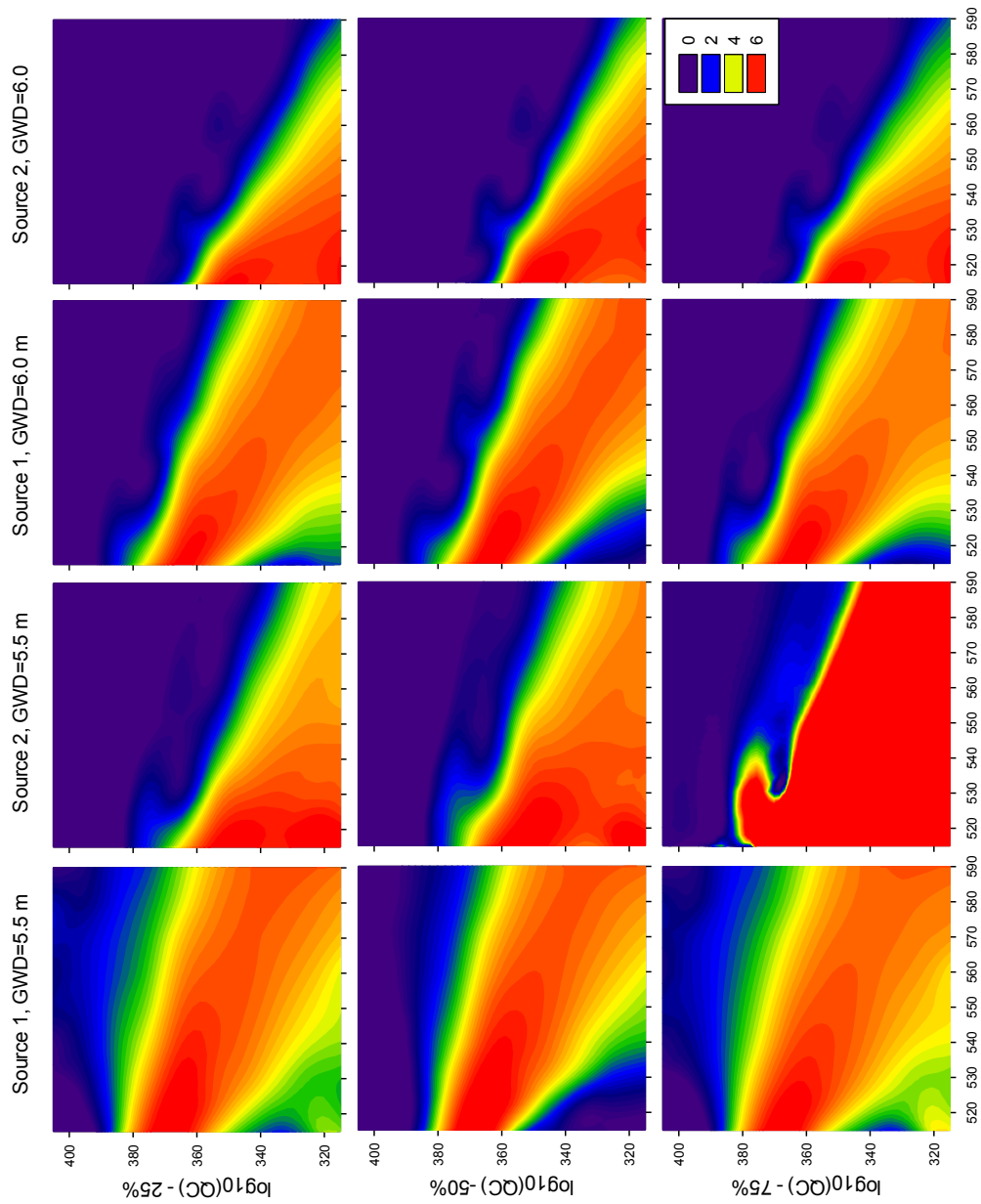
Figure 5-30 Spatial distributions of log10(QC [ppm m day-1]) performance indicator obtained in HYDRUS-3D simulations with the profile aggregated parameters during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
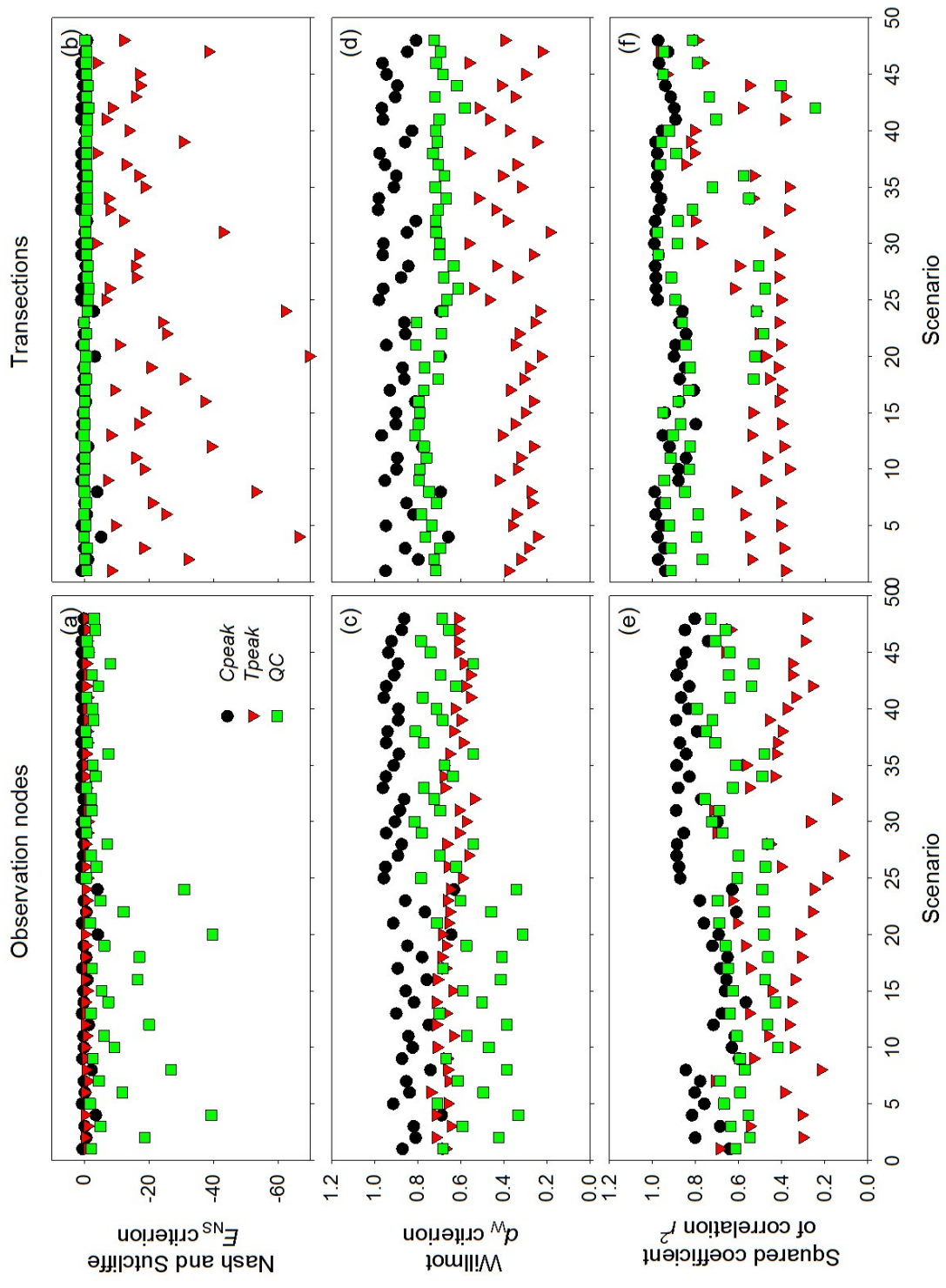
Figure 5-31 Efficiency indices of the profile aggregation abstraction.
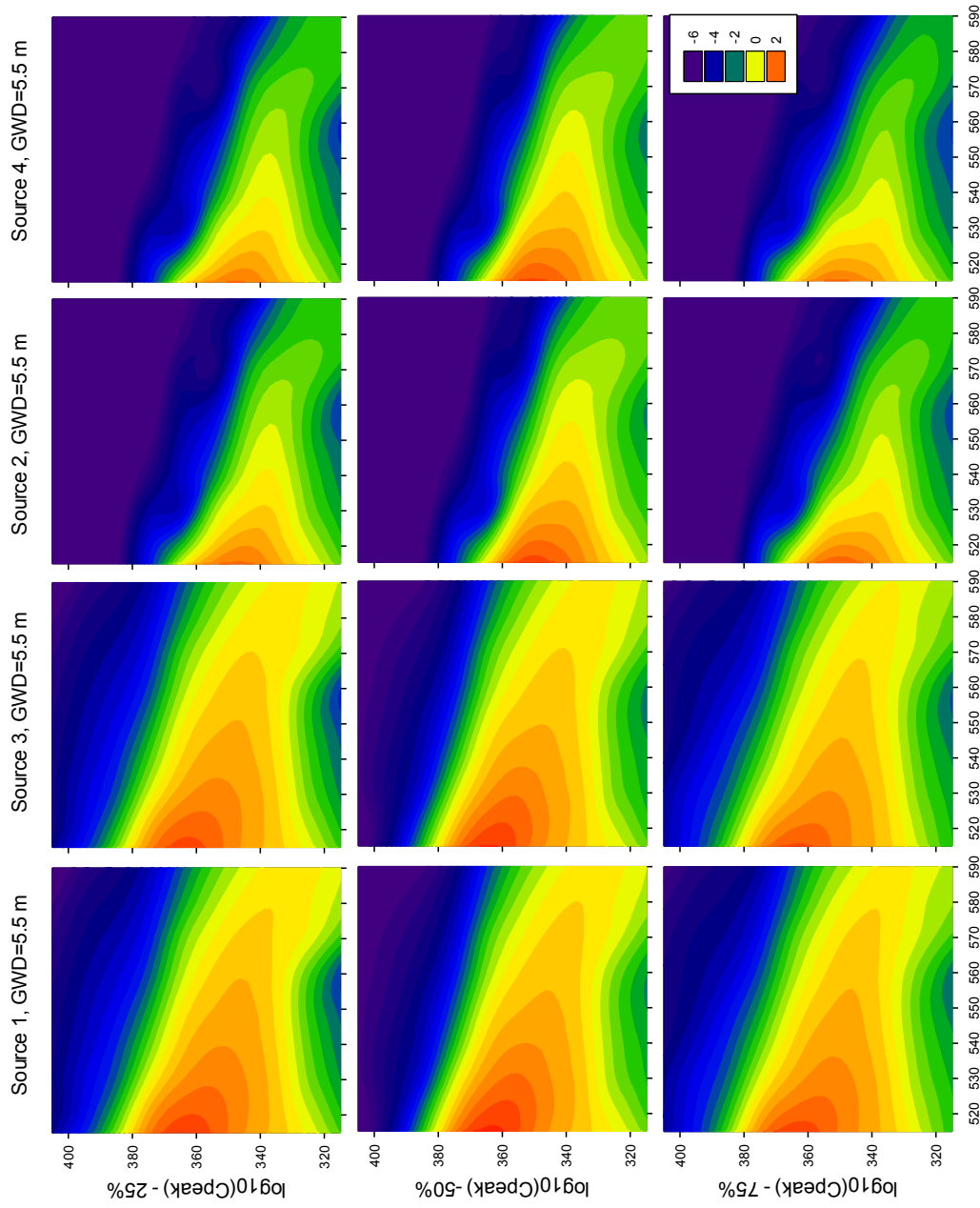
Figure 5-32 Spatial distributions of $\log_{10}(C_{peak}[ppm])$ performance indicator obtained in HYDRUS-3D simulations with ignoring unsaturated zone during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
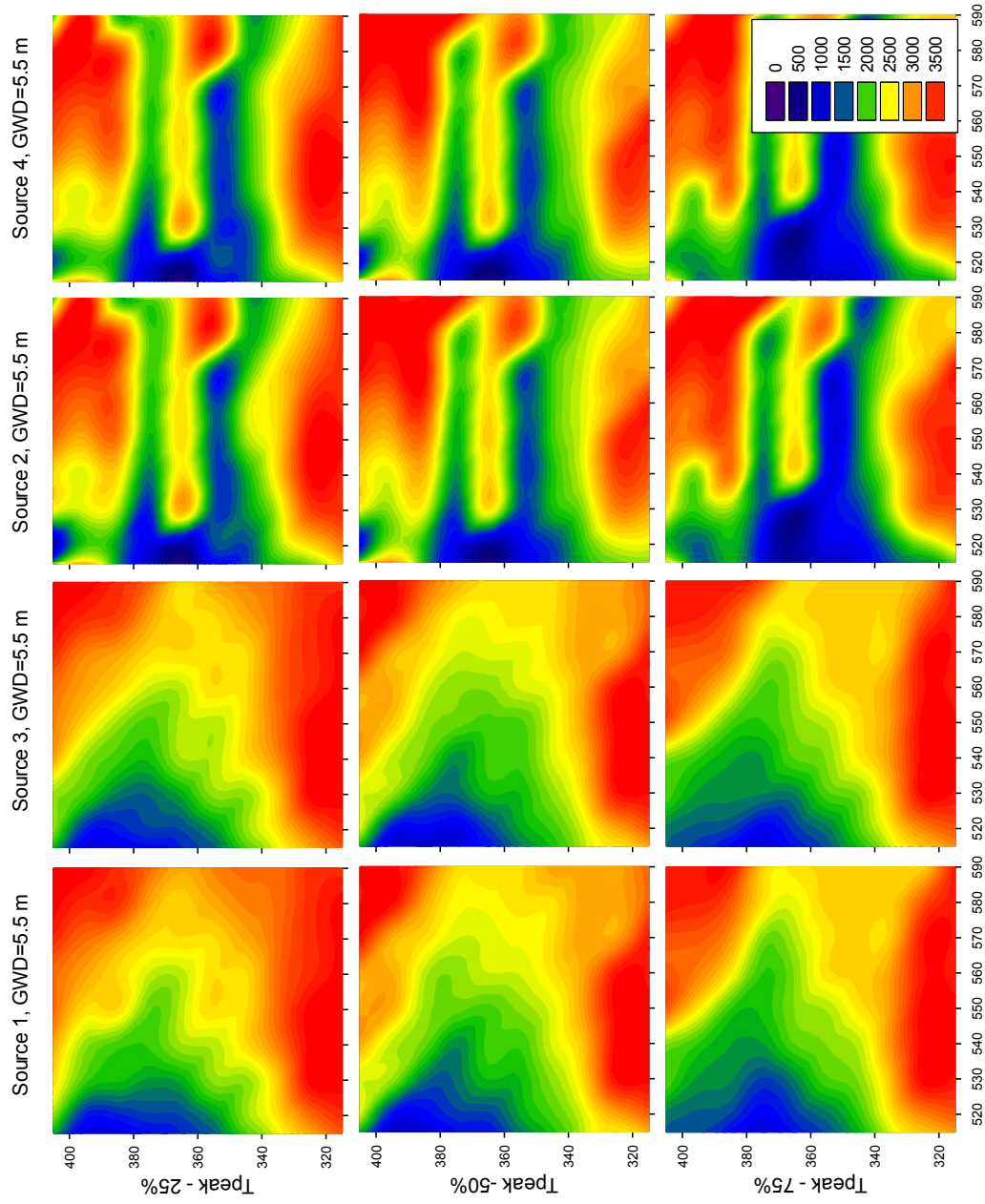
Figure 5-33 Spatial distributions of T $_{peak}$[day] performance indicator obtained in HYDRUS-3D simulations with ignoring unsaturated zone during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
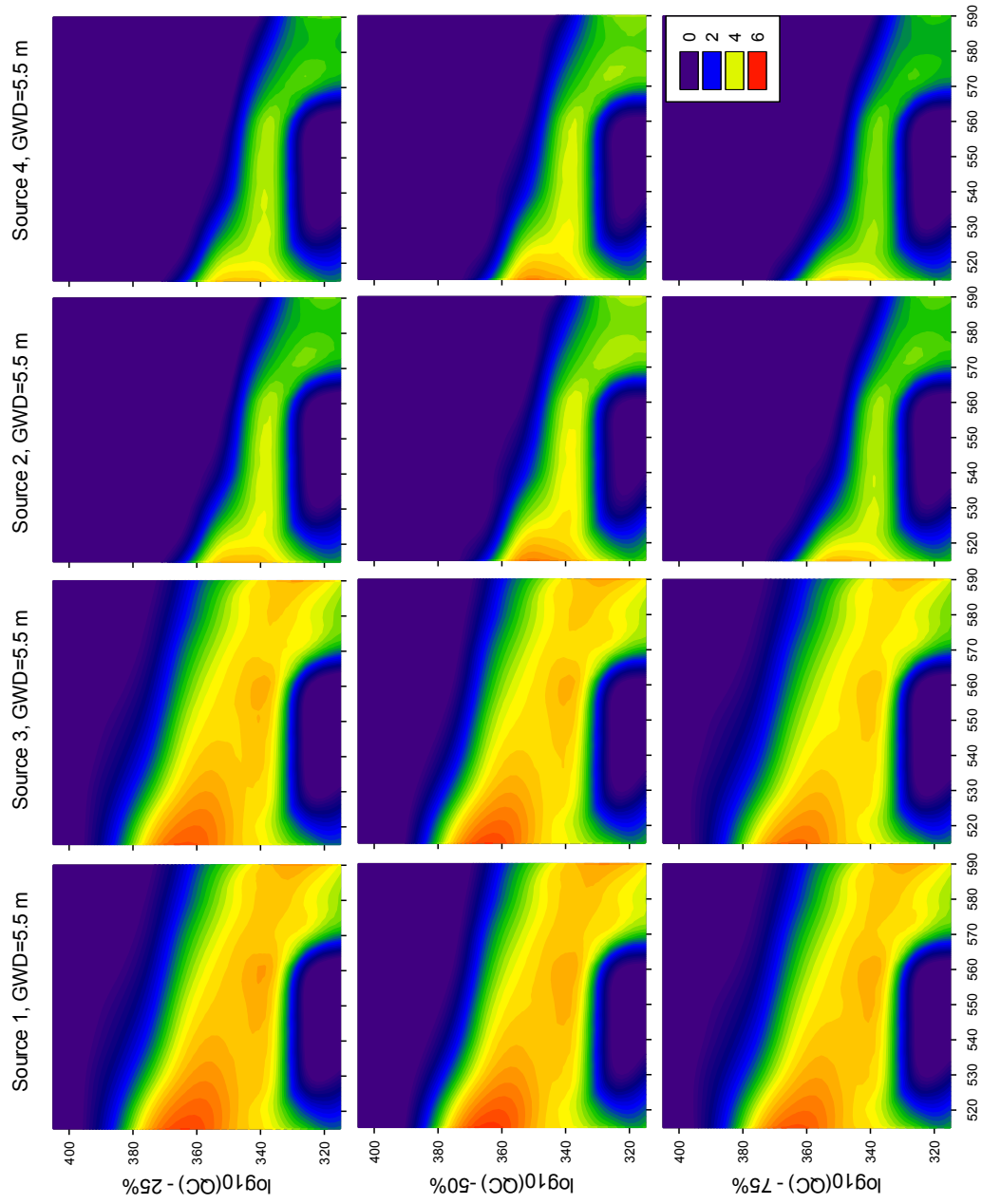
Figure 5-34 Spatial distributions of $\log_{10}$(QC [ppm m day-1]) performance indicator obtained in HYDRUS-3D simulations with limiting input domain abstraction during 10-year simulation period in two scenarios for contaminant source locations, two GWD scenarios and precipitation probabilities of 25%, 50% and 75%.
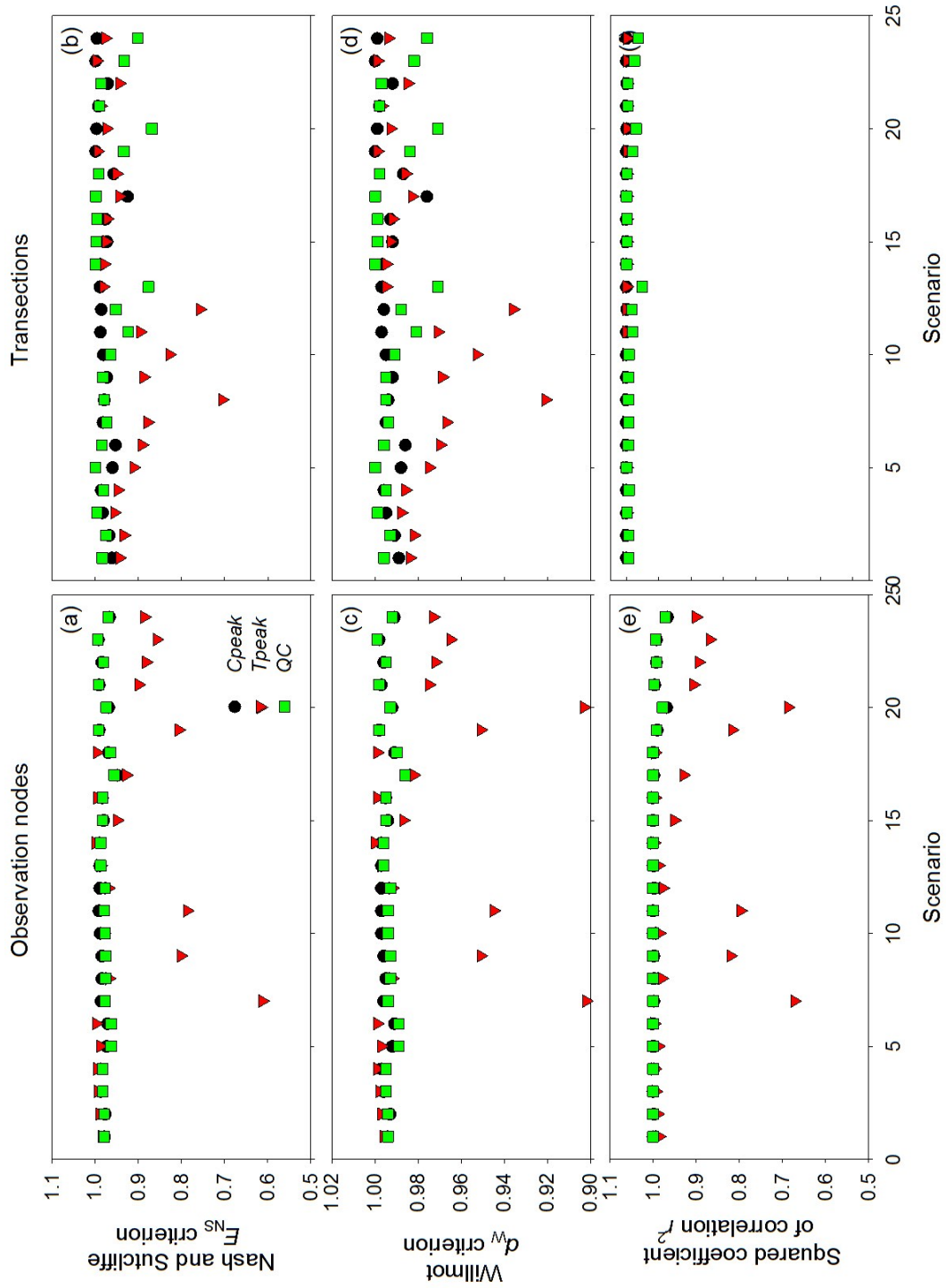
Figure 5-35 Efficiency indices for ignoring the unsaturated zone abstraction.

5.2.8 Abstraction of Parameter Determination: Pedotransfer Functions and Hydraulic Conductivity Scaling

Systematic model abstraction can include the sequential or concurrent application of two or more categories of model abstraction. The ARS study team implemented this using pedotransfer functions and scaling, thus using two categories of parameter estimation abstraction (Fig. 2-1). Simulations were carried out for the smaller flow domain used for model calibration (Section 4.2). This enabled us to compare results of the parameter estimation abstraction with the model calibration results.

Scaling abstraction strives to provide compatibility of the parameter estimation scale with the scale of model resolution. Neuman et al. (2003) suggested that this can be done by upscaling over computational grid cells. Accordingly, scaling abstraction for parameter estimation was implemented by upscaling PTF-estimated hydraulic conductivity values to the computational grid cells. A relatively minor effect of scale on soil water retention was observed (Pachepsky et al., 2001) for soils having similar textures as those in this study, For this reason the ARS study team applied upscaling to the saturated hydraulic conductivity, K_s.

Schulze-Makuch et al. (1999) suggested using a power-law relationship to correlate the increase in K_s with the volume V_m of tested aquifer material as follows

$$K_S = AV_s{}^B \tag{5.11}$$

in which A and B are empirical parameters. The ARS study team reviewed the literature of data sources used by Rawls et al. (1998) to develop pedotransfer functions to estimate K_s as a function of scale and volume. The ARS study team noted a great variety of shapes of aquifer material. The shapes were also quite different from the shapes of computational cells in our numerical study. Direct application of the above volume-based scaling (5.11) was hence not possible. For this reason the ARS study team replaced the volume of affected media by the characteristic length of the medium and calculated the scaled hydraulic conductivity based on reference K_{sR} values estimated using the PTFs of Rawls et al. (1998) using the following relationship

$$K_S^* = K_{SR} \left(\frac{L_g}{L_s}\right)^b \tag{5.12}$$

where L_g is characteristic length of the numerical grid used to solve flow and transport problem, L_s represents the size of the soil cores used to determine K_{sR}, and b is an exponent.

Substantial uncertainty exists in Eq. (5.12) stemming from variations in the lengths of samples used to measure K_s in the Rawls et al. (1998) database. The lengths mostly varied from 0.05 to 0.2 m. The value of the exponent b was also uncertain; Schulze-Makuch et al. (1999) encountered values of B in Eq. (5-11) from 0.40 to 0.95 for homogeneous porous media, from 0.62 to 0.96 for heterogeneous dual- porosity media, and from 0.82 to 1.13 in heterogeneous, fractured media. The ARS study team assumed six abstraction scenarios (Table 5.2) based on a single value for K_{sR} for each soil material, one value (5 m) for L_g, two values (0.05 and 0.2 m) for L_s, and three values (0.3, 0.5 and 0.7) for b. The K_s value in the SiCL3 material was not altered in this abstraction since SiCL3 served as the flow restrictive layer at the bottom of the flow domain.

Results of the simulations are presented in Fig. 5-36 as plots of simulated versus observed concentrations. The concurrent use of the PTF abstraction and scaling abstraction gave results that were very similar to those of the calibration. Visual inspection of Fig. 5-36 suggests that scaling cases A1, A4 and A5 produced relationships between the simulated and observed concentrations that were quite similar to the calibration (Fig. 5-36, CAL). The linear trend lines in Fig. 5-36 represent the correspondence between simulated and observed concentrations

$$C_{obs} = a C_{sim} \qquad (5.13)$$

where a is the slope of linear fit. This slope differed most significantly from unity for scaling scenarios A2, A3, and A6.

Figure 5-37 shows values of the determination coefficient (R^2) of the regressions and values of the slopes a in (5.13). Values of R^2 for scenarios A1, A4 and A5 were relatively close to the calibration scenario, with slopes closest to 1. The ARS study team concluded that the K_s values of Rawls et al. (1998) must be scaled using $(L_g/L_s)^b$ from the soil core size for which hydraulic conductivity was determined to the representative numerical grid size, with the exponent b varying between 0.3 and 0.5 for small to large soil core sizes.
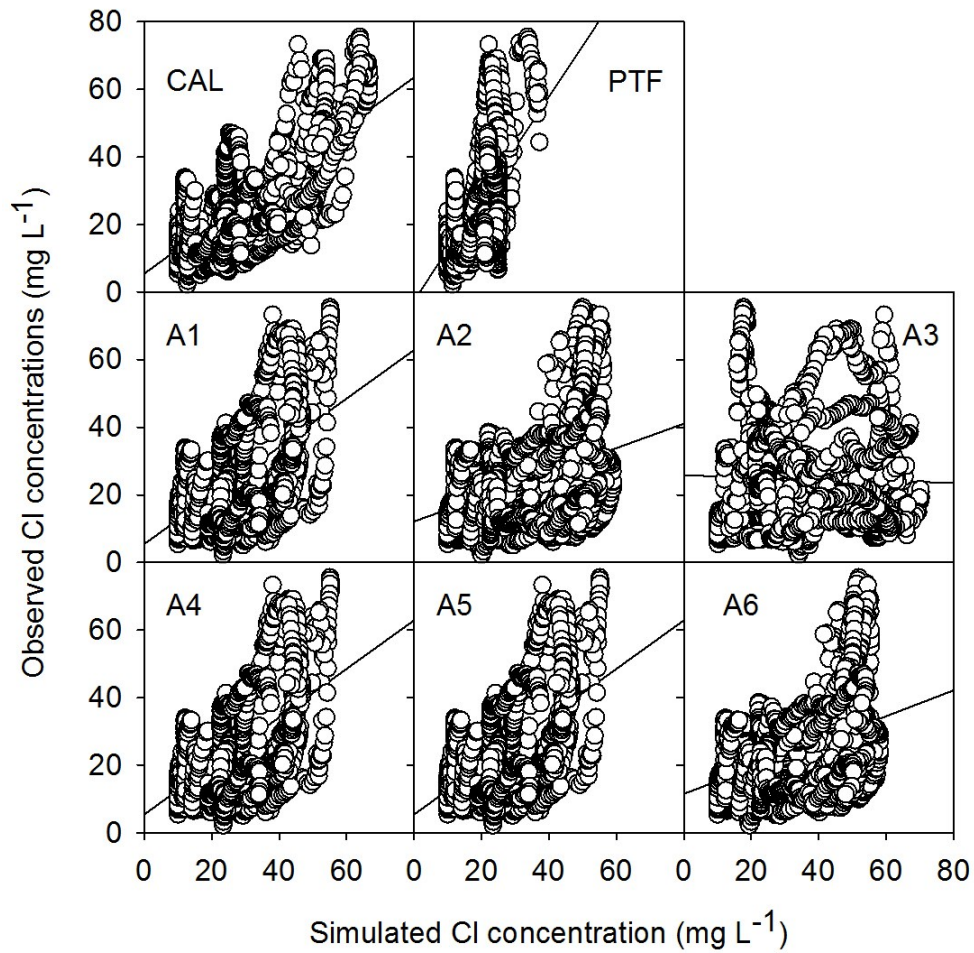
Figure 5-36 Observed and simulated chloride concentrations in locations 5 through 9 with saturated hydraulic conductivity values obtained with different parameterization methods: CAL – manual calibration, PTF - using pedotransfer functions of Rawls et al. (1998), A1 through A6, PTF-estimated valued scaled to the size of the computational cells.
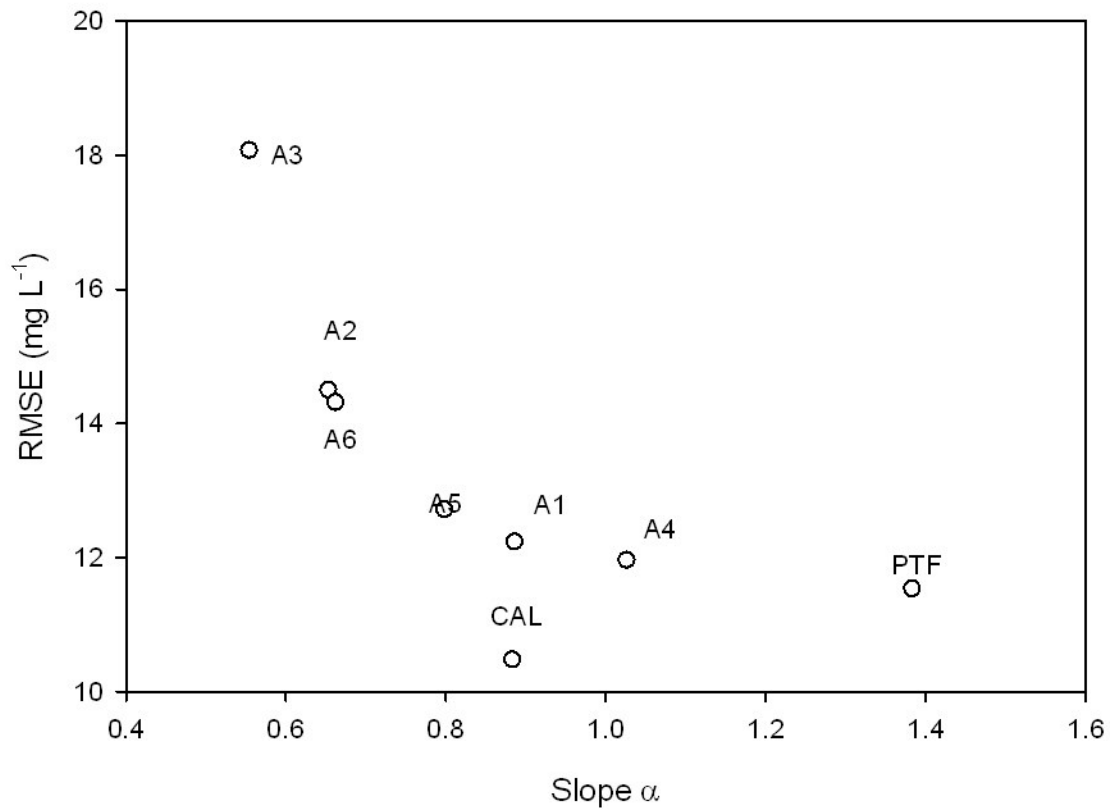
Figure 5-37 Root-mean-square errors (RMSE) and slopes of regressions ' simulated vs. observed chloride concentrations' in locations 5 through 9 with saturated hydraulic conductivity values obtained with different parameterization methods: CAL – manual calibration, PTF - using pedotransfer functions of Rawls et al. (1998), A1 through A6 - PTF-estimated valued scaled to the size of the computational cells. Regressions were computed with the intercept set to zero.

6 APPLICATIONS OF MODEL ABSTRACTION IN MONITORING NETWORK DESIGN

Models developed using abstractions based on model structure will have different underlying conceptual representations. Since a monitoring network is always designed based on a conceptual model, designs may be different for different abstracted models. As evidence is growing that conceptually different models can provide similar accuracy, developments in environmental modeling technologies is moving in two directions: (a) multimodeling in which several models are used simultaneously, in particular using ensembles of models, and (b) ranking models and selecting the most plausible model or models. Systematic model abstraction enables one to provide sets of models and ensembles of models for both of these modeling strategies.

Integrating model abstraction in monitoring network design follows the same two strategies. Monitoring network design can either jointly use designs resulting from several conceptual models, or develop a design to discriminate models with the goal of selecting the more plausible one. This section describes two methods that the ARS study team developed to implement both strategies.

6.1 Selecting Monitoring Locations to Decrease Prediction Uncertainty

6.1.1 A Linear Statistical Inference Method to Assess Data Importance to Model Predictions, Sensitivity and Model Abstraction

The purpose of conducting sensitivity analyses for model calibration is to assess the value of the data collected from monitoring wells in terms of their usefulness for model calibration and to identify new locations that could improve the reliability of the calibration. Generally three-dimensional models require spatially distributed parameters that characterize transport properties of the materials composing the simulation domain. Ideally, the monitoring locations should be selected so as to provide sufficient and reliable information for determining the hydraulic and transport parameters of each soil material. To comply with these requirements, data collected at specific locations must vary temporally within a range that corresponds to the sensitivity of the model to the transport properties of the selected material, and have minimum interference with the parameters of other materials. Relatively small temporal variability typically can lead to high parameter uncertainty, while the interference produces correlation between the parameters of different soil materials. The best possible information therefore could be obtained from a dense monitoring network covering areas close to the tracer application point as well as wells and depths embracing all possible soil materials within the simulation domain. Unfortunately this implies a spatial density of the monitoring network that is impractical. Monitoring data are normally collected from a limited number of locations with soil properties generally not well known at the time of data collection. In this case a sensitivity analysis can be used to assess the value of the collected data for model calibration and to identify new locations that are most likely to produce information necessary for improving the calibration process.

To evaluate the importance of observations for improving model predictions, Draper and Smith (1981), and Hill (1998) proposed the use of a linear statistical inference equation for calculating prediction standard deviations:

$$S_{Y_l'} = \left[S^2 (\underline{A}_Y \left(\underline{A}^T \omega \underline{A} \right)^{-1} \underline{A}_Y^T)_{ll} \right]^{1/2} \tag{6-1}$$

where

Y_l' is the lth predicted value;

$S_{Y_l'}$ is the standard deviation of Y_l';

S^2 is the calculated error variance from the model calibration;

$\underline{A}_Y$ is the sensitivities of the predictions with respect to the model parameters;

$\underline{A}$ is the matrix of sensitivities of the simulated equivalents of the observations (Y_i') with respect to the model parameters;

T indicates the transpose of the matrix;

$\left(\underline{A}^T w \underline{A}\right)^{-1}$ is a symmetric square linear approximation of the parameter variance-covariance matrix;

X_j is the jth parameter; $\underline{\omega}$ is the matrix of weights on observations used in the calibration.

Equation (6-1) was used to evaluate the importance of existing hydraulic-head observation locations to advective-transport predictions for the Death Valley regional flow system in California and Nevada (Hill et al., 2001) for the stationary flow regime. Elements of the matrix of sensitivities $\underline{A}_Y$ were calculated as derivatives:

$$\underline{A}_Y = \frac{\partial Y_l'}{\partial X_j} \tag{6-2}$$

Standard deviations describing the precision of the water level measurements were used to calculate the weights in $\underline{\omega}$ in that study.

To evaluate the effect of omitting an observation location on the prediction uncertainty, Hill et al. (2001) used a modified version of equation (6-1) as follows:

$$S_{Y_{l(-i)}'} = \left[S^2 (\underline{A}_Y \left(\underline{A}_{(-i)}^T \underline{w}_{(-i)} \underline{A}_{(-i)}\right)^{-1} \underline{A}_Y^T)_{ll} \right]^{\frac{1}{2}} \tag{6-3}$$

in which $S_{Y_{l(-i)}'}$ is the standard deviation of the l simulated value, Y_l', calculated without the ith observation location; $\left(\underline{A}_{(-i)}^T \underline{w}_{(-i)} \underline{A}_{(-i)}\right)^{-1}$ is a symmetric, square parameter variance-covariance matrix calculated with the information for the ith observation omitted;

$\underline{A}_{Y(-i)}$ is the matrix of sensitivities of the simulated equivalents of the observations (Y_i') with respect to the model parameters, with the sensitivities for the ith observation omitted;

$\underline{\omega}_{(-i)}$ is the diagonal matrix of weights on observations used in the calibration, with the value for the ith observation omitted.

The relative importance of an observation for a prediction was defined as the percent increase in uncertainty caused by omitting an observation, and was termed the Observation-Prediction (OPR⁻) statistic:

$$OPR^- = \left[(S_{Y_{l(-i)}'}/S_{Y_l'}) - 1.0 \right] \times 100 \tag{6-4}$$

Similarly to equation (6-4) the relative importance of a new observation for a prediction can be calculated as:

$$OPR^+ = \left[(S_{Y_{l(+i)}'}/S_{Y_l'}) - 1.0 \right] \times 100 \tag{6-5}$$

where $S_{Y_{l(+i)}'}$ is the standard deviation of the l simulated value, Y_l', calculated with the ith observation location added;

For the case of adding observations, $S_{Y'_{l(+i)}}$ is calculated as:

$$S_{Y'_{l(+i)}} = \left[S^2 (\underline{A}_Y \left(\underline{A}^T_{(+i)} \underline{w}_{(+i)} \underline{A}_{(+i)} \right)^{-1} \underline{A}^T_Y)_{ll} \right]^{\frac{1}{2}} \tag{6-6}$$

where index *(+i)* identifies the added location.

The applicability of the linear statistic method (equations (6-1) through (6-6)) to evaluation of the importance of the existing and new water head and flow observations for a hypothetical groundwater flow system was demonstrated in Tonkin et al. (2007).
The applicability of the linear statistic method given by equations (6-1) through (6-6) to evaluation of the importance of existing and new pressure head and flow observations for a hypothetical groundwater flow system was demonstrated in Tonkin et al. (2007). The approach could not be literally transferred to our work since it assumes flow stationarity, which is not the case in our study. Therefore, the ARS team used the first-order sensitivity coefficient (Eq. 6-3) as a substitute for the parametric sensitivity (6-2). Values of the first-order sensitivity coefficient were averaged over the observation period.

6.1.2 Observation-Prediction Statistics Application

The Observation-Prediction statistics (OPR) method described in section 6.1.1 was implemented to evaluate the importance of observations collected in the monitoring wells during the lateral flow experiment for the HYDRUS-3D calibration. HYDRUS-3D was run with the parameters and boundary conditions obtained with the model calibration. To evaluate the importance of the observation locations for the HYDRUS-3D calibration, and specifically for saturated hydraulic conductivity, values of K_s of 8 soil materials within the simulation domain were varied similarly as for the model sensitivity analysis in section 5.1.3. The only difference between the simulations in section 5.1.3 and the current HYDRUS-3D runs was the initial profile distribution of the tracer. Initial C_0 values were set to zero to eliminate the influence of spatially variable initial concentrations on the sensitivity index values. A total of 256 simulations were carried out assuming two combinations of the K_s parameters of 8 materials. The computed tracer concentrations were recorded at 100 observation nodes to provide input for the sensitivity and OPR statistics. The observation (output) nodes were placed on the regular rectangular grid within a 27x14 m rectangular area at depths *z* of 2.07 and 2.49 m (Fig. 6-1). The eight locations (wells 5 through 12) used for the model calibration and sensitivity analysis in section 5.1.3 constituted the existing observation network for the OPR analysis.

Values of the sensitivity index $S_i(K_s)$ were found to be different for the same soil material at two depths and 50 locations, indicating different model sensitivities to K_s associated with the specific location of the observation node in the simulation domain. Generally, zero S_i values were obtained for times smaller than the tracer arrival times at each location. For instance, nonzero S_i values were observed starting from days 88 and 118 of the simulation in wells at distances of 3.5 and 14 m from the irrigated plot, respectively (Fig. 6-2). This shows that the distance from the irrigated plot was the major factor influencing S_i values at selected locations. Sensitivity values also changed with depth. Temporal persistence in the S_i time series was observed at some locations. For example, the, S_i values of well 38 decreased for 8 soil materials in the order L1, SiCL1, SCL2, SCL1, SL1, L2, SL1 and SL3 at both computational layers (i.e. *z*=2.07 and 2.49 m). Persistence with depth at other locations existed only for some materials. For example, S_i values were persistent for 6 out of 8 soil materials at well 5 (Fig. 6-2).

The effect of depth and distance can be clearly seen in Fig. 6-3. Generally, the number of simulation days with nonzero S_i values, and S_i values averaged among the materials, were greater for the computational layers at 2.49 m as compared to 2.07 m layer (Fig. 6-3). The number of nonzero S_i days, the range and the average S_i values tended to decrease with distance from the irrigated plot, reflecting a general decrease in the model sensitivity with tracer travel distance. Notwithstanding this tendency, several points among the observation wells located at equal distances from the irrigated plot showed consistently high and low S_i values. Since locations with high S_i values are more informative and more important compared to locations having low S_i values, they could be included into the existing monitoring network for improvement of the model calibration. The potential locations selected solely based on the S_i values are shown in Fig. 6-4.

The importance of existing observation locations for the HYDRUS-3D calibration was evaluated using the OPR$^-$ statistic. Averaged over time, the S_i values obtained with HYDRUS-3D were used to calculate the matrix of sensitivities $\underline{A}_Y$ in Eq. 6-1. The standard deviations were computed separately for each well used for the HYDRUS-3D calibration (wells 5 through 12) based on Eq. 6-1. The matrix of weights $\underline{\omega}$ on the observations was represented by the proportion of days showing nonzero S_i values during the 131 day simulation. Consequently, wells with fewer nonzero S_i values had higher weights in the computations. To evaluate the effect of omitting an observation location on the prediction uncertainty, the deviations $S_{Y'_{l(-i)}}$ were computed with weights of the omitted location set equal to zero in Eq. (6-1). Finally, the ARS study team calculated the relative importance of each observation on the predictions using Eq. 6-4.

The percent of increased uncertainty OPR$^-$ ranged from 0 to 28.7%, and was greater for z=2.49 m than for the z=2.07 m layer (Fig. 6-5). Relatively high OPR$^-$values were obtained at the locations of wells 6, 11 and well 12 (z=2.49 m layer) and at well 6 (z=2.07 m layer), indicating the importance of data collected at those places for model calibration. Data collected near wells 7 and 8, and at all locations of the z=2.07 m layer, except well 6, were the least important value for model calibration. This result implies that it is not necessary to collect data at two depths for the model calibration, and emphasizes the overall importance of monitoring at the depth corresponded to the 2.49 m computational layer.
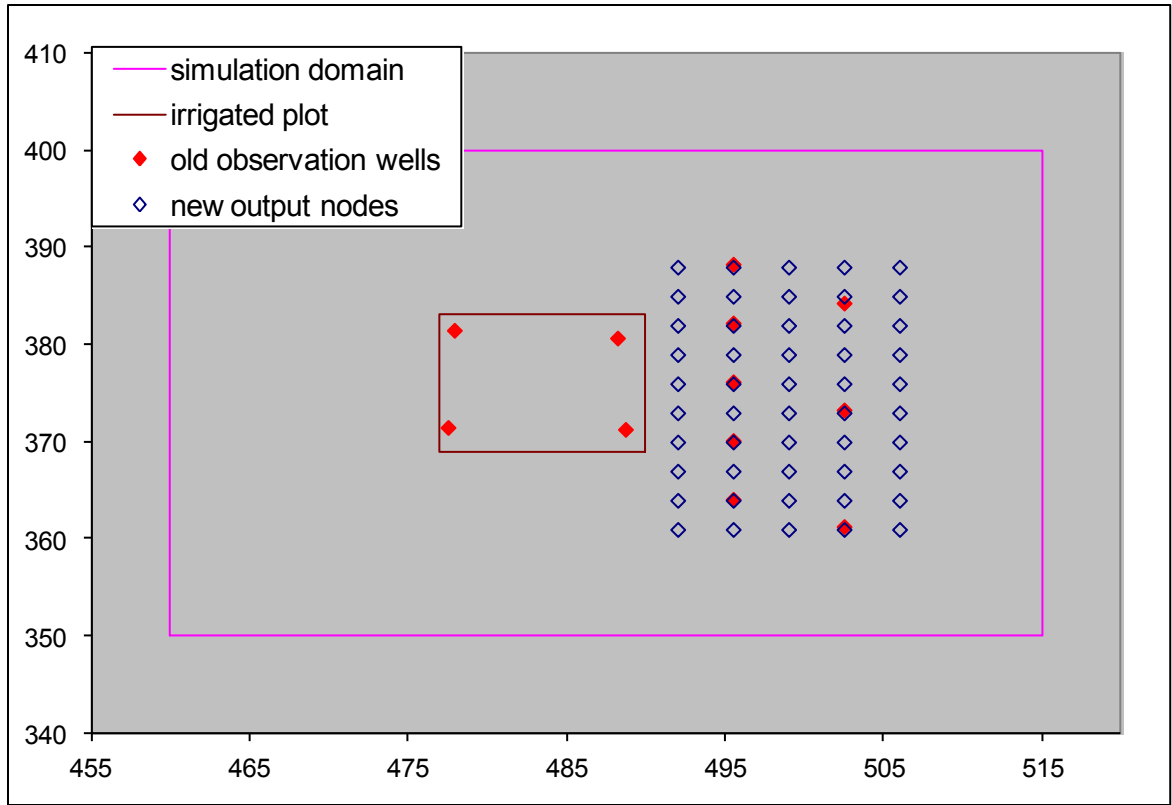
Figure 6-1   Simulation domain with existing (red symbols) and new (hypothetical) observation wells. The numbers at the plot identify new well numbers.
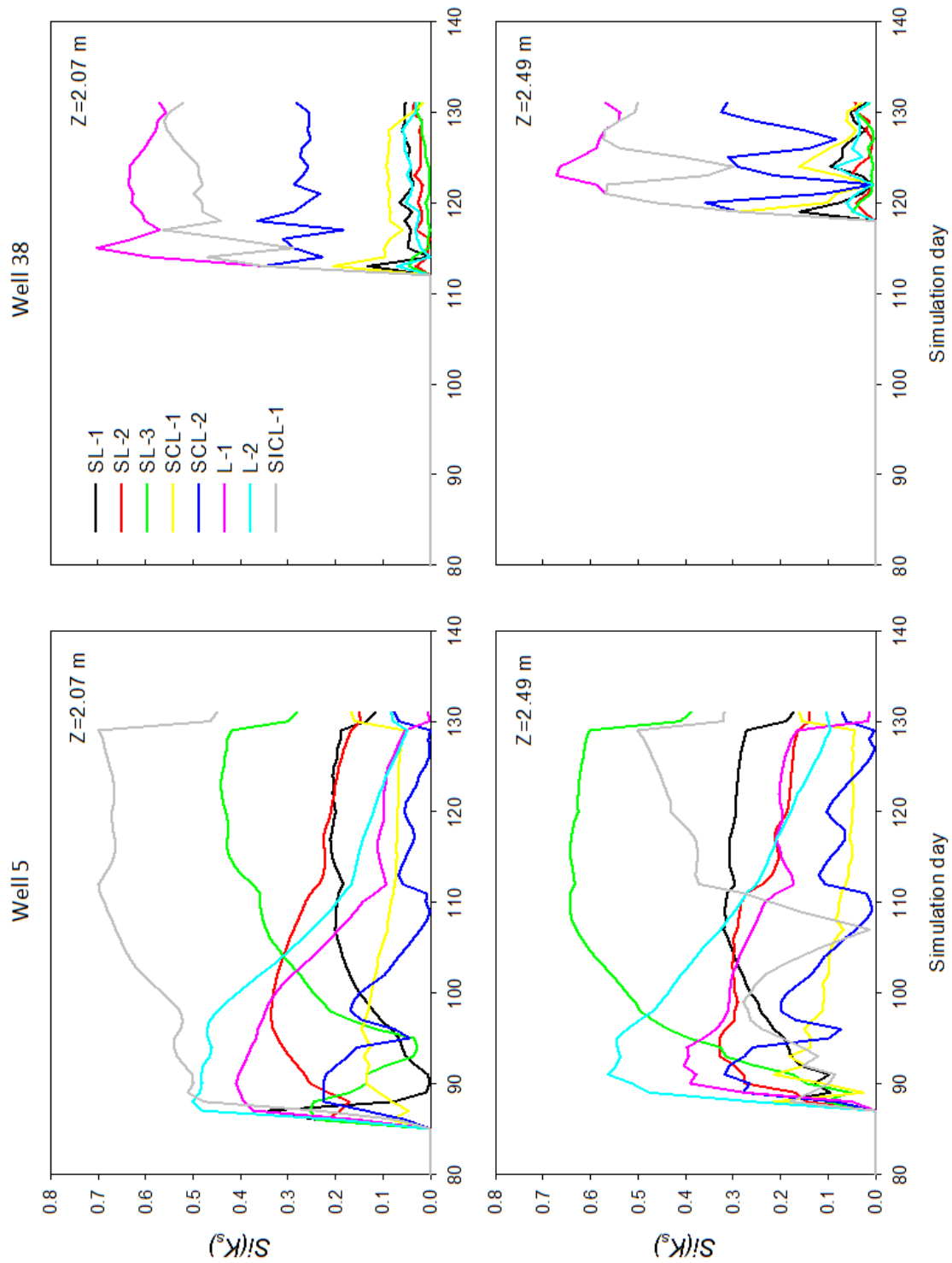
Figure 6-2   Examples of time series of the sensitivity index Si(Ks) for eight materials in well 5 and well 38 locations.
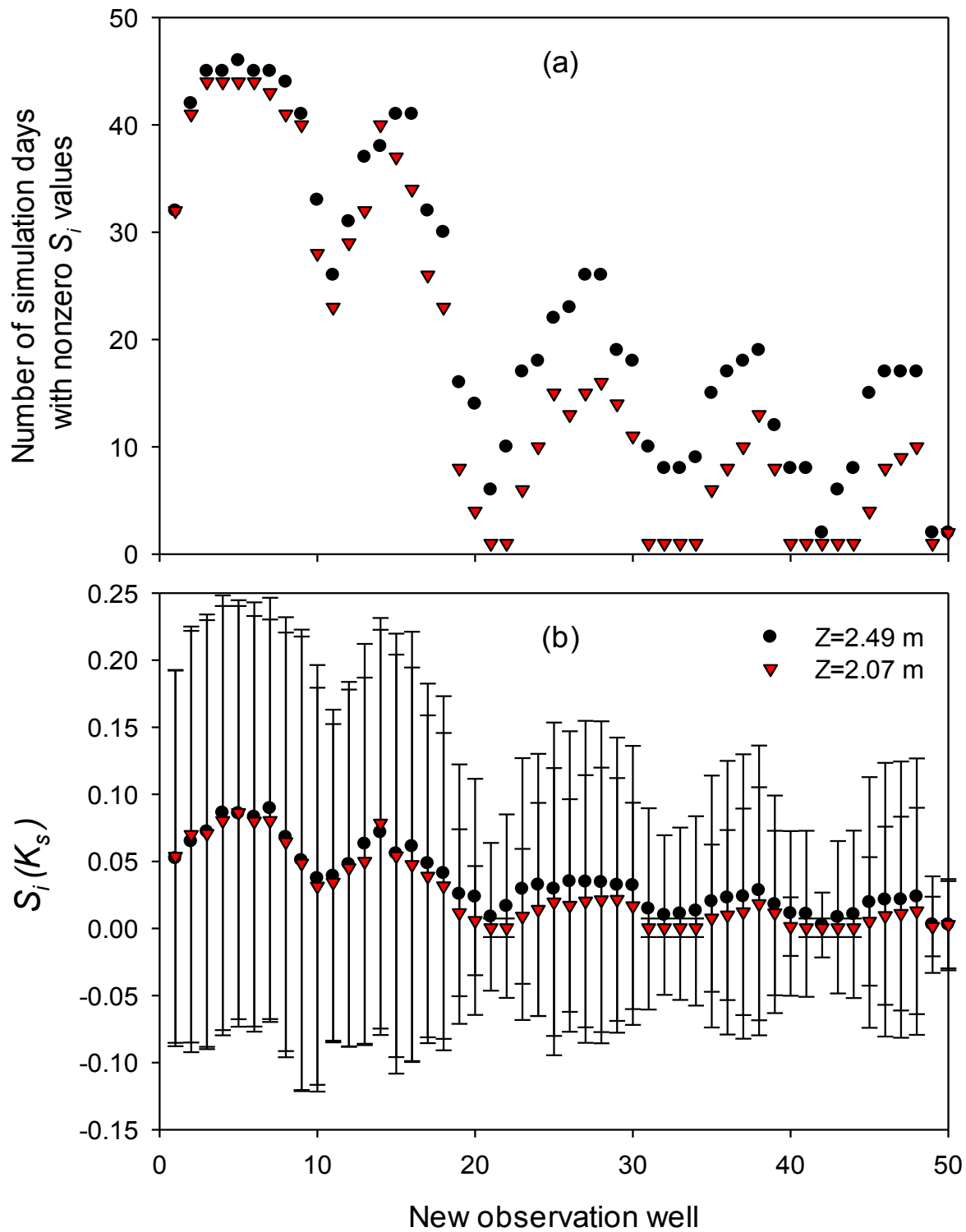
Figure 6-3 Number of days with nonzero Si values for each material (a) and Si variations in 50 observation wells.
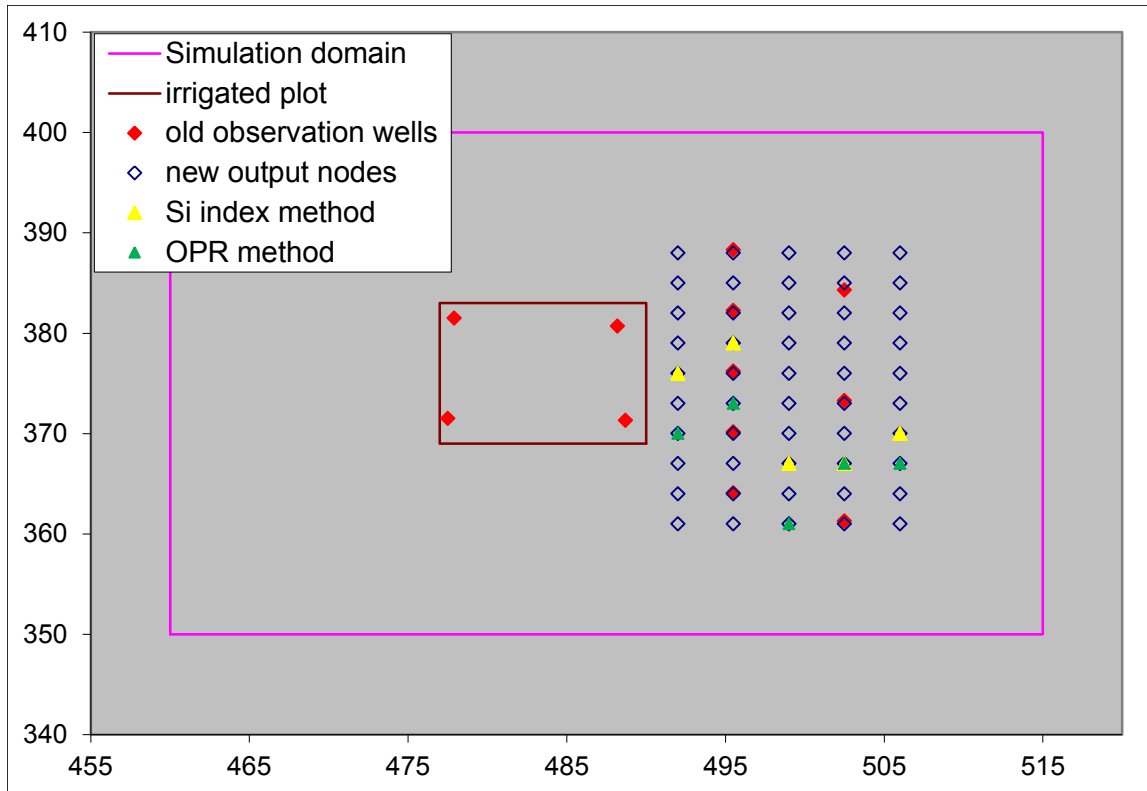
Figure 6-4   New observation locations selected based on maximum Si values (yellow triangles) and on averaged OPR statistics (green triangles).

The importance of the new observation locations for the HYDRUS-3D calibration was evaluated using the OPR$^+$ statistic. The S_i values computed for two depths at 50 locations shown in Fig. 6-1 were averaged over time periods showing nonzero values and used to calculate the matrix of sensitivities $\underline{A}_Y$ in Eq. 6-6. The matrix of weights $\underline{\omega}$ on the observations was represented by the proportion of simulation days with nonzero S_i values over the total simulation period (131 days). The standard deviations computed for locations at wells 5 through 12 for the OPR$^-$ statistics were used for the denominator in Eq. 6-5. The numerators $S_{Y'_{l(+i)}}$ in Eq. 6-5 were computed by adding one output node at a time to the existing nodes. The OPR$^+$ statistics computed for the added nodes for two computational layers at 50 locations ranged from 0% to 46.6% and from 0% to 54.8% for the 2.49 m and 2.07 m computational layers, respectively (Fig. 6-1). The OPR$^+$ values averaged across 8 existing observation wells did not change significantly with distance from the irrigated plot. However, they were noticeably smaller for some wells at z=2.49 m than at z=2.07 m, indicating that data collected from the z=2.49 m layer were less important.

To identify new locations for improving the HYDRUS-3D calibration, the average OPR$^+$ values were ranked in a descending order. The ranks of the new wells were different for the two computational layers (Fig. 6-7), with the percentage decrease in uncertainty generally being higher for the same rank for the z=2.07 m layer as compared to the 2.49 m layer. The locations of the new observations wells were therefor selected based on OPR$^+$ values computed for z=2.07 m. The new monitoring locations selected on the basis of high S_i values and using the OPR$^+$ statistics along with the existing monitoring locations are shown in Fig. 6-4. The locations obtained using both methods were very close to each other and indicated the preferential flow direction of the tracer. Nevertheless, the locations selected using the OPR$^+$ method were slightly northward from the locations having high S_i values in the first three transects (wells 1-10, wells 11-20 and wells 21-30) . These mismatches can be caused by differences in the approaches used in the two methods. The S_i-based method analyses values of the sensitivity separately at each observation location and indicates the locations with the highest sensitivity. These locations were between wells 6 and 7 in the existing monitoring network, where the highest concentrations were calculated (Fig. 5-2). The OPR$^+$ method examines the spatial distribution of S_i values weighed by the number of nonzero values and selects the locations were S_i values were not necessarily high, but were changing spatially. For instance, new observation location 16 between wells 7 (location 15) and 8 (location 17) in the existing monitoring network was in a transitional zone from relatively high concentrations in well 7 to low concentrations in well 8, indicating considerable uncertainty in the S_i values computed for location 17.
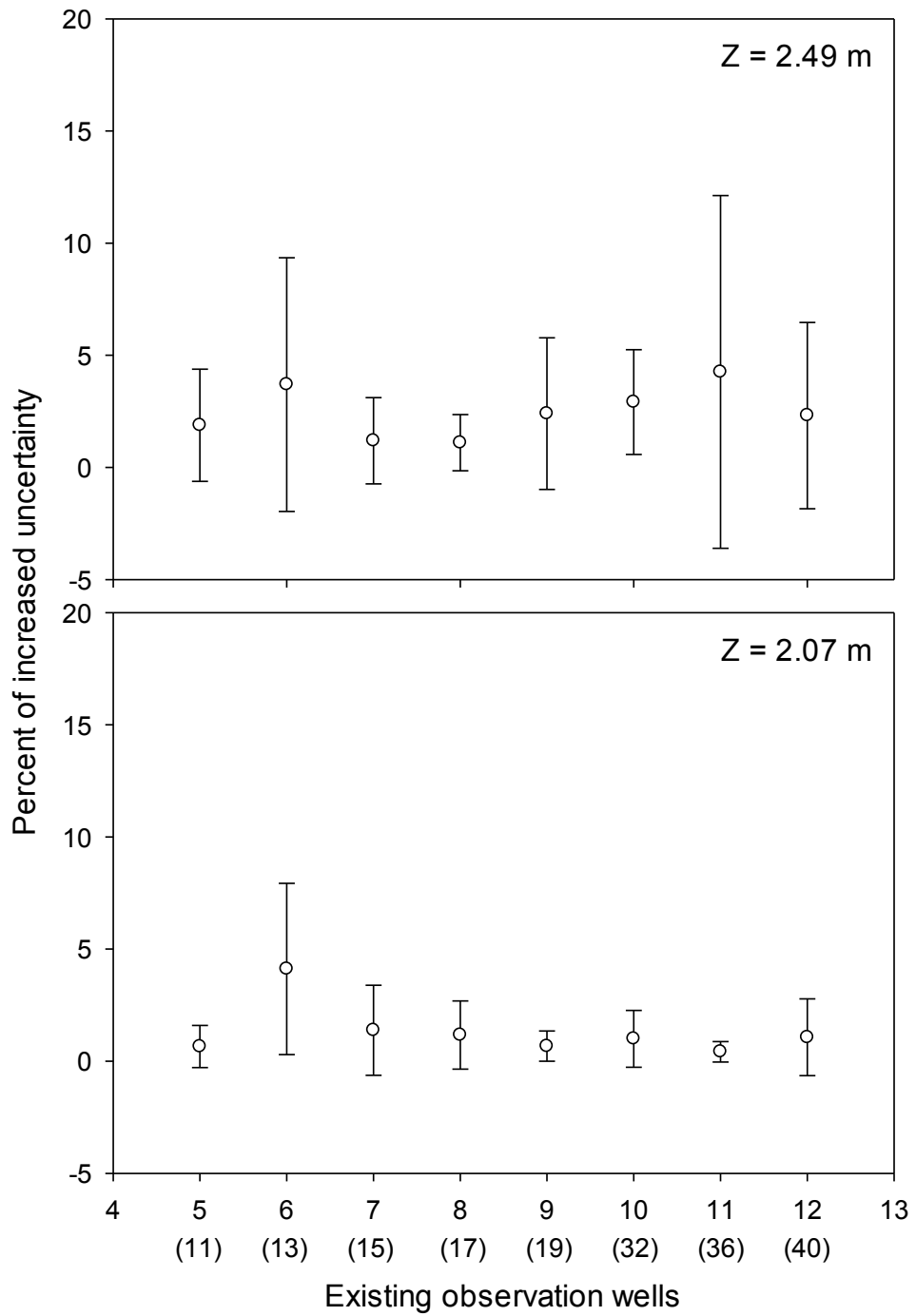
Figure 6-5   Percent of increased uncertainty caused by removal one existing observation well. New observation well numbers are shown in parenthesis on X-axis.
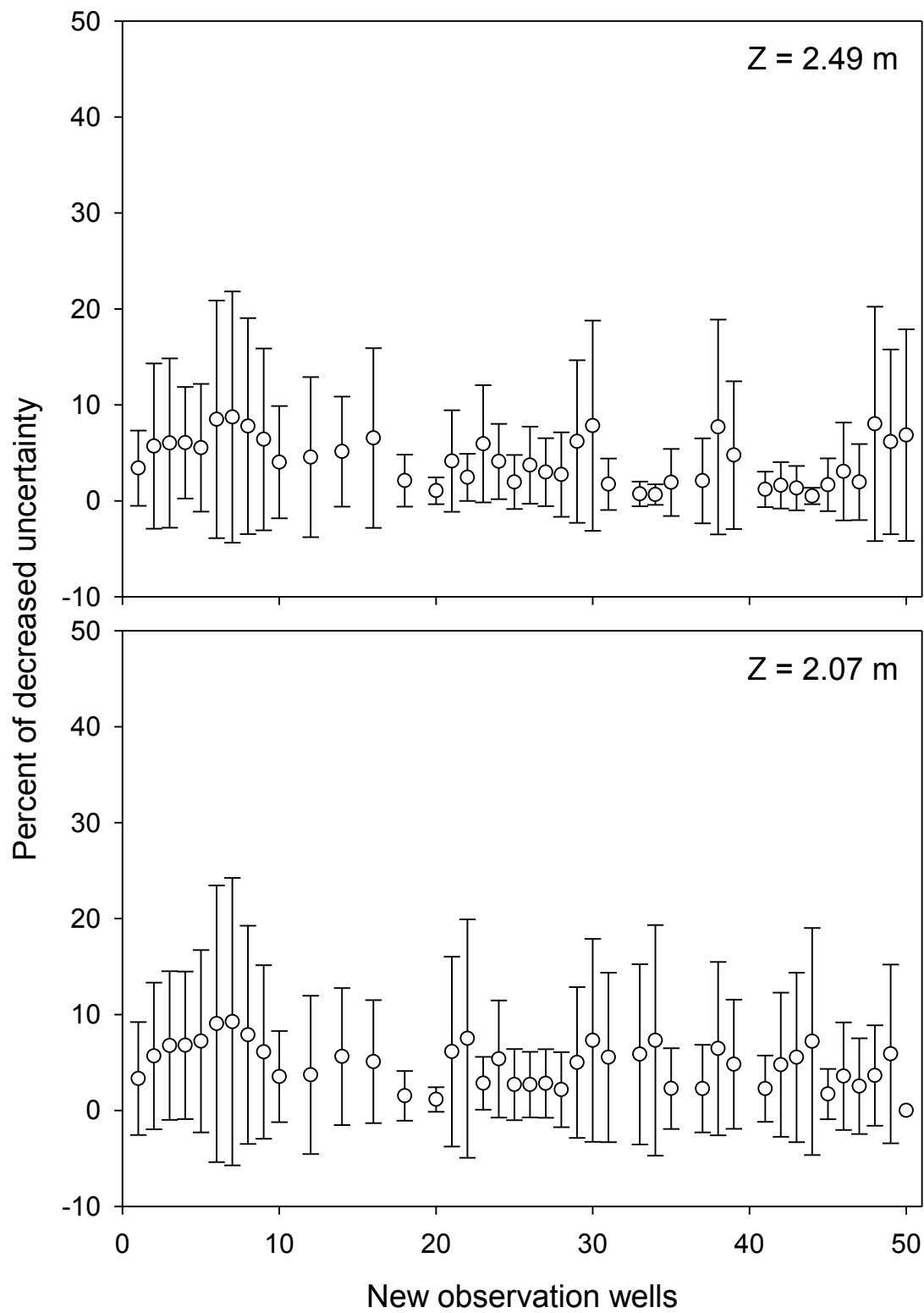
Figure 6-6   Percent of decreased uncertainty caused by adding one new observation well.
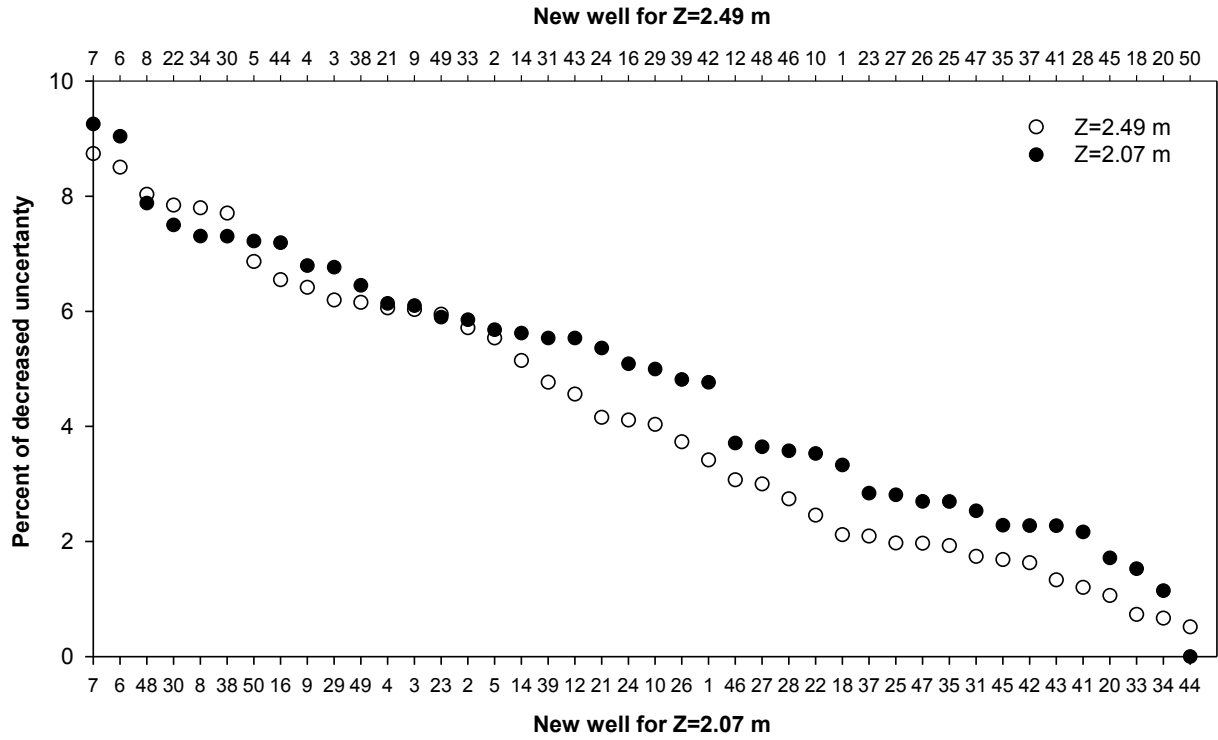
Figure 6-7   Ranks of new observation wells at two depths computed based on the percent of decreased uncertainty caused by adding one new well.

Overall, both statistical methods provided useful additional information for optimization of the monitoring network, specifically for model calibration purposes. The method using the observation-prediction (OPR) statistic accounts for both model sensitivity to the parameters and their spatial structure. Based on results of this study, the ARS study team recommend using the OPR[+] method for selecting monitoring locations to improve model calibration.

6.1.3 Selecting Monitoring Locations Based on Statistical Analysis of the Performance Indicators

To identify optimal locations for monitoring propagation of the contaminant plume from the source, the ARS study team used the results of 48 simulations for different weather, groundwater level and release scenarios as described earlier in section 5.2.3. Plume propagation was evaluated using again three performance indicators: C_{peak}, T_{peak} and QC computed for each of the 100 output nodes. Two approaches were used to select the monitoring locations. The first approach was based on the frequencies of appearance of extreme values of the performance indicators evaluated at different distances from the release points, while the second approach used the probability analysis.

To implement the first approach the minimum values of T_{peak} (T_{peak}^{min}), and maximum values of C_{peak} (C_{peak}^{max}) and QC (QC_{peak}^{max}) were computed for each transect passing through the output nodes located at the same distance from the release points. These values were computed using the results of HYDRUS-3D simulations with the calibrated parameters separately for each scenario. One location within each transect with the minimum value of T_{peak}^{min}, and maximum value of C_{peak}^{max} or QC_{peak}^{max} was chosen as a potential monitoring location for this transect. Locations were selected in the same manner for each performance indicator at all 11 transects and for all 48 scenarios. The ARS study team next analyzed the frequency of appearance of the selected locations at each transect using results of the 48 simulations. The frequency distributions for the C_{peak} performance indicator at 11 transects is shown in Fig. 6-8a. The distributions typically had a sharp peak at one of the locations. The maximum frequency values were 0.68±0.17, 0.78±0.13 and 0.65±0.21 for C_{peak}, T_{peak} and QC, respectively, which indicates high probabilities of detecting the contaminant plume using the selected locations.

Monitoring locations identified based on the frequency approach are mapped in Fig. 6-8b. Overall, the locations selected using the QC indicator were in good agreement with those selected using the C_{peak} performance indicator, while T_{peak}-based locations were shifted from 10 to 20 m to the north from the C_{peak}- and QC-based locations. The shift between the locations selected using different performance indicators increased with the tracer travel distance (Fig. 6-8b). The 20 m shift is significant for a 100 m travel distance, and can be even more pronounced for large-scale simulations.

To implement the second approach for selecting new monitoring locations, now based a probability analysis, the ARS study team analyzed the spatial distributions of performance indicators obtained with the 48 simulations using the calibrated HYDRUS-3D model. The ARS study team mapped for this purpose
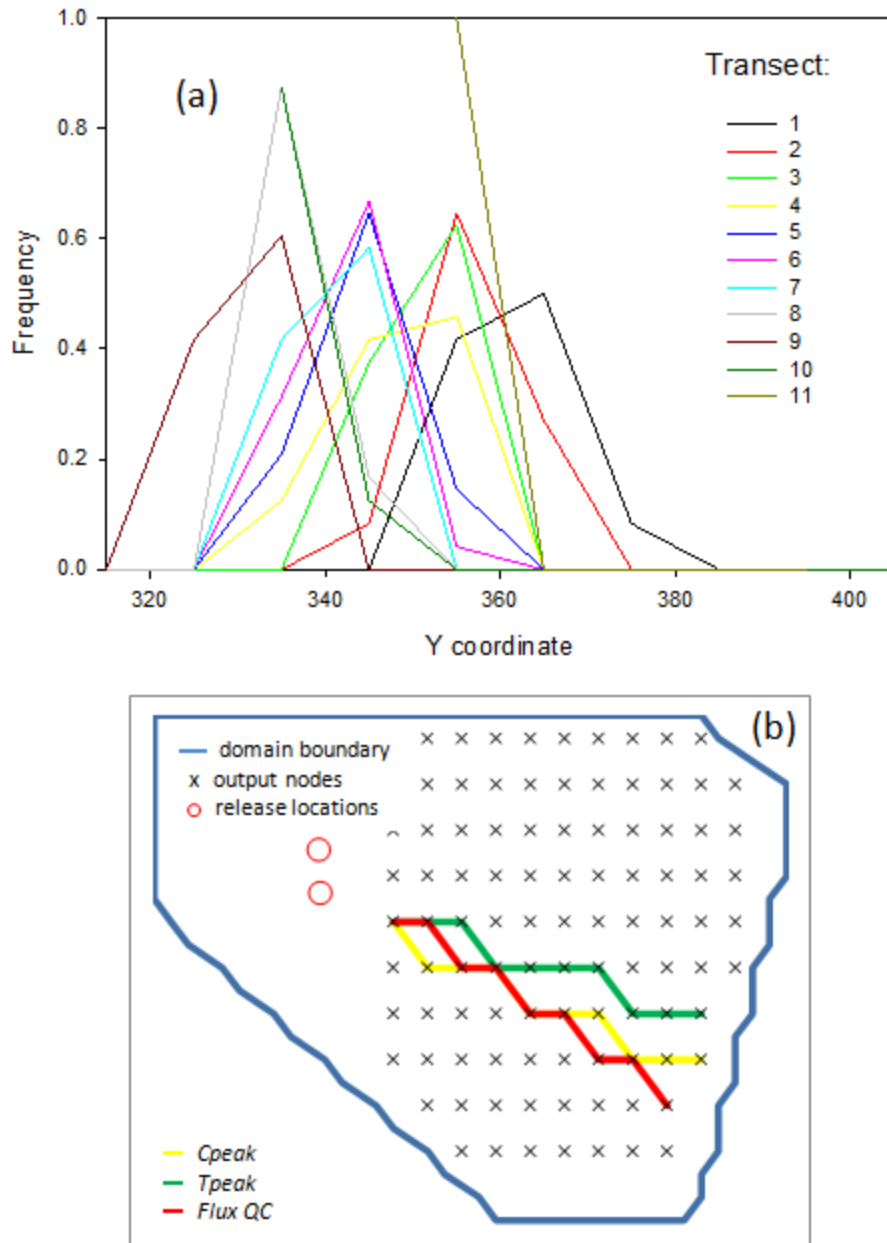
Figure 6-8 Frequency distributions for the C_{peak}^{max} in 11 transects (a), and monitoring locations selected based on the frequency distributions of 3 performance indicators (b).

mapped the values of 3 performance indicators at 25%, 50% and 75% probability level. These levels will specify for each output node the probability that the performance indicator obtained with the 48 simulation scenarios will not exceed the value of the performance indicator at that probability level. For instance, $C_{peak}^{75\%}$ means that the *Cpeak* values computed for 36 scenarios at this node will be smaller than $C_{peak}^{75\%}$, while *Cpeak* values will be greater than $C_{peak}^{75\%}$ value at this node only for 12 scenarios.

The spatial distributions of $\log_{10}(C_{peak})$, T_{peak} and $\log_{10}(QC)$ for three probability levels are shown in Fig. 6-9. The spatial variability generally decreased for C_{peak} and QC, and increased for the T_{peak} performance indicator with an increase in the probability level. For instance C_{peak} varied within 7, 5 and 4 orders of magnitude for the 25%, 59% and 75% probability levels, respectively (Fig. 6-9). High C_{peak} variability translated into considerably variability in the QC values, which ranged within 7.5, 6.5 and 5.5 orders of magnitude for the same probability levels. To identify potential locations for the monitoring network examined persistence of high PI values within each transect. The persistence was assessed as an interquartile range ($PI_{peak}^{75\%} - PI_{peak}^{25\%}$), while the values of the performance indicators were taken at 50% probability level.

Generally, an increase in the interquartile range with the $PI_{peak}^{50\%}$ increase was observed in this study, which indicates variation in the flow direction associated with the source location and groundwater scenarios. Therefore very few locations with PI values smaller than $PI_{peak}^{50\%}$ were selected, while in most cases locations with maximum $PI_{peak}^{50\%}$ were chosen for *Cpeak* and *QC* PI within each transect. The locations with the smallest travel times $T_{peak}^{5o\%}$ were selected for *Tpeak* performance indicator. The selected locations are shown in Fig. 6-10b. The selected locations are shown in Fig. 6-10b. The spatial position of the observation network was very similar to that obtained using the frequency approach (Fig. 6-10a), which confirms the reliability of the implemented approaches.

Overall, both the frequency and probability of peak performance indicators appear to be appropriate statistics to identify monitoring locations with the highest probability of contaminant plume detection. The observed differences between the time-based (T_{peak}) and concentration (C_{peak}) or flux-based (QC) performance indicators can be attributed to the different nature of these indicators. The arrival time is not always a good predictor of plume propagation in highly heterogeneous media due to combined effect of the contaminant dispersion and dilution, which may lead to very rapid transport of relatively small amounts of contaminant. As such, the C_{peak} and QC performance indicators appear to be more reliable for identifying monitoring locations.

6.1.4 Model Abstraction Applicability for Selection of Monitoring Locations

The applicability of the model abstraction techniques described in sections 5.2.5-5.2.7 is tested in this section for selection of monitoring locations. To do this the ARS study team examined the spatial distributions of the performance indicators obtained with HYDRUS-3D using the abstracted parameters (section 5.2.5), using profile aggregation (section 5.2.6) and using the abstraction of ignoring the unsaturated zone (section 5.2.7). Similar to section 6.2.1, values of the performance indicators at probability levels of 25%, 50% and 75% were plotted as contour maps. Spatial distributions obtained for the three model abstractions were found to be different. Patterns of the C_{peak} and QC indicators obtained with parameter abstraction (Fig. 6-11) in general resembled the patterns of the calibrated HYDRUS-3D model. However, the absolute values of the performance

Figure 6-9 Spatial distributions of the performance indicators at 3 probability levels obtained in HYDRUS-3D simulations with the calibrated parameters.
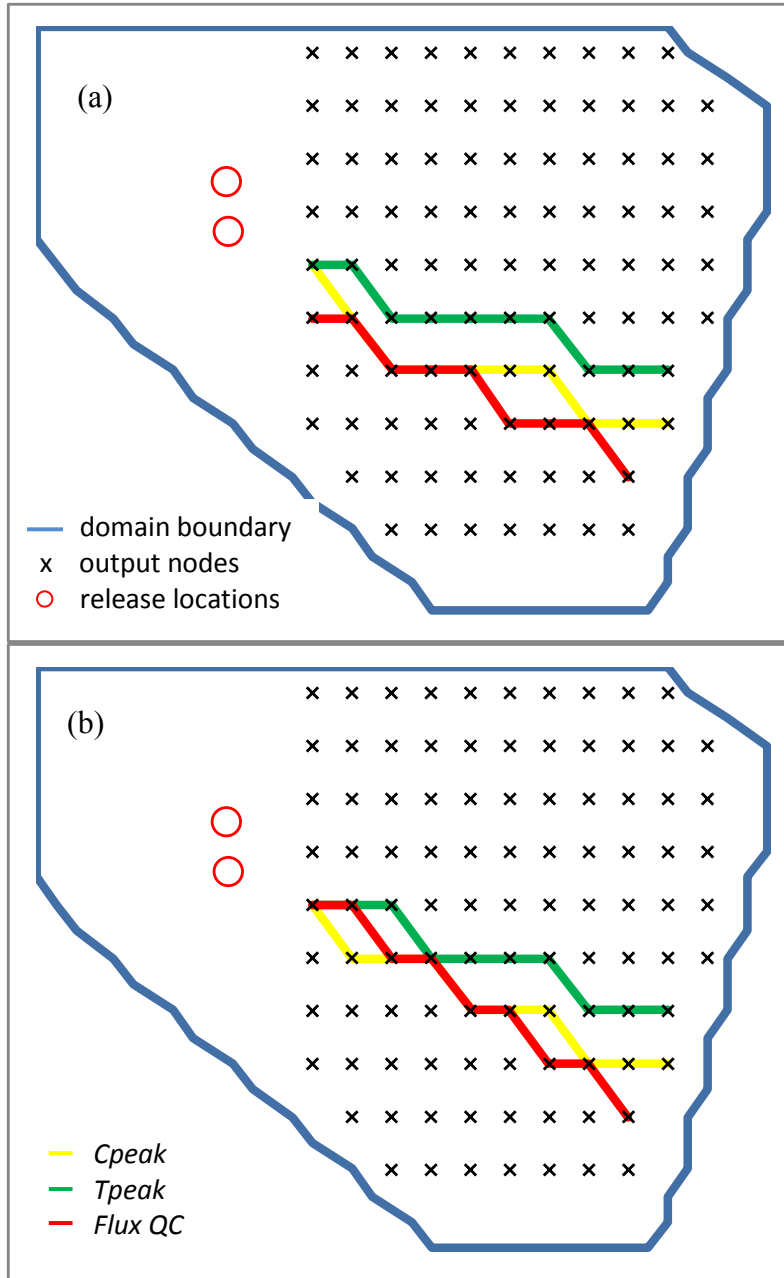
Figure 6-10 Monitoring locations selected using the frequency (a) and probability (b) approaches for 3 performance indicators.

indicators were somewhat different. Specifically the abstracted model systematically underestimated peak concentrations and overestimated total fluxes (Fig. 6-9 and Fig. 6-11). Little or no similarities were found between the travel times of the calibrated and abstracted models.

Abstraction to a homogeneous soil profile produced even greater differences with the original model. The shapes of the plume and absolute values of the performance indicators for the homogeneous profile model significantly deviated from the calibrated heterogeneous model (Fig. 6-8). The flow direction, viewed as a line with the smallest gradient in the performance indicator, was shifted towards south in the abstracted model compared to the original one.

No differences in the spatial patterns of the performance indicators between the original and abstracted model were found when the abstraction of ignoring the unsaturation zone was implemented. The spatial distributions of the three performance indicators were reproduced by the abstracted model with high accuracy at the 25%, 50% and 75% probability levels (Fig. 6-13). This result was expected since this abstraction adequately reproduced predicted concentrations and fluxes as was shown in section 5.2.7. To test the applicability of the three abstraction techniques to selecting monitoring locations, both the frequency and probability approaches were implemented to the spatial distributions of performance indicators. Peak values of the performance indicators were used for the frequency approach, while the peak and inter quartile range values were used both in the probability approach as described in section 6.2.1.

Monitoring locations for the two abstractions, along with the original model, are shown in Fig. 6-14. The parameter abstraction generated similar monitoring networks for the C_{peak} and QC performance indicators, while locations identified using T_{peak} values were shifted 10 to 20 m to the north (Fig. 6-14b). This shift was not systematic, which indicates a weakness of the T_{peak} indicator. Differences between the two approaches for location selection were minor for the C_{peak} and QC indicators, but noticeable for T_{peak}. This could be attributed to the different sensitivity of the different approaches to the shape of the performance indicator distributions. In spite of the observed differences between the absolute values of C_{peak} and QC, the observation locations for the abstracted model were very close to those obtained with the calibrated model. For the frequency approach, 6 and 7 out of 10 locations for the C_{peak} and QC indicators, respectively, were located 10 m to the north in the abstracted model as compared to the original model (Fig. 6-14 a,b). These differences were even smaller for the probability approach. Only 3 and 4 out of 10 locations for the abstraction model did not match the locations of the original model (Fig. 6-14 d,e). The difference in all cases was 10 m, which is comparable with the output grid resolution. Interestingly, the locations selected on the basis of T_{peak} values for the abstracted model were not systematically shifted from those of the calibrated model. The monitoring locations for the abstracted model oscillated around the network obtained for the calibrated model with a 10 m amplitude. Overall, in spite of inaccuracies in the parameter abstracted model for predicting contaminant concentrations and fluxes, this abstraction approach appeared to be appropriate for selecting monitoring location in our study.

The abstraction by ignoring the unsaturated zone appeared to be the most accurate in predicting the tracer concentrations and fluxes, leading to essentially the same monitoring network as the calibrated model. Monitoring locations obtained using both the frequency and probability approaches for the abstracted model matches for 100% the locations generated by the original model (Fig. 6-14 c,f). As discussed in section 6.2.7, this was possible since the unsaturated zone was very thin in this study and since spatial heterogeneity in the soil properties caused vertical transport to dominate at the release points. These hydrologic conditions created a specific flow regime in which the tracer was transported preferentially in the vertical direction before reaching
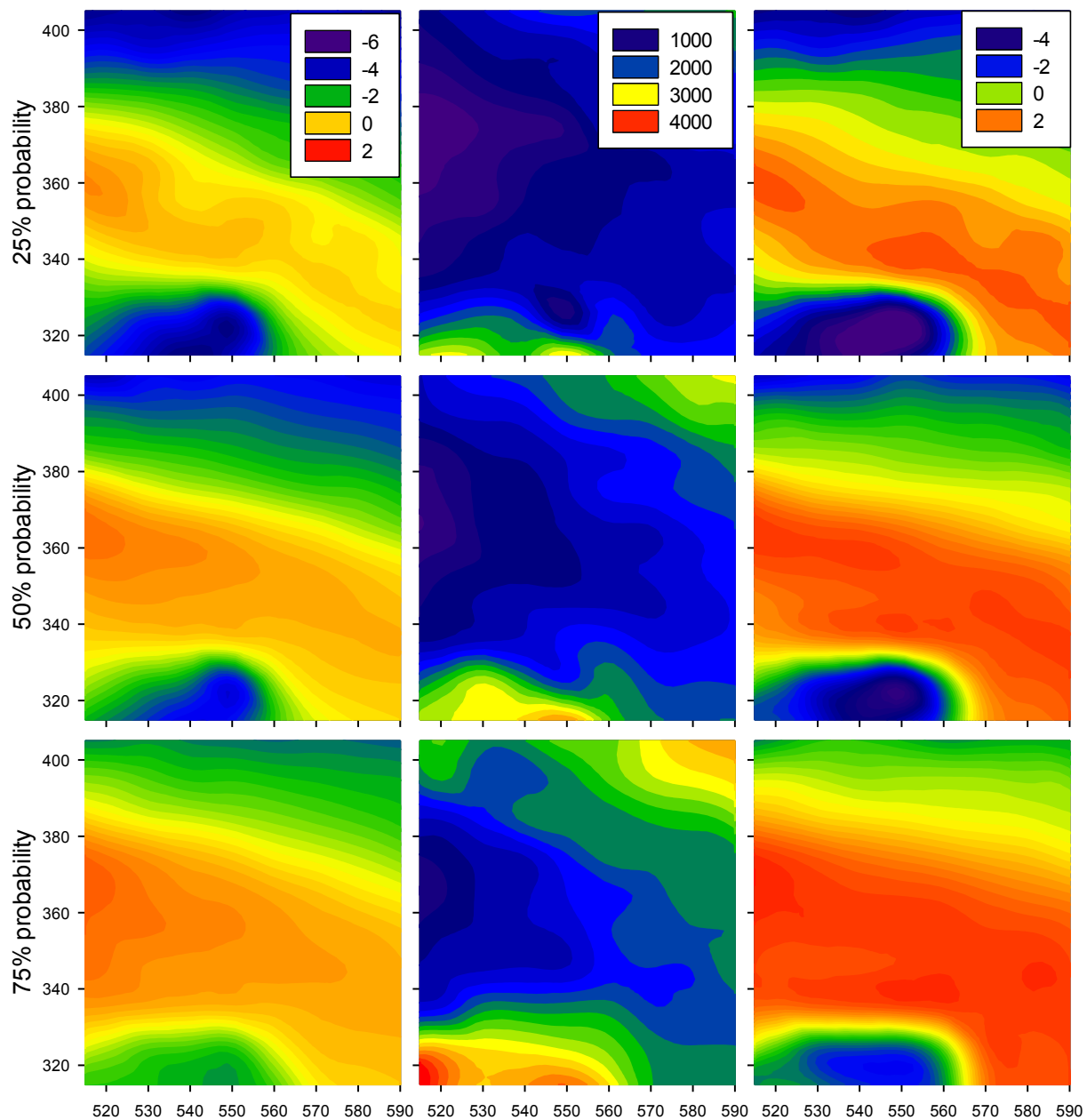
Figure 6-11   Spatial distributions of the performance indicators at 3 probability levels obtained in HYDRUS-3D simulations with the abstracted parameters.
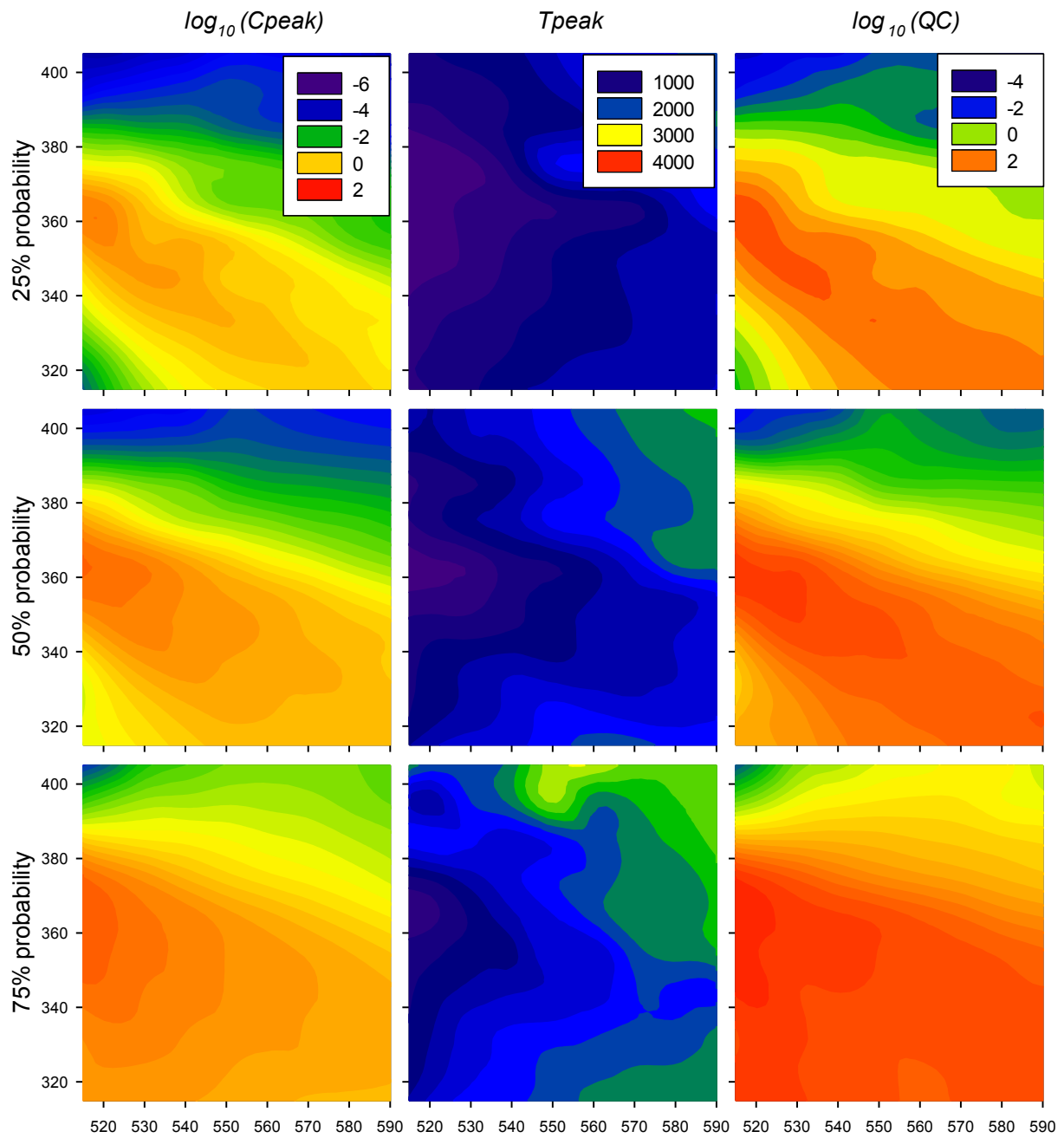
Figure 6-12 Spatial distributions of the performance indicators at 3 probability levels obtained in HYDRUS-3D simulations with the abstracted soil profile.
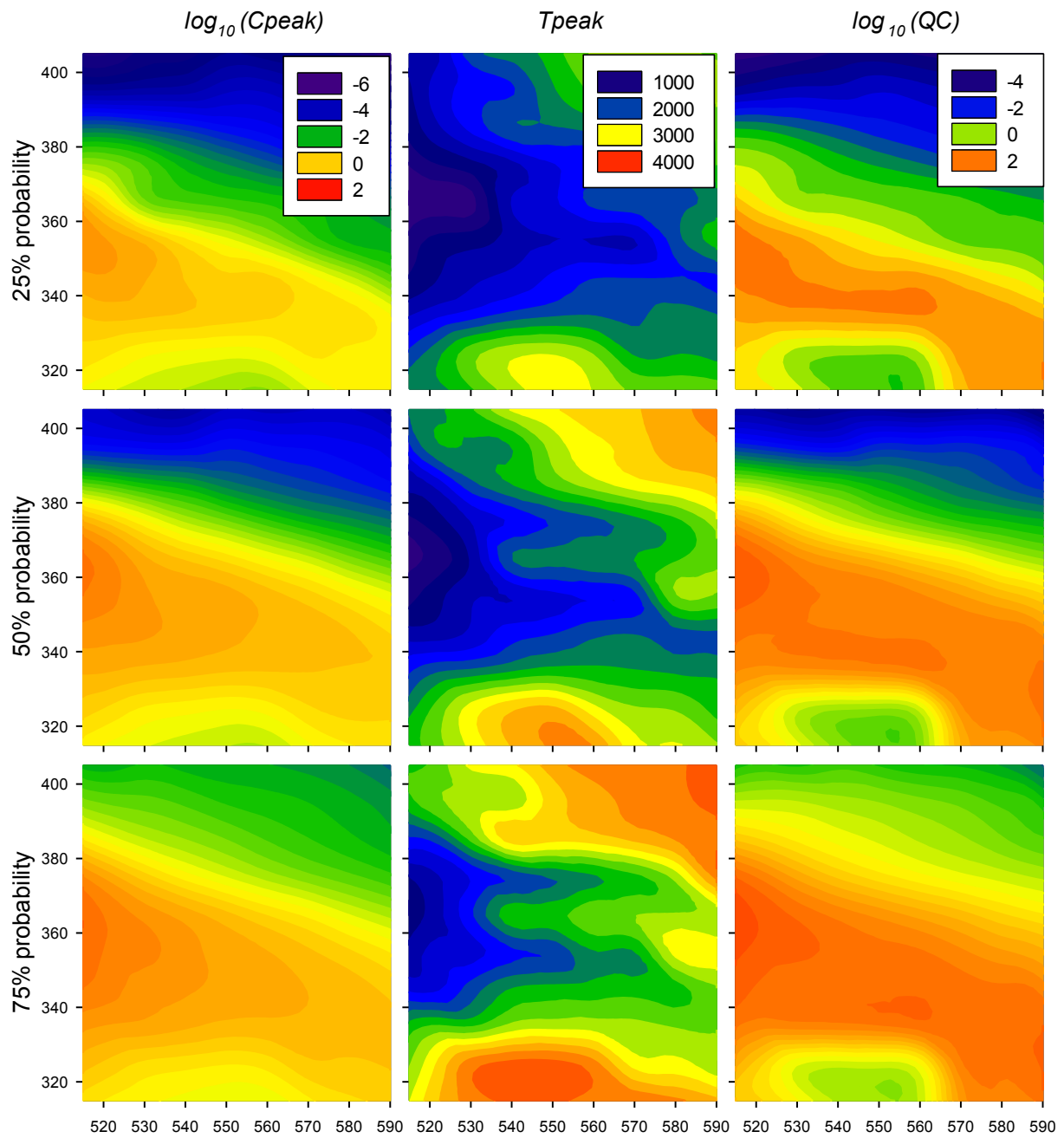
Figure 6-13 Spatial distributions of the performance indicators at 3 probability levels obtained in HYDRUS-3D simulations with the discarded unsaturated zone.
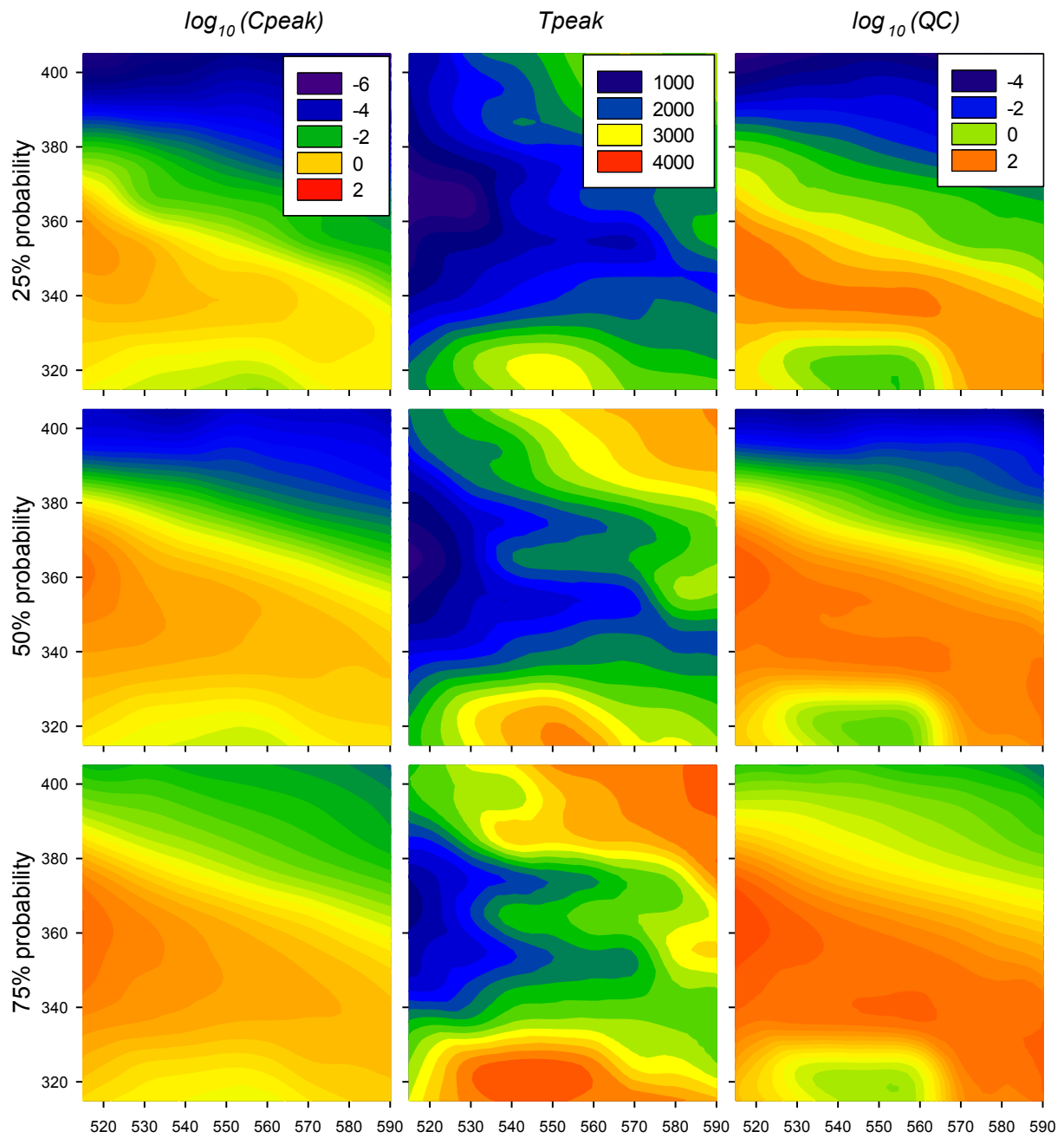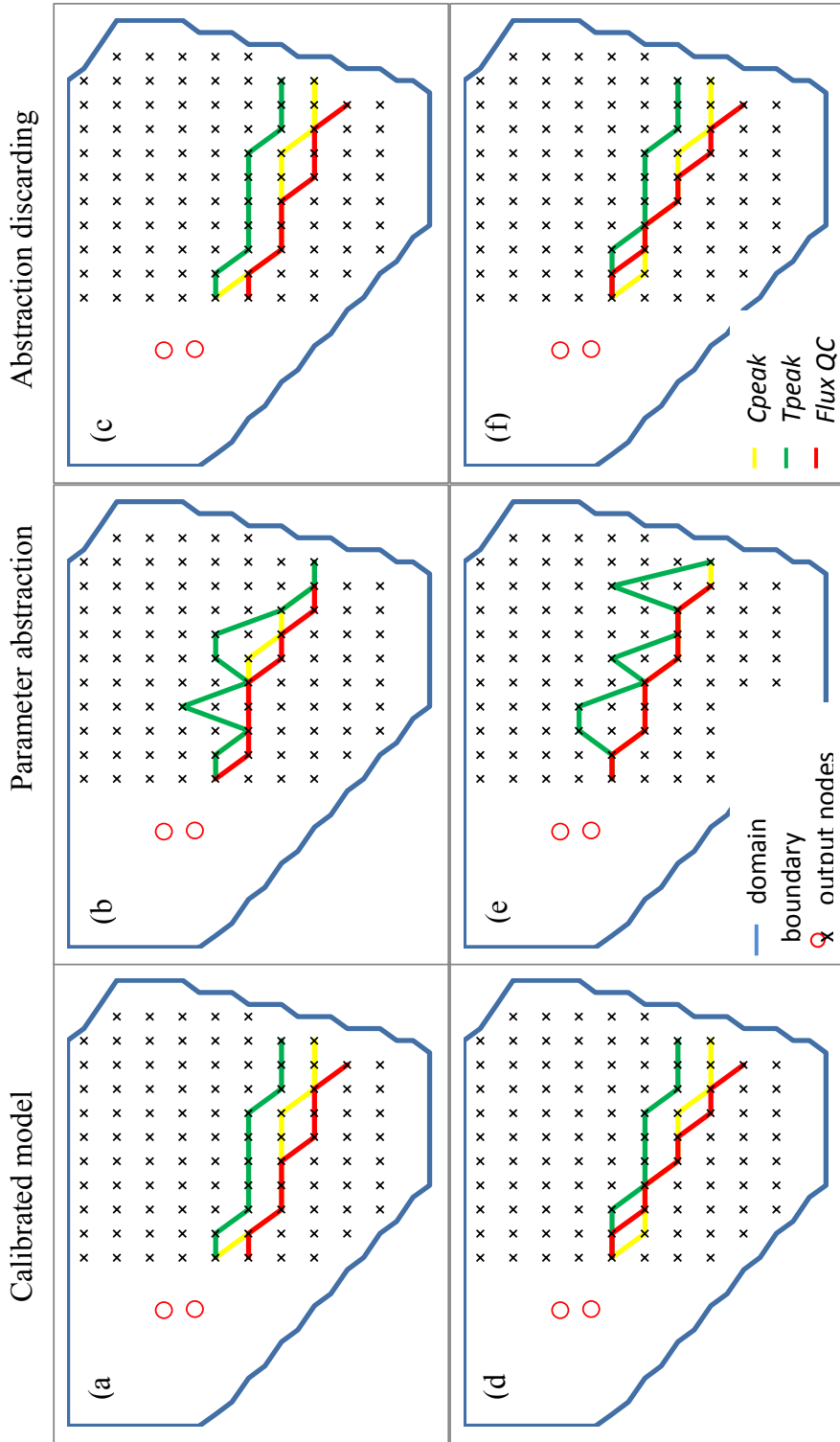
Figure 6-14 Monitoring locations selected for the calibrated model and two abstractions using the frequency (a, b, c) and probability (d, e, f) approaches.

groundwater. The results of this abstraction may differ when a thicker vadose zone is present, or when horizontal lenses may slow down vertical transport.

Using a homogeneous soil profile was the only abstraction technique that generated a monitoring network dissimilar to the network obtained with the calibrated HYDRUS model. The selected monitoring locations were systematically shifted to the south from those of the calibrated model, with the difference increasing with the distance from the source of release (data not further shown). This occurred primarily due to existence of preferential lateral flow within the heterogeneous simulation domain. These preferential flow path were completely discarded in the homogeneous profile abstraction process. Homogeneous soil profile abstraction generally may be more appropriate for small scale applications or for situations where vertical soil stratification is not important. The natural succession of soil genetic horizons of large scale field projects may produce different hydrological regimes at different depths that cannot be ignored when water and chemical transport is modeled. The results of this study provide a very remarkable illustration of this general concept.

6.2 Selecting Monitoring Locations for Model Discrimination

6.2.1 Kullback-Himmelblau Sequential Design Method

Coupling of modeling and monitoring can be conducted for various purposes. One of the purposes can be selection of a better model from a set of models originating, for instance, from different conceptualizations of subsurface, from different levels of model complexity, from different lists of processes included, etc. The model comparison and discrimination can be improved if additional observations will be included. A new observaPtion points can be selected at the location where the maximum difference between model predictions is achieved.

A fruitful approach to evaluate the usefulness of a measurement for distinguishing between two hypotheses was proposed in the seminal paper (Kullback and Leibler, 1951) in which the information in a measurement for discrimination between two hypotheses was first defined. The mean value of this Kullback-Leibler information represented the information gain that could be encountered if the true hypothesis were accepted rather than the wrong one. This mean value eventually was termed Kullback-Leibler divergence, information gain, relative entropy, or information divergence, and was computed for model predictions an estimate of the information loss when full truth is approximated by the model (Poeter and Anderson, 2005). This measure was proven to be useful in (a) evaluation of predictive capabilities of hydrological models when observations presented the 'true' distribution that was approximated by model predictions (Weijs et al., 2010) , (b) improving inverse solutions of groundwater flow models (Szucs et al, 2006), and (c) assessment of improvement in modeling results with data assimilation (Bulygina and Gupta, 2009). Various approximations of the Kullback-Leibler divergence resulted in development of the family of model discrimination criteria, such as Akaike criterion AIC, and later AICc, (Burnham and Anderson, 2004) that were used in groundwater modeling to rank calibrated models (Poeter and Anderson, 2005; Foglia et al., 2007; Ye et al., 2008;), and in inverse groundwater modeling for hydraulic conductivity estimation using Bayesian model averaging (Tsai and Li, 2008). All applications of the Kullback-Leibler divergence to model discrimination, however, relied on the existing set of observations and did not attempt to seek additional measurements.

Kullback (1959) showed that Kullback-Leibler information could be also applied for the selection of a new observation location to better discriminate between models without making an assumption that one of models generates the 'true' distribution whereas another one does not.

Himmelblau (1970) implemented this suggestion for the case of non-linear models. The method is as follows:

Let Y be a random variable that is distributed with a probability density $p_1(y)$ when hypothesis H₁ is true (Model 1) and distributed with a probability density $p_2(y)$ when hypothesis H₂ (Model 2) is true. Then in some sense the quantity $ln[p_1(y)/p_2(y)]$ can be said to be a measure of the odds in favor of choosing H₁ over H₂ or, from the information theory viewpoint, of the information in favor of hypothesis H₁ as opposed to hypothesis H₂. The expected information in favor of choosing either H₁ or H₂ respectively is given by Kullback-Leibler divergences

$$I(1:2) = \int_{-\infty}^{\infty} p_1(y)\, ln\frac{p_1(y)}{p_2(y)}\, dy \tag{6-7}$$

$$I(2:1) = \int_{-\infty}^{\infty} p_2(y)\, ln\frac{p_2(y)}{p_1(y)}\, dy \tag{6-8}$$

Kullback (1959) proposed that the total information gain due to selecting one model instead of another, i.e. value

$$J(1,2) = I(1:2) + I(2:1) = \int_{-\infty}^{\infty} [p_1(y) - p_2(y)]\, ln\frac{p_1(y)}{p_2(y)}\, dy \tag{6-9}$$

be maximized to distinguish between two states of nature. It means that the new monitoring point has to be selected where the $J(1,2)$ value reaches maximum.

From the linearized analysis, one can assume that the $(n+1)^{st}$ observation is normally distributed about the expected value for the model, $\mathcal{E}\left\{Y_r^{(n+1)}\right\} = y_r^{(n+1)}$, with a variance of σ_Y^2 (Himmelblau, 1970). Furthermore, $y_r^{(n+1)}$ is distributed in a local (linearized) region about a predicted value, $\hat{Y}_r^{(n+1)}$, with a variance of σ_r^2. Consequently, $Y^{(n+1)}$ is distributed about $\hat{Y}_r^{(n+1)}$ with a variance of $\sigma_Y^2 + \sigma_r^2$. Thus the probability density of $Y^{(n+1)}$ for the rth model is

$$p_r\left(y^{(n+1)}\right) = \frac{1}{\sigma_Y^2 + \sigma_r^2} exp\left[-\frac{1}{2}\frac{\left(Y^{(n+1)} - \hat{Y}_r^{(n+1)}\right)^2}{\sigma_Y^2 + \sigma_r^2}\right], \ r=1,\ 2 \tag{6-10}$$

The quantities $\hat{Y}_r$ and σ_r^2 represent the mean and variance of Y for the rth model. Kullback (1959) showed that substituting (6) into (3) and (4), after completing the integrations yields

$$I(1:2) = \frac{1}{2}ln\frac{\sigma_Y^2 + \sigma_2^2}{\sigma_Y^2 + \sigma_1^2} + \frac{1}{2}ln\frac{\sigma_Y^2 + \sigma_1^2}{\sigma_Y^2 + \sigma_2^2} - \frac{1}{2} + \frac{1}{2}\frac{\left(\hat{Y}_1^{(n+1)} - \hat{Y}_2^{(n+1)}\right)^2}{\sigma_Y^2 + \sigma_2^2} \tag{6-11}$$

$$I(2:1) = \frac{1}{2}ln\frac{\sigma_Y^2 + \sigma_1^2}{\sigma_Y^2 + \sigma_2^2} + \frac{1}{2}ln\frac{\sigma_Y^2 + \sigma_1^2}{\sigma_Y^2 + \sigma_2^2} - \frac{1}{2} + \frac{1}{2}\frac{\left(\hat{Y}_1^{(n+1)} - \hat{Y}_2^{(n+1)}\right)^2}{\sigma_Y^2 + \sigma_1^2} \tag{6-12}$$

$$J(1,2) = \frac{1}{2}(\sigma_1^2 - \sigma_2^2)\left(\frac{1}{\sigma_Y^2 + \sigma_2^2} - \frac{1}{\sigma_Y^2 + \sigma_1^2}\right)\left(\hat{Y}_r^{(n+1)} - \hat{Y}_s^{(n+1)}\right)^2 \tag{6-13}$$

Box and Hill (1967) described an improved version of Kullback (1959) discriminant function in which the prior probabilities are included. To make full use of prior information available about the validity of each model, in addition to computing the estimation coefficients used in $\hat{Y}_r^{(n+1)}$, it would seem reasonable to weight $I(1:2)$ and $I(2:1)$ by the respective prior probabilities of Model 1 or Model 2 being the correct model, $P_1^{(n)}$ and $P_2^{(n)}$, respectively.

Himmelblau (1970) extended the previous approach by assuming that several competing models exist, among which the (n+1)st experiment is to discriminate. A scalar discriminant function can be formed using the matrix of relative likelihoods and a vector of prior probabilities as follows:

$$K_v = \begin{bmatrix} P_1^{(n)} & \cdots & P_v^{(n)} \end{bmatrix} \cdot \begin{bmatrix} I(1:1) & \cdots & I(1:v) \\ \vdots & \ddots & \vdots \\ I(v:1) & \cdots & I(v:v) \end{bmatrix} \begin{bmatrix} P_1^{(n)} \\ \vdots \\ P_v^{(n)} \end{bmatrix} \tag{6-14}$$

Each element on the main diagonal of the $I(r:s)$ a matrix is zero so the discriminant function becomes:

$$K_v = \frac{1}{2}\sum_{r=1}^{v}\sum_{s=r+1}^{v} P_r^{(n)} P_s^{(n)} \left[\frac{(\sigma_r^2 - \sigma_s^2)^2}{(\sigma_Y^2 + \sigma_r^2)(\sigma_Y^2 + \sigma_s^2)} + \left(\hat{Y}_r^{(n+1)} - \hat{Y}_s^{(n+1)} \right)^2 \left(\frac{1}{\sigma_Y^2 + \sigma_r^2} + \frac{1}{\sigma_Y^2 + \sigma_s^2} \right) \right] \tag{6-15}$$

One way to obtain the posterior probability that model *r* is correct after taking *n* observations is to apply successively for each model Bayes' theorem in the following form:

$$P_r^{(n)} = \frac{P_r(y(n))}{\sum_{r=1}^{v} P_r^{(n-1)} p_r(y(n))} \tag{6-16}$$

where $p_r^{(n-1)}$ is the prior probability associated with the *r*th model. The initial probabilities $p_r^{(0)}$ can be set equal to 1/*v* if not known.

In general, neither σ_Y^2 nor σ_r^2, *r*=1,…*v*, will be known; hence the values of these variances must be estimated. Himmelblau (1970) suggests obtaining the variance of $\hat{Y}_r$ for a model, by using the intermediate values (variances and co-variances at the optimal point) obtained in least square estimates of the models' parameters.

The sequential procedure to discriminate among *v* models is summarized as follows (Himmelblau, 1970):
1. Based on experimental design selected in some arbitrary or suboptimal way, collect *n* data points.
2. Estimate the parameters in the *v* models by linear or nonlinear regressions; estimate σ_Y^2 and calculate each σ_r^2.
3. Calculate the prior probabilities for the (*n*+1)st run which are equal to the posterior probabilities for the nth run by using equations (6) and (12) with *n* substituted for (*n*+1) in the latter. The initial *P*'s can all be equal to 1/*v* if no better choice is available.
4. Select the vector of experimental conditions for the (*n*+1)st run (the vector **x**$^{(n+1)}$ by maximizing K_v using a numerical optimization routine.
5. Run an experiment at **x**$^{(n+1)}$ and repeat starting with step 2.

The sequential procedure continues until one (or more) $P_r^{(n)}$ reaches a value acceptance of the model by some criterion. Or the experimenter can just observe the trend of the changes in the P_r as the number of experiments increases, drop models with low values of P_r, and add models, if he wishes, terminating the experiments when he feels satisfied with the discrimination actually achieved.

Limitations of applying the above procedures to subsurface flow and transport problems are that (1) the data are rarely sufficient to calibrate models, and (2) few if any numerical codes are available for calibrating 3D models if a variably-saturated subsurface is simulated (even if such codes exist, calibration procedure can consume enormous time). As mentioned above, the calibration is needed to estimate values of the means and variances for the simulated variables.

The latter problem can be overcome by using the assumptions of the ensemble Kalman filter (EnKF) methodology. The centerpiece of the EnKF methodology is the hypothesis that the covariance matrix of predictions can be obtained by running an ensemble of models and computing statistics, including variances and covariances on the predictions. Models in the ensemble have the same system of equations (as such it is a single model) but differ in the initial conditions, parameter values, and forcings, which may vary in feasible ranges with realistic probability distributions to form an ensemble. Using this postulate, one can run ensembles for each model to obtain the required values of $\hat{Y}_r$ and σ_r^2.

6.2.2 Applications of the Kullback-Himmelblau Method

In this section the ARS study team presents two examples to demonstrate using the Kullback-Himmelblau sequential design (KHSD) method for augmentation of a monitoring network for a contaminated aquifer based on data of (1) synthetic simulations, and (2) the OPE3 field tracer experiment carried out in 2008. It is assumed that n (≥1) observation locations are available where concentrations are measured to obtain breakthrough curves (BTCs) to be used for the model calibration. As independent variable, the ARS team selected coordinates ($\mathbf{x}_i$, $i=n+1,..,N$) of "imaginary" observation points that compose a grid covering some area downstream to the contaminant source location. As response function (Y), the ARS team selected the maximum concentration that can be reached at each observation point. The variance (σ_r^2) and mean ($\hat{Y}$) for the rth model at each point is calculated by running models using randomly varying parameters within a certain range and computing statistics. The value of the discriminant function (K_v, Eq. (6-17)) is subsequently calculated at each location. The maximum value of K_v specifies the location of the new observation point.

6.2.2.1 Example of Synthetic Simulations

Consider the aquifer to be composed of relatively homogeneous sediment characterized by a loamy sand texture. The simulation domain is presented by a rectangular parallelepiped having the size of 50x30x10 m in the x, y and z directions, respectively (Figure 6-15). Flow in the saturated zone is essentially one-dimensional in the x-direction. Groundwater is very shallow. Constant total hydraulic head values of 9.5 m and 9 m are prescribed along the boundaries at $x=0$ and $x=50$ m respectively; thus creating a hydraulic head gradient of 0.01 in the saturated zone. Hydraulic heads along the boundaries $y=0$ and $y=50$ m decrease linearly from 9.5 to 9.0 m with increasing x. The boundaries at $z=0$ and $z=50$ m were subjected to no flux conditions. Initially at $t=0$, the contaminant concentration in the aquifer is assumed to be zero.

A total of 20 L of contaminated water with a concentration C_0 of 100 mg L^{-1} was injected at the location of $x_0=5$, $y_0=15$ and $z_0=8$ m during 0.01 day. Two observation wells were installed at $W_{o1}(x=10m, y=13m)$ and $W_{o1}(x=10m, y=18m)$ where groundwater was sampled from a depth of 7.5 m. To create synthetic BTCs, the ARS study team performed simulations with the HYDRUS-3D model by imposing the above initial and boundary conditions, as well as the contaminant injection. The porous medium hydraulic parameters characterizing loamy sand were taken from the HYDRUS database as: $\theta_s=0.41$, $\theta_r=0.057$, $\alpha=12.4$ m-1, $n=2.28$, and $K_s=3.5$ m day^{-1}. The longitudinal (a_L) and transverse (a_T) dispersivities were assumed to be 0.75 and 0.15 m, respectively. The hydraulic parameters of most or all field problems exhibit spatial variability. To account for this the ARS study team generated a lognormally distributed hydraulic conductivity field using by using linear scaling transformations (Simunek, et al. 2011) with a scaling parameter of 0.25 (Fig. 6-16).

The finite element mesh consisted of 33,201 nodes and 60,000 3D elements with mesh sizes of 1 m in the *x*- and *y*-directions, and 0.5 m in the *z*-direction. Contaminant transport was simulated with HYDRUS-3D for a time period of 60 days (the initial observation period). Simulated concentration distribution is shown in Figure 6-17.

To apply the KHSD method for augmentation of the monitoring network, one needs at least two models. As our first model the ARS study team used the same HYDRUS-3D model; however, unlike the synthetic forward simulation, the random hydraulic conductivity field was not introduced and the saturated hydraulic conductivity was assumed to be constant. The second model was selected based on the assumption that flow in the saturated zone is essentially one dimensional and uniform. Solute transport for such conditions can be described with the linear advection-dispersion equation

$$
\begin{aligned}
R\frac{\partial C}{\partial t} = D_x\frac{\partial^2 C}{\partial x^2} + D_y\frac{\partial^2 C}{\partial y^2} + D_z\frac{\partial^2 C}{\partial z^2} - v\frac{\partial C}{\partial x} - R\lambda C \\
+ \frac{M}{\theta}\delta(x - x_0)\delta(y - y_0)\delta(z - z_0)\delta(t - t_0)
\end{aligned}
\tag{6-17}
$$

subject to initial condition

$$
C(x, y, z, 0) = 0
\tag{6-18}
$$

and boundary conditions

$$
C(\pm\infty, y, z, t) = C(x, \pm\infty, z, t) = C(x, y, \pm\infty, t) = 0
\tag{6-19}
$$

where *C* is contaminant concentration, *R* is retardation factor, D_x, D_y, D_z and are hydrodynamic dispersion coefficients in *x*, *y* and *z* directions, respectively, *v* is groundwater flow velocity, λ is the first order decay rate, *M* is the contaminant mass released instantaneously at point (x_0, y_0, z_0) at time t_0, θ is aquifer porosity and δ is the Dirac delta function.

The analytical solution for the problem (6-17) - (6-19) is (Lenda and Zuber, 1970):

$$
\begin{aligned}
C(x, y, z, t) = \frac{M}{\theta}\frac{1}{8\pi^{3/2}}\frac{1}{\sqrt{D_x D_y D_z/R}}\frac{1}{t\pi^{3/2}}exp[-\lambda(t - t_0)] \\
\cdot exp\left[-\frac{\{(x-x_0)-v(t-t_0)/R\}^2}{4D_x(t-t_0)/R}\right] \cdot exp\left[-\frac{(y-y_0)^2}{4D_y(t-t_0)/R}\right] \cdot exp\left[-\frac{(y-y_0)^2}{4D_y(t-t_0)/R}\right]
\end{aligned}
\tag{6-20}
$$
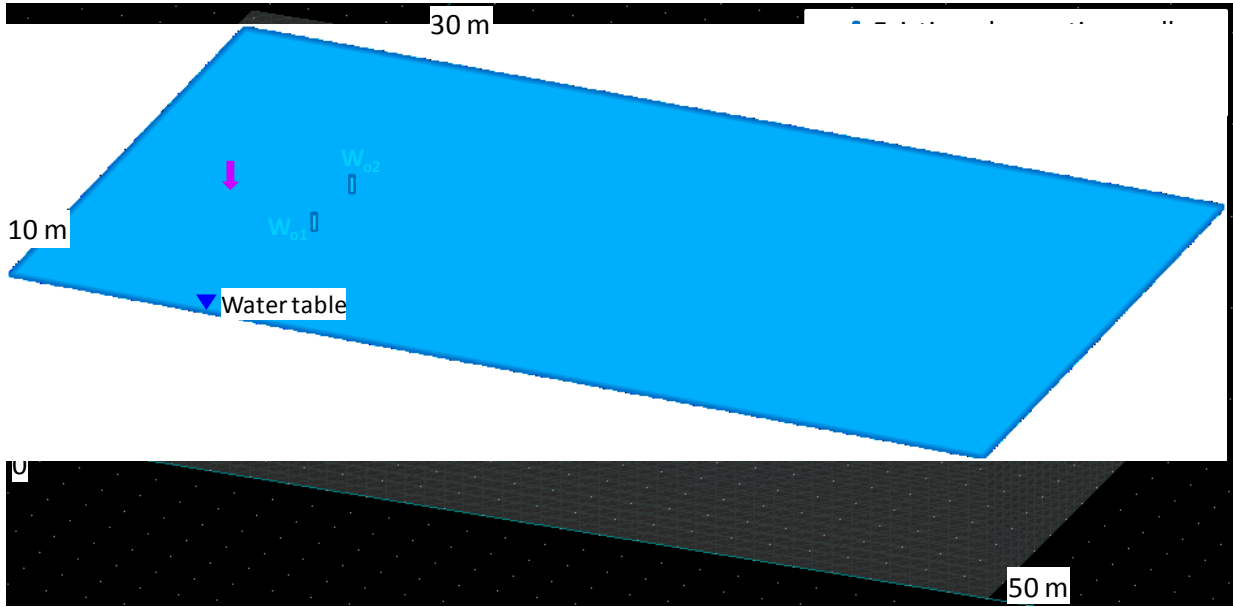
Figure 6-15 Simulation setup for synthetic example.
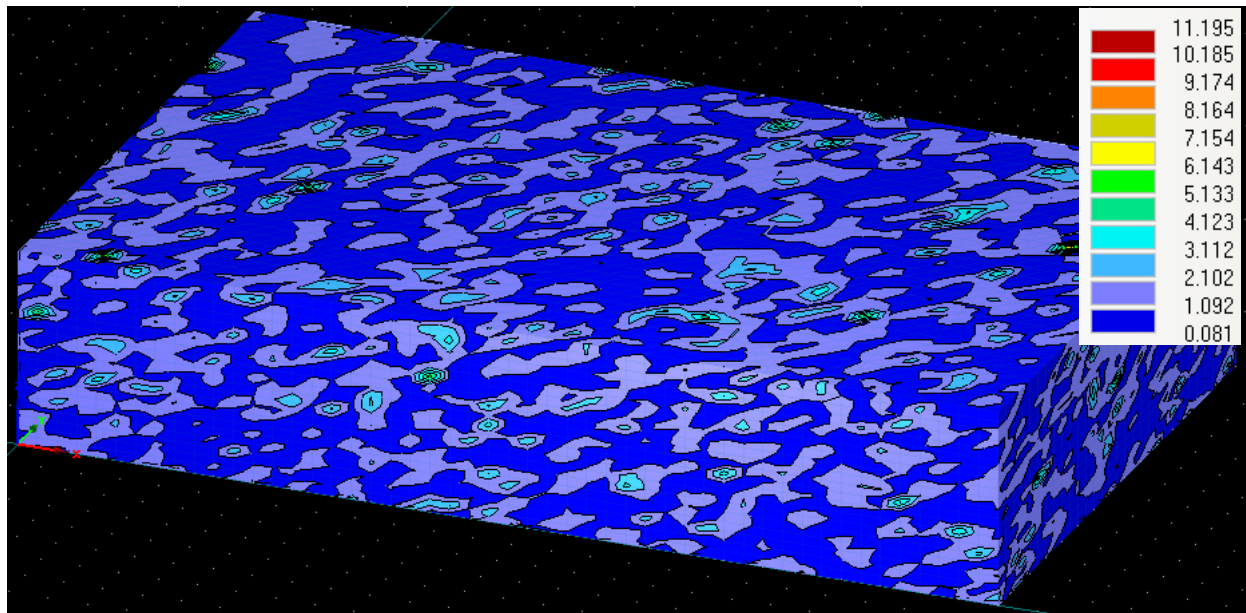
Figure 6-16 Distribution of scaling factor of saturated hydraulic conductivity in the simulation domain.
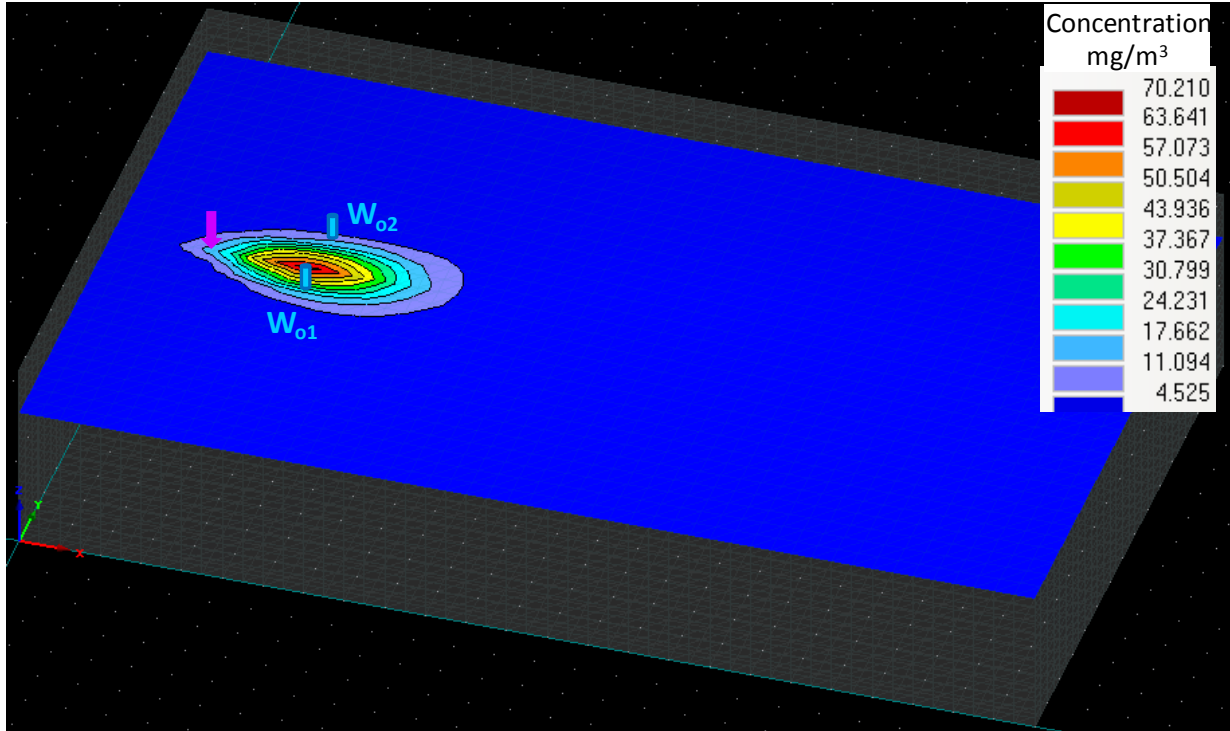
Figure 6-17 Simulated contaminant distribution after 60 days, at the layer of z=7.5 m.

Parameters of HYDRUS-3D and the above analytical model must be calibrated using the observed BTCs at the observation wells. A manual trial and error calibration procedure was used for HYDRUS-3D, while the Levenberg-Marquardt nonlinear optimization method (Marquardt, 1963) was used to calibrate the analytical model.

Parameters obtained for the HYDRUS 3D model were K_s=3.2 m day^{-1}, a_L=1.0 m and a_T=0.17 m. These values were not much different from those used to generate the BTCs at observation wells, mostly because relatively moderate values of the scaling parameter for log(K_s) was used. Results of the HYDRUS trial and error fit for BTCs at the two wells are shown in Figure 6-18a. Transport parameters found for analytical solution (6-20) were v=0.123 m/day, D_x=0.202 m^2 day^{-1}, D_y=0.015 m^2 day^{-1} and D_z=0.002 m^2 day^{-1}. Results obtained with the analytical model are shown in Figure 6-18b.

Ninety six imaginary observation wells at z=7.5 m, in the rectangular area bounded by (x,y) coordinates: (12,8) m and (34,22) m were chosen to calculate the discriminant function values, K_v (equation (6-16). The distance between nodes was 2 m in both x- and y-directions (Fig. 6-15). To estimate variance (σ_r^2) and mean ($\hat{Y}_r$) of the maximum concentrations in those nodes, the ARS study team performed 25 runs for each model using randomly varying parameters within a range 20% from the values found by the inversion. Simulations were executed for a period of 365 days.

To calculate the K_v values, the ARS study team used the variance and mean of the concentrations of the first observation well (W$_{o1}$), i.e. =8.4 and =16.1, respectively, as a starting point (the nth experiment). The initial values of prior probability were 0.5. Calculated distribution of the discriminant function presented in Figure 6-19 indicate a clear maximum at x=22 m and y=16 m, which is supposed to become the location of a new observation point.

Calculation of K_v using data from the second well (W$_{o2}$) with =0.85 and =3.4 as the starting point did not provide an apparent maximum value within the area of imaginary wells. The ARS team assumed that was because of the use of the maximum concentration in the response function. The mean maximum concentration in well W$_{o2}$ (3.4 µg L^{-1}) was approximately 5 times smaller than in W$_{o1}$ (16.1 µg L^{-1}).
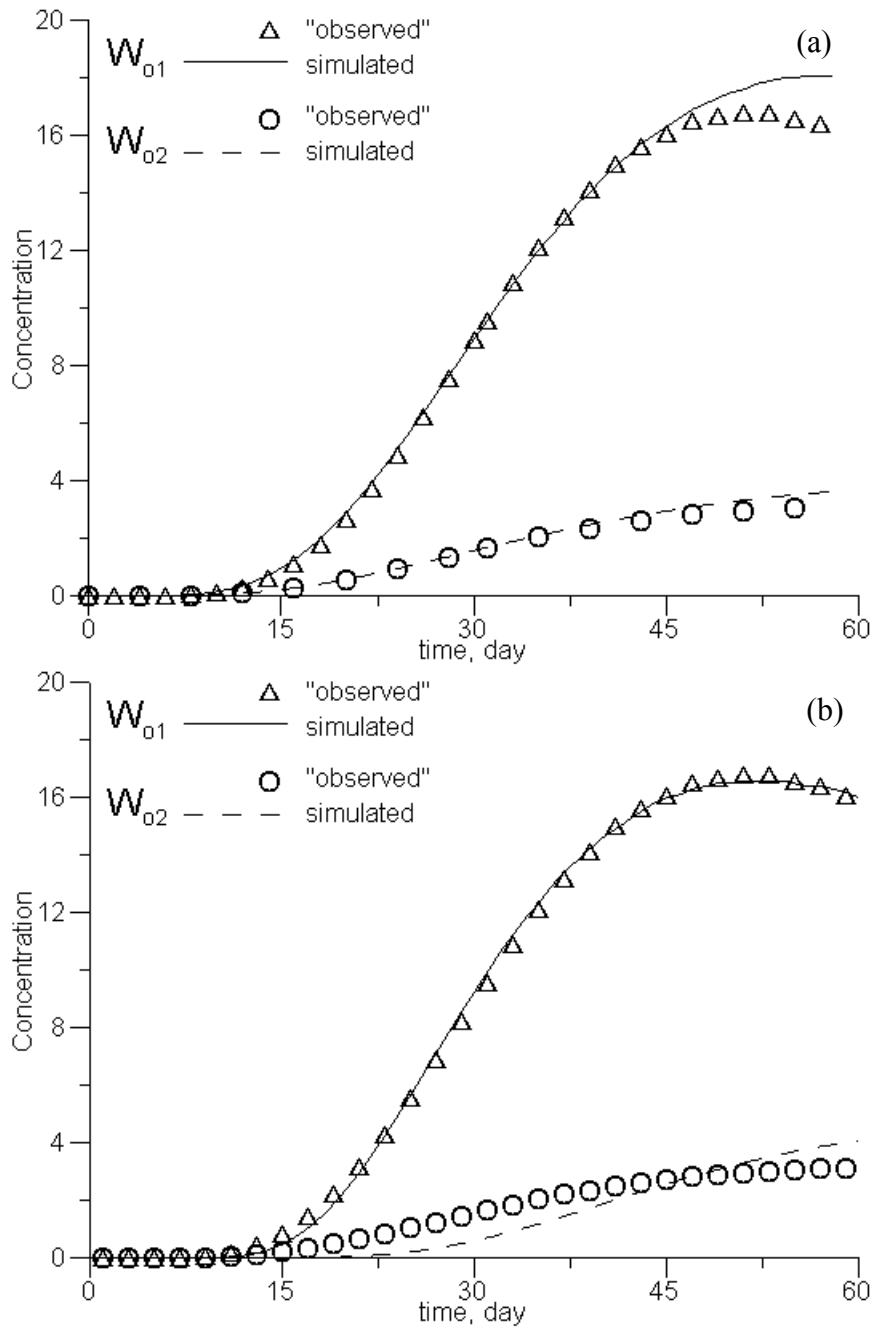
Figure 6-18 Comparison of the BTCs obtained in the HYDRUS simulations (a) and using the analytical model (b) with the observed in two observation wells.
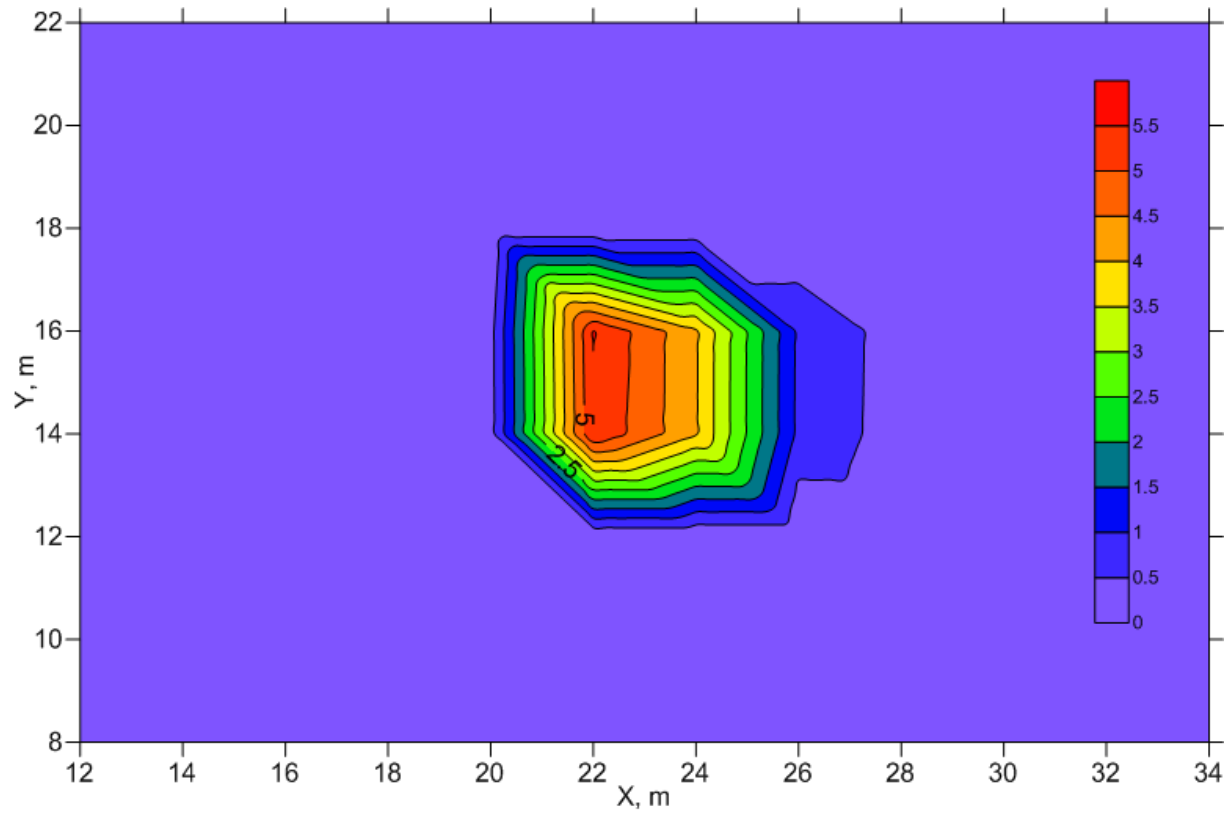
Figure 6-19   Distribution of the calculated values of discriminant function, Kv.

6.2.2.2 Application to the OPE3 Field Experiment

The Kullback-Himmelblau sequential design (KHSD) methodology was tested using data from the tracer experiment carried out in 2008 at the USDA-ARS OPE3 integrated research site (Fig. 3-18). A pulse of KCl solution was applied to an irrigation plot, and chloride concentrations were measured in groundwater at three sampling depths with 12 observations wells.

The spatial distribution of soil materials was obtained from cores taken from depths of 0-200 cm at 20 cm increment during installation of the observation wells. Soil texture measured at 12 locations between depths of 0.2 m and 2.0 m at 0.2 m increments was grouped into 9 clusters such that each class was represented by soil texture averaged over the samples belonging to this class texture (Fig. 4-3). These classes included: three sandy loam classes (SL1, SL2 and SL3), two sandy clay loam classes (SCL1 and SCL2), two loam classes (L1 and L2), and two silty clay loam classes (SiCL1 and SiCL2).

A 3D conceptual flow and transport model was developed to simulate water flow and chloride transport for the tracer experiment at the OPE3 site. The ARS team considered a 3-D layered domain that extended laterally for 60 m in the *x* (east-west) direction and for 50 m in the *y* (north-south) direction. The thickness (*z* direction) of the domain varied from 2.5 m to 4.3 m (Fig. 6-20). The finite-element mesh was composed of 6670 nodes and 11088 3D elements.

The KHSD method was used to search for the optimal location of the augmentation wells on a 2D grid including 120 imaginary wells (Fig. 6-20). For the response function the ARS study team calculated the maximum observed concentration (C_{max}) in each observation well during the 120 day simulation using HYDRUS-3D. The initial and boundary conditions were prescribed as follows:
- Equilibrium pressure head distributions relative to the groundwater levels measured in the observation wells at the beginning of the tracer experiment;
- Spatially variable, but constant in time, pressure head profiles specified along the lateral boundary;
- Variable initial Cl- distribution as measured in observation wells prior to the experiment;
- Third type (Cauchy) boundary conditions for transport along the soil surface;
- Cl- concentrations of groundwater along the lateral boundaries were set to the initial concentrations observed before the tracer experiment;
- Daily-averaged atmospheric boundary conditions at the soil surface beyond the irrigated plot, and daily-averaged irrigation depth less runoff at the soil surface within the irrigated plot (Fig. 4-7);
- A zero flux boundary condition along the bottom of the simulated domain;
- Cl- concentrations of 4940 mg L^{-1} and 8 mg L^{-1} in the irrigation water during and after the tracer application, respectively.

The ARS team considered discrimination between two models differing in terms of their subsurface structure lithology as follows (see also Fig. 6-21): Model 1 involving a layered medium containing 10 layers and 5 materials (every layer was assumed to be homogeneous), and Model 2 involving a medium with 10 layers and lenses, and 7 soil materials. Locations of lenses were delineated manually.
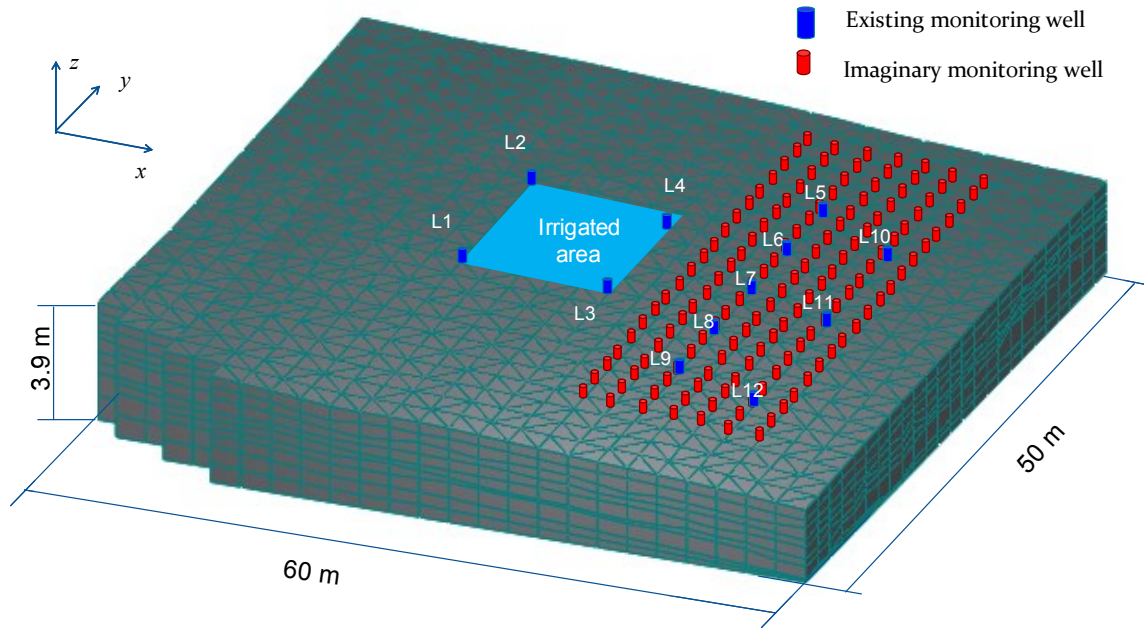
Figure 6-20   Modeling domain with finite element mesh and observation locations.
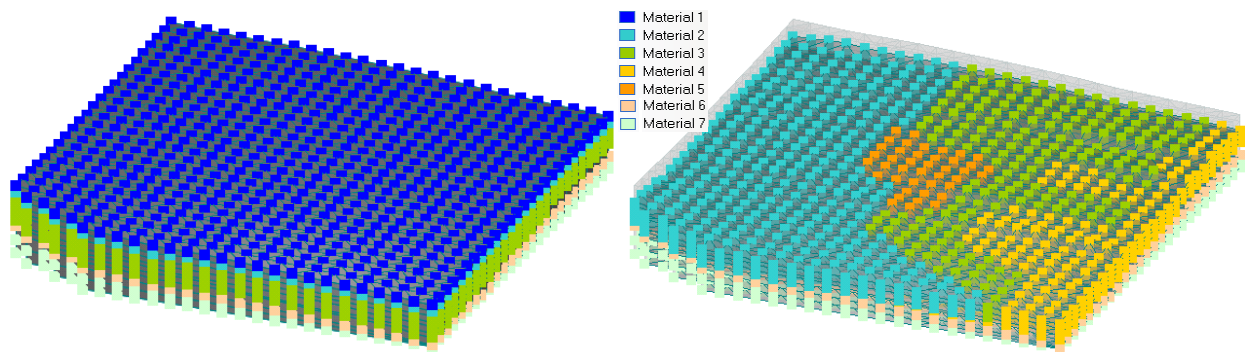
Figure 6-21   Soil materials in two subsurface structure models: a) Model 1 - Layered media, b) Model 2 - Layers and lenses media (shown layers 4-10)

The hydraulic conductivities and dispersivities were calibrated manually, while pedotransfer functions were used to condition the calibration results to produce an ensemble of models. Model parameters are presented in Table 6-1. The ensemble modeling was used to estimate variances of the means of model response (C_{max}) distributions. The ensembles of 25 realizations for each model were simulated by generating a lognormally distributed saturated hydraulic conductivity field with a prescribed standard deviation of $\log_{10}(K_s)$ equal to 0.25. An example stochastic realization of the hydraulic conductivity scaling factor is shown in Fig. 6-22).

Figure 6-23 shows the spatial distribution of the tracer concentration at the end of the simulation time. The ARS team noted that the distribution simulated using Model 2 exhibited a more definite preferential flow and transport patterns as compared to Model 1. Simulated BTCs at eight existing monitoring well (L5-L12) are presented in Fig. 6-24 for the two models. Values of the discriminant function, K_v, were evaluated at 120 locations of the imaginary monitoring wells (Fig. 6-20) on the 5th grid layer.

Five simulation scenarios were carried out by assuming that the nth experiment was performed in different existing monitoring well (those at locations L5 to L9). When using Bayesian estimates to calculate the discriminant function (Eq. 6-17), specific locations of max(K_v) were very different in each of the five simulations. The ARS team suggests that this was because of a lack of a good calibration of the model affecting the calculation of prior and posterior probabilities. Applying the simpler equation (6-13), which disregards Bayesian estimates, provided similar results for all five simulations as shown in Fig. 6-25. The maximum of K_v was located at x=498-500 m and y=370 m, east of well L8. The distinction between the BTCs simulated with two models was most significant at L8 (Fig. 6-17). Another local maximum was found between wells L6 and L10. The difference between the two BTCs was also considerable at well L6 (Fig. 6-24).

Table 6-1 Model parameters found by trial-and error method.

Parameter Texture	θ_r	θ_s	α, m^{-1}	n	K_s m/day	a_L m	a_T m
Material 1	0.049	0.361	2.31	1.37	0.3	0.25	0.05
Material 2	0.046	0.331	3.06	1.26	0.3	0.25	0.05
Material 3	0.0422	0.317	3.54	1.24	1.0	0.25	0.05
Material 4	0.044	0.313	3.30	1.22	0.2	0.25	0.05
Material 5	0.047	0.314	2.51	1.23	0.1	0.25	0.05
Material 6	0.060	0.328	1.22	1.32	0.01	0.25	0.05
Material 7	0.060	0.328	1.22	1.32	0.0001	0.25	0.05
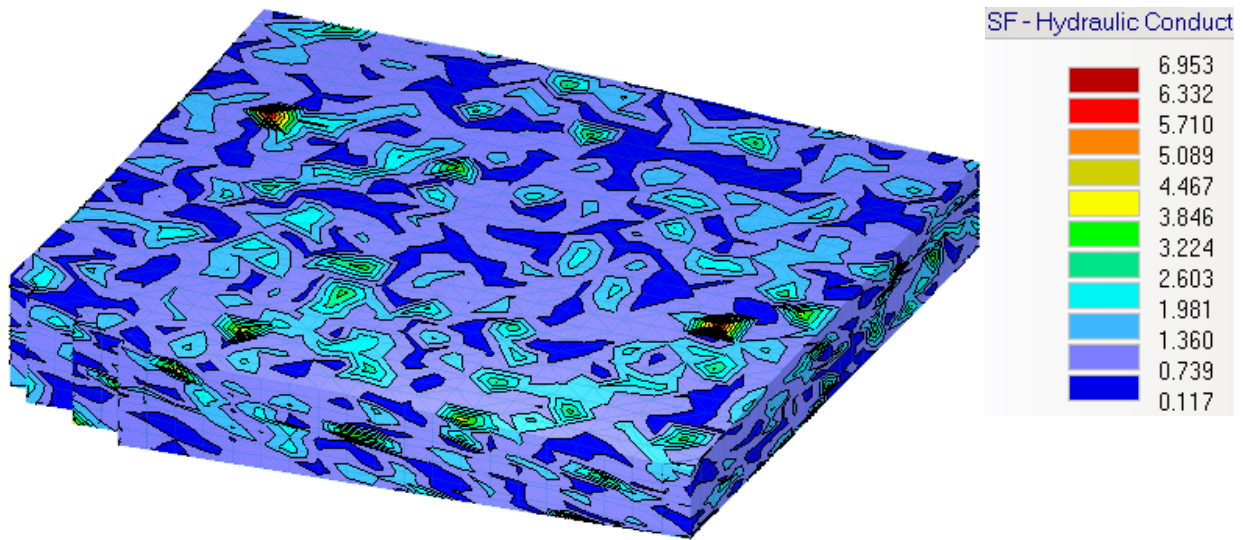
Figure 6-22   Example of a stochastic realization for the hydraulic conductivity scaling factor (SF).
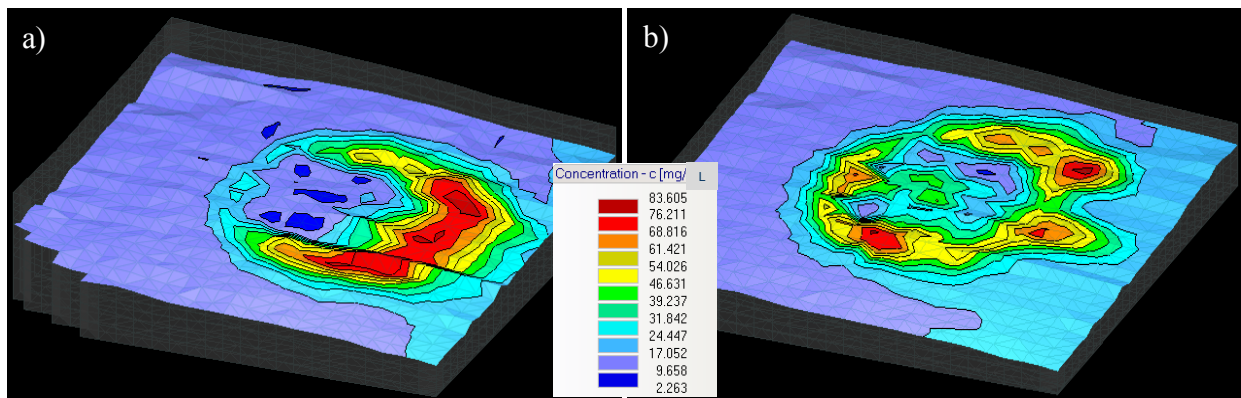
Figure 6-23   Spatial distribution of tracer concentration after 120 days in 5th layer (z=2.75 m): a) Model 1, b) Model 2.
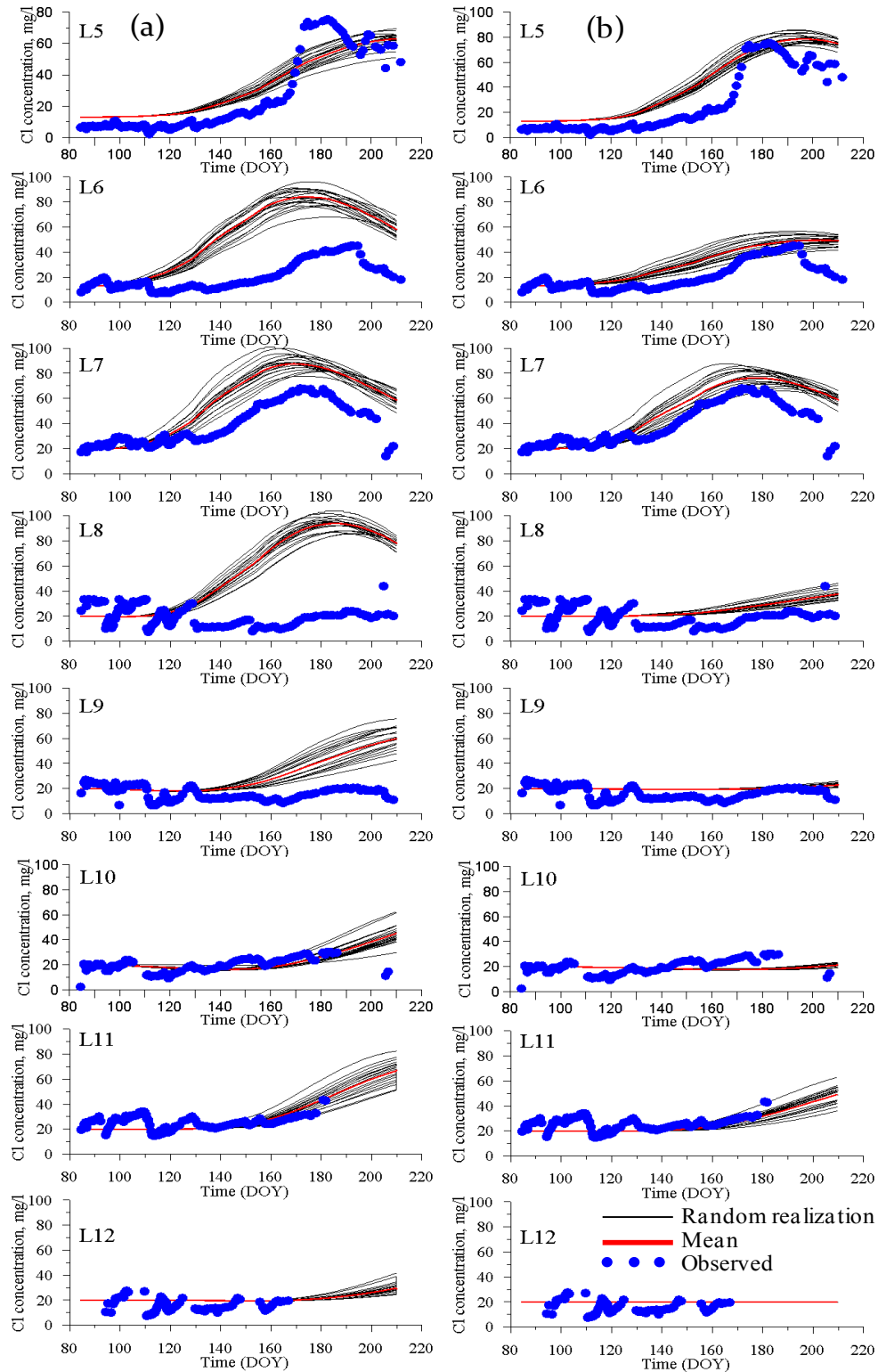
Figure 6-24   Simulations of the tracer BTCs in locations 5-12: a) Model 1; b) Model 2.
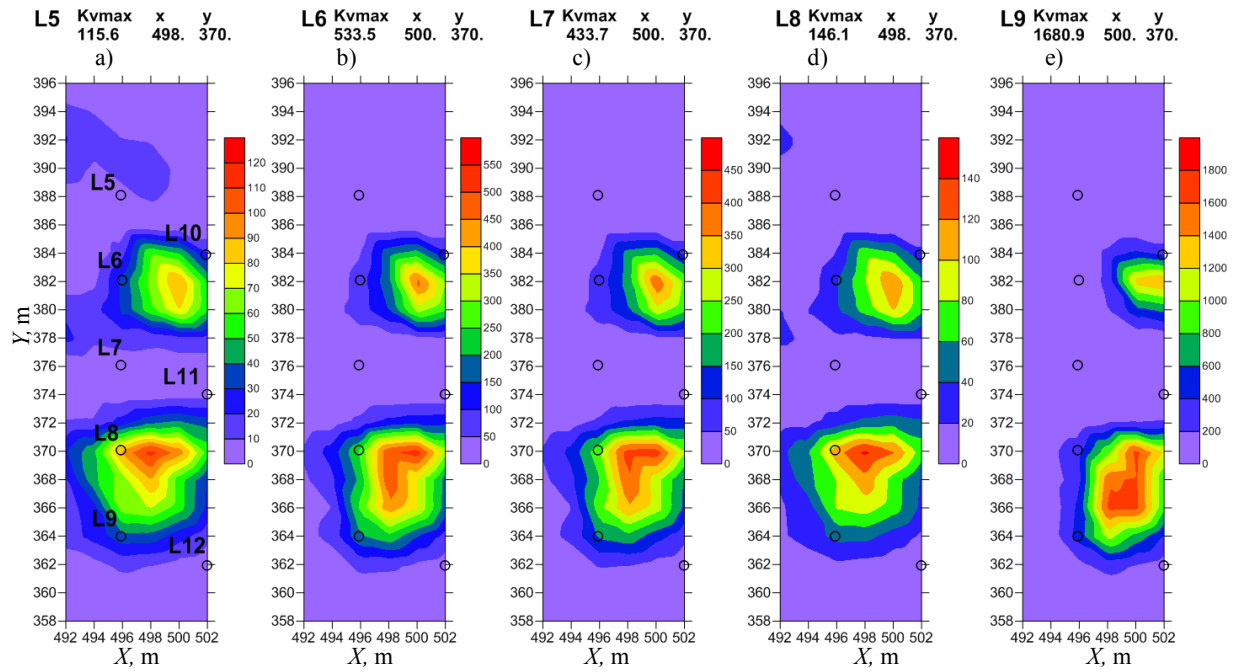
Figure 6-25 Discriminant function (Kv) distribution in five simulations: (a)-(e) scenarios of the _n_th experiment was performed in wells L5-L9, respectively.

7 SUMMMARY AND CONCLUSIONS

The objective of this project was to demonstrate applicability of model abstraction techniques for optimization of monitoring networks at US NRC licensing sites. The study focused on two aspects of network optimization: (1) identification of performance indicators providing reliable information about subsurface contaminant transport, and (b) defining monitoring locations that would improve model calibration and reduce risk of incorrect detection of contaminant releases.

To accomplish the project goal, the ARS study team used data from the OPE3 experimental field site near Beltsville, MD. This site has been extensively studied for more than 10 years using physical, geophysical, biophysical, remote sensing, and groundwater monitoring methods. All available data were analyzed in a systematic way and used to construct a conceptual site model and simulation domain. Special attention was paid to the spatial heterogeneity and organization of soil structural units composing the OPE3 site.

A complex set of methods was employed to study the heterogeneity. A ground penetrating radar survey was conducted with 2x2 m spatial resolution and a scanning depth of approximately 3 m. Results of the survey were analyzed jointly with the water content monitoring data to delineate soil hydrologic layers. Results of the soil surveys conducted within the 10-year study period revealed considerable heterogeneity at the site, and provided detailed information about spatial organization of the soil structural units. The spatial heterogeneity was confirmed by direct and remote measurements at different scales. At the soil core scale these methods included soil chemical analysis, soil texture analysis, and measurements of the soil bulk density, soil water retention and saturated hydraulic conductivity. At the soil profile scale the ARS team used soil water content, pressure head and groundwater monitoring. At the field scale, methods of electromagnetic induction, electrical resistivity, color infrared imaging and direct measurements of corn biomass were implemented.

At the soil core scale laboratory water retention data clearly indicated the presence of soil macroporosity that potentially could cause fine-scale high-conductivities in parts of the soil pore space. Hydraulic conductivity measurements at the same scale showed that, in actuality, such high conductivity regions would require not only the presence of macropores but also of a connectivity between these pores, which could be found only in a relatively small percentage of locations.

At the soil profile scale soil moisture and soil pressure head monitoring data provided the required information about field hydraulic properties of unsaturated soils. Field water retention was well defined by those measurements. This indicated that the Richards model would be an appropriate conceptualization of the flow processes in the soils under study.

Soil electrical resistivity monitoring at the field scale further revealed the existence of narrow hydrologically active vertical zones. The obtained data indicated that some parts of the subsurface did not fully participate in the prevailing flow and transport processes. Biophysical monitoring of the research area allowed delineation of relatively large structural units in the soil cover that provided distinctly different conditions for plant growth. The size of those units exceeded the scale of investigation of this project.

To validate applicability of the model abstraction techniques for model calibration, a field experimental study was designed and carried out. The experiment involved the application of a pulse of conservative tracer on a 10x10 m plot equipped to monitor soil water contents,

pressure heads, groundwater depths and runoff. The tracer concentrations were monitored at 3 depths in 5 and 3 wells installed at distances of 7 m and 14 m from the irrigated plot, respectively. The concentration measurements were supplemented with water content and groundwater depth data for all wells. The data collected within the 130-day monitoring period provided information for model calibration. The collected dataset included 28 breakthrough curves measured at 12 locations (involving a total of 7,280 tracer concentrations). The data indicated the presence of a fine-textured, low-permeability layer diverting the tracer according to the topography of the layer. There were no indications that the Buckingham-Darcy flow model and corresponding Richards equations were inapplicable at the larger scale, likely because flow and transport at that scale seemed to be driven mostly by the hydraulic gradients in groundwater. The complexity of the flow and transport domain required the use of a full 3D model. For this reason, the ARS study team used the HYDRUS-3D and TOUGH2 software packages to simulate flow and transport in the lateral flow experiment. For the TOUGH2 simulations, the radionuclide/tracer transport module T2R3D was coupled with the EOS9 module describing a saturated-unsaturated water flow in a porous medium.

Since both models do not have a calibration option, trial and error calibration was applied in this study. The calibration objective functions consisted of deviations of simulated groundwater depths and measured concentrations. The calibrated parameters were the saturated hydraulic conductivity and the longitudinal dispersivity of 8 soil materials composing the simulation domain. The ratio of longitudinal to transversal dispersivity was set to 5 in all simulations to reduce the number of fitting parameters. The water retention parameters were estimated from basic soil properties using the ROSETTA software for all materials. The trial-and-error calibration was reasonably successful for HYDRUS-3D, yielding satisfactory description of the peak concentrations and their arrival times. However, both models failed to reproduce accurately the shape of the breakthrough curves, while the peaks in concentration could not be predicted with the TOUGH2 simulations.

A sensitivity analysis was conducted to evaluate applicability of the different abstraction techniques to simplify the calibrated model. Using modeling results, the ARS team examined the combination of the simulation domain composition with specific transport properties and fluxes on the domain boundaries. The variance-based first order sensitivity coefficient (S_i) and the derivative-based sensitivity index (D_i) were used in the sensitivity analysis. Four abstraction techniques were tested using HYDRUS-3D and TOUGH2: (1) parameter abstraction, (2) subsurface textural units, (3) boundary conditions, and (4) time influence. The accuracies of the abstraction techniques were evaluated in terms of three performance indicators: (1) the tracer time series, (2) the peak concentrations (C_{peak}), and (3) the times to peak concentration (T_{peak}). Results of the sensitivity analysis showed that introducing the peak concentration and the time to peak concentration as performance indicators for the LSA allowed a significant reduction in the input information for the sensitivity analysis, and definition of the soil material that could be used for model simplification. The results also revealed a limitation of the C_{peak} and T_{peak} performance indicators in that they needed having peaks on the simulated BTCs. Both parameter abstraction and alteration of the boundary between soil materials (abstraction of subsurface textural units) identified the same material as a candidate for model simplification. Further analysis showed that this material could be replaced with another material with similar soil texture.

Among the boundary conditions, averaging evapotranspiration rates did not affect the performance indicators, while runoff and groundwater depth did have an effect and therefore were not a subjected to model simplification. The sampling time influenced the performance indicators when the sampling intervals were longer than 2 and 10 days within and beyond the irrigated plot,

respectively. Overall, the sensitivity analysis appeared to be a useful tool to identify the directions for model abstraction.

Results of the sensitivity analysis were also used to evaluate the importance of observations collected in the monitoring wells during the lateral flow experiment, and to identify new locations to improve the HYDRUS-3D calibration. To do this, the ARS study team implemented the Observation-Prediction statistics (OPR) method, which computes the percent of changed uncertainty OPR caused by omitting or adding one observation node. The OPR method indicated that there was no need for data collection at two depths for model calibration, and emphasized the overall importance of monitoring at a depth corresponding to the z=2.49 m computational layer. The OPR method also appeared to be sensitive to values of the sensitivity parameters and their spatial structure. New monitoring wells identified using the OPR method were located in zones of transition from relatively high to low concentrations where the maximum model predicting uncertainty can be expected. Based on results of this study, the OPR method is recommended for selecting monitoring locations to improve model calibration.

The applicability of different abstraction approaches was also evaluated at the scale of a decommissioning or license NRC site. The size of the simulation domain for this purpose was increased to 118x190 m to mimic the scale of a NRC site. Soil properties and boundary conditions of the extended domain were set based on a 10-year soil survey and groundwater monitoring data. Three precipitation scenarios were generated based on 60-year weather data in Beltsville, MD. These scenarios included decades with 25%, 50% and 75% probabilities of the precipitation. The ARS study team assumed that the 10-year simulation period was sufficient for the tracer travel distance between the irrigation plot and the eastern edge of the simulation domain, and that the three selected probabilities adequately represent the temporal variability of precipitation.

Four release scenarios and four groundwater scenarios represented uncertainties associated with the source of the contaminant release and the subsurface flow conditions. The release was simulated as a 10-day contaminant pulse from either the east-north or east-south points within the irrigated plot located on the surface or at the depth of 2.3 m. A total of 48 different scenarios were generated for HYDRUS-3D simulations, which included: 4 release scenarios x 4 groundwater scenarios x 3 precipitation scenarios. These scenarios provided input for the model abstraction techniques.

The simulated tracer concentrations and fluxes were chosen in this study to characterize the tracer spread from the potential points of accidental release, while the simulated peak concentration (C_{peak}) and time to peak concentration (T_{peak}) at the observation nodes and transects at different distances from the release points characterized the direction and velocity of the contaminant plume. The accuracy of the model abstraction techniques was evaluated using the Nash and Sutcliffe efficiency criterion, the Willmot efficiency criterion and the coefficient of determination.

Three abstraction techniques were validated for this scale. They were: (1) abstraction using pedotransfer functions, (2) profile aggregation, and (3) abstraction by ignoring the unsaturated zone. In abstraction (1) the calibrated values of the saturated hydraulic conductivities were replaced by their PTF estimates, in abstraction (2) the heterogeneous soil profile was replaced with a homogeneous profile composed of the dominant soil material in the domain, while in abstraction (3) unsaturated contaminant transport was discarded by placing the release source on the groundwater table. Techniques (1) and (2) were found to be inappropriate for model abstraction. This was confirmed by all performance indicators. Technique (1) altered significantly

the contaminant concentrations, flow velocities and total mass transported, while technique (2) altered the flow direction. Ignoring the unsaturated zone appeared to be the only abstraction applicable for model simplification in our study. This result was attributed to the presence of perched water in our study and should be used cautiously at NRC sites with thick unsaturated zones.

The same abstraction techniques were used to identify monitoring locations that had high probability to detect contaminant release and transport from a NRC decommissioning or licensing site. Results of HYDRUS-3D simulations for 48 scenarios were used to identify locations at different distances from the release points with (a) more frequent, and (b) more probable and persistent appearance of extreme values of the performance indicators, while the second approach involved a probability analysis. Overall, both the frequency and probability of the peak performance indicators appeared to provide appropriate statistics to identify monitoring locations with highest probability of contaminant plume detection. The observed difference between the time-based (T_{peak}) and concentration- (C_{peak}) or flux- (QC) based performance indicators were attributed to the different nature of the indicators. Arrival time did not appear to be a good predictor of plume propagation in high heterogeneous media due to the combined effect of contaminant dispersion and dilution on contaminant transport. The C_{peak} and QC indicators were more reliable to identify the monitoring locations in this study.

Validation of the three model abstraction techniques to identify monitoring locations confirmed applicability of the abstraction by ignoring the unsaturated zone for contaminant transport monitoring. However, the results of this abstraction may differ when a thicker vadose zone is present, or when horizontal lenses may slow down vertical transport. Monitoring locations obtained using the abstracted model matched the network identified by the calibrated model. In case of parameter abstraction, the abstracted model systematically underestimated peak concentrations and overestimated total fluxes. Little or no similarities were found between the travel times of the calibrated and abstracted models. Surprisingly, in spite of inaccuracies in the parameter abstracted model, the parameter abstraction approach appeared to be appropriate for monitoring location selection in this study. The homogeneous soil profile was the only abstraction technique which generated a monitoring network dissimilar to the network obtained with the calibrated HYDRUS model. The selected monitoring locations were systematically shifted to the south from the calibrated model, with the difference increasing with the distance from the source of release. This occurred primarily due to existence of the preferential lateral flow in the heterogeneous simulation domain, which was entirely discarded using homogeneous profile abstraction.

Overall, this study demonstrated that model abstraction is an important component for development of monitoring strategies at NRC decommissioning or licensing sites. At the stage of data collection and analysis, model abstraction can identify sources of predictive uncertainty and optimize monitoring networks for the model calibration specific for the selected performance indicators. At the monitoring stage, model abstraction serves as a tool for model conceptualization and calibration. The calibrated model in turn provides most probable directions and velocities of contaminant transport from a potential release point under different scenarios, even when the source and the released mass are not well defined. Based on the model predictions some old locations can be excluded and new locations can be included into the monitoring network to provide more reliable information about potential contaminant transport from the release points. This multistep approach can be iterative since new monitoring data can provide important information for model conceptualization and improvement of the model calibration.

The Kullback-Himmelblau sequential design method was used for augmentation of the monitoring well network assuming that comparison and discrimination of models can be improved if additional observations are included. The method is based on (1) generalization of Kullback's discriminant function and "weights of evidence" for the case of available prior probabilities, and (2) ensemble modeling to estimate the variance of the predicted values. The method was tested using data from (1) a synthetic example generated with HYDRUS-3D by simulating pulse injection of a conservative pollutant into an aquifer composed of relatively uniform sediments, and (2) the tracer experiment at the USDA-ARS OPE3 integrated research site, carried out in 2008. In each case two models were used for the simulations: a complex and an abstracted model. Models used for the synthetic example were the HYDRUS-3D model, and an analytical solution of Lenda and Zuber (1970) for 1D uniform flow and 3D transport.

To simulate flow and chloride transport for the tracer experiment at the OPE3 site two three-dimensional flow and transport models were used based on the HYDRUS-3D code: a model representing a layered medium; and a model accounting for many layers and lenses. Calibration of hydraulic conductivities and dispersivities was performed, and pedotransfer functions were conditioned to calibration results to build an ensemble of models. A search of the optimal location for augmentation wells was carried out on a 2D grid downstream of the tracer injection. A single response function of the maximum observed concentration was used to discriminate between the models. Results indicate that the KHSD method can be used to find the optimum location of additional observation wells. Well-calibrated models provide a single optimal location, but if models are not well calibrated the Bayesian estimates strongly depend on the sequence of observation wells. Generalization of the method is possible for multiple response cases, such as maximum concentrations and arrival times. The information theory-based KHSD method to augment the observation network shows promise for improving the monitoring of flow and transport for discrimination of conceptual models of the subsurface. Outcomes of this study can provide information for future data collection and monitoring efforts that further reduce the uncertainty.

8   REFERENCES

Abit, S.M., A.Amoozegar, M.J. Vepraskas, and C.P. Niewoehner, "Solute Transport in the Capillary Fringe and Shallow Groundwater: Field Evaluation," *Vadose Zone Journal ,* 7: *890-898,* 2008.

Allen, G., L.S. Pereira, D. Raes, and M Smith. 1998. "Crop Evapotranspiration - Guidelines for Computing Crop Water Requirements, FAO Irrigation and Drainage Paper 56, 1998. Available at http://www.fao.org/docrep/X0490E/x0490e00.htm (contents accessed on August 20, 2009.

Altman, S.J., B.W. Arnold, R.W. Barnard, G.E. Barr, C.K. Ho, S.A. McKenna, and R.R. Eaton, "Flow Calculation for Yucca Mountain Groundwater Travel Time,"(*GWTT-95*). *SAND96-0819*, Sandia National Laboratories, Albuquerque, NM, 212pp., 1996.

Arndt, O., T.Barth, B. Freisleben, and M. Grauer, "Approximating a finite element model by neural network prediction for facility optimization in groundwater engineering," *European Journal of Operational Research,* 166 :769–781, 2005.

Aziz, J.J., M. Ling, H. S. Rifai, C. J. Newell, and J. R. Gonzales, "MAROS: A decision support system for optimizing monitoring plans," *Groundwater,* 41:355-367, 2003.

Bandurraga, T.M., and G.S. Bodvarsson, "Calibrating Hydrogeologic Parameters for the 3-D Site-scale Unsaturated Zone Model of Yucca Mountain, Nevada," *Journal of Contaminant Hydrology*, 38(1-3): 25-46, 1999.

Barthel, R., "Common problematic aspects of coupling hydrological models with groundwater flow models on the river catchment scale," *Adv. Geosci.,* 9" 63–71. 2006.

Bashi-Azghadi, S. N., and R. Kerachian,." Locating monitoring wells in groundwater systems using embedded optimization and simulation models," *Science of the Total Environment*,408: 2189–2198, 2010.

Bierkens, M. F. P. , "Designing a monitoring network for detecting groundwater pollution with stochastic simulation and a cost model," *Stochastic Environmental Research and Risk Assessment*, 20: 335–351, 2006.

Bigelow, J.H., and P.K. Davis, *Implications for Model Validation of Multiresolution, Multiperspective Modeling (MRMPM) and Exploratory Analysis*, MR-1750, RAND, 2138, Santa Monica, CA, 2003

Bloomfield, J. P., "*Rationale for groundwater monitoring in the context of the Lowland Permeable Catchment Research Programme (LOCAR)*", Technical Report WD/OO/05, British Geological Survey, Natural Environment Research Council Keyworth, Nottinghamshire, 2000

Blöschl, G., and Sivapalan, M., "Scale Issues in Hydrological Modeling", *Hydrological Processes*, 9: 251-290, 1995.

Boyle, D.P., H.V Gupta, J., Soroshian, V. Koren, Z. Zhang, and M. Smith, "Toward Improved Streamflow Forecasts: Value of Semidistributed Modeling," *Water Resources Research*, 37: 2749-2759, 2001.

Box, G.E. and W.J. Hill, "Discrimination among mechanistic models". *Technometrics* 9, 57-71, 1967.

Brooks R., P. Jamieson, and M. Semenov, "Simplifying Sirius: sensitivity analysis and development of a meta-model for wheat yield prediction", Available at: http://eprints.lancs.ac.uk/29412/, 2010.

Bulygina, N., H. Gupta, H., "Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation", *Water Resources Research*, 45, W00B13, 2009.

Burnham, K. P., and D.R. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection", *Sociological Methods & Research*, 33, 261-304, 2004

Cariboni, J., D. Gatelli, R. Liska, and A. Saltelli, " The role of sensitivity analysis in ecological modeling," Ecological *Modelling*, 203(1-2): 167-182, 2007.

Celia, M.A., E.T. Bouloutas, and R.L. Zarba, "A general mass-conservative numerical solution for the unsaturated flow equation," *Water Resource Research*, 26: 1483–1496, 1990.

Chadalavada, S., and B. Datta, "Dynamic Optimal Monitoring Network Design for Transient Transport of pollutants in Groundwater Aquifers," *Water Resources Management*, 22:651–670, 2008.

Chen, J., S. Hubbard, and Y. Rubin, "Estimating The Hydraulic Conductivity at the South Oyster Site From Geophysical Tomographic Data Using Bayesian Techniques Based on the Normal Linear Regression Model," *Water Resources Research*, 37(6): 1603–1613, 2001.

Chen, J., S. Hubbard, Y. Rubin, C. Murray, E. Roden, and E. Majer, "Geochemical Characterization Using Geophysical Data and Markov Chain Monte Carlo Methods: a Case Study at the South Oyster Bacterial Transport Site in Virginia," *Water Resources Research*, 40(12), W12412, 2004.

Chinkuyu,, A. Guber, T. Gish, D. Timlin, J. Starr, T. Nicholson, R. Cady, A. Schwartzman. *Field Studies to Confirm Uncertainty Estimates of Ground-Water Recharge*, NUREG/CR-6946. U.S. Nuclear Regulatory Commission. Washington, D.C. 20555-0001, 2008.

Cieniawski, S. E., J. W. Eheart, and S. Ranjithan, "Using genetic algorithms to solve a multiobjective groundwater monitoring problem," *Water Resources Research*, 31(2): 399–409, 1995.

Confalonieri R, G. Bellocchi, S. Bregaglio, M. Donatelli, and M.Acutis, "Comparison of sensitivity analysis techniques: A case study with the rice model WARM," *Ecological Modelling*, 221(16):9, 2010, Available at: http://dx.doi.org/10.1016/j.ecolmodel.2010.04.021

Coumou, D., I. Mattha, S. , Geiger, S., Driesner, T., "A parallel FE– FV scheme to solve fluid flow in complex geologic media," *Computers & Geosciences,* 34: 1697– 1707, 2008.

Dane, J.H. and G.C. Topp (Eds.) "*Methods of soil analysis. Part 4. Physical Methods*". ASA and SSSA, Madison WI, 2002.

Dane, J.H. and J.W. Hopmans, "Pressure cell," *In* Dane, J.H., and G.C. Topp - ed., *Methods of Soil Analysis. Part 4. Physical Methods*, pp. 684-688, ASA and SSSA, Madison, WI, 2002.

Datta, B., D. Chakrabarty, and A. Dhar, A. "Optimal Dynamic Monitoring Network Design and Identification of Unknown Groundwater Pollution Sources," *Water Resources Management*, 23:2031–2049, 2009.

Davis, P.K., and J. H. Bigelow, "*Motivated Metamodels: Synthesis of Cause-effect Reasoning and Statistical Metamodeling. MR-1570*," RAND, 2138, Santa Monica, CA, 2003.

Deiana, R., G. Cassiani, A. Villa, A. Bagliani, and V. Bruno, "Calibration of a Vadose Zone Model Using Water Injection Monitored by GPR and Electrical Resistance Tomography," *Vadose Zone Journal*, 7(1): 215 – 226, 2008.

Dhar, A. and B. Datta, "Multiobjective Design of Dynamic Monitoring Networks for Detection of Groundwater Pollution," *Journal of Water Resources Planning and Management*, 133(4): 329–338, 2007

Do, H., and G. Rothermel, "Using Sensitivity Analysis to Create Simplified Economic Models for Regression Testing," ISSTA'08, July 20–24, Seattle, Washington, USA. 2008.

Durner, W., "Hydraulic Conductivity Estimation for Soils with Heterogeneous Pore Structure", *Water Resources Research*, *32*(9): 211-223, 1994.

Draper, N.R., and H. Smith, H., "*Applied regression analysis"* (2nd ed.): New York, John Wiley, 709 p., 1981

Facchi, A., B. Ortuani, D. Maggi, and C. Gandolfi, "Coupled SVAT–groundwater model for water resources simulation in irrigated alluvial plains," *Environmental Modelling & Software*, 19: 1053–1063, 2004.

Foglia, L., S.W. Mehl, M.C. Hill, P. Perona, and P. Burlando, "Testing alternative ground water models using cross-validation and other methods", *Ground Water,* 45, 627-641, 2007.

Gee, G.W. and Or, D., "Particle-Size 2004.Analysis" In Dane, J.H. and G.C. Topp (ed.) *Methods of soil analysis. Part 4. Physical Methods*, ASA and SSSA, Madison WI:278-282, 2002.

Gerke, H.H., and M.Th. van Genuchten, "A Dual-Porosity Model for Simulating the Preferential Movement of Water and Solutes in Structured Porous Media," *Water Resources Research*, 29: 305-319, 1993.

Gish, T.J. and K.-J.S. Kung, "Procedure for Quantifying a Solute Flux to a Shallow Perched Water Table," *Geoderma*, 138:57–64, 2007.

Gish, T.J., C.L. Walthall, C.S.T. Daughtry, and K.-J.S. Kung, "Using Soil Moisture and Spatial Yield Patterns to Identify Subsurface Flow Pathways," *Journal of Environmental Quality*, 34:274-286, 2005.

Gish, T.J., W.P. Dulaney, C.S.T. Daughtry, and K-J.S. Kung, "Use of Ground-penetrating Radar to Identify Ground Water Pathways at the Watershed Scale", *Soil Science Society of America Journal*, 66: 1620-1629, 2002.

Gish, T.J., C.L. Walthall, C.S.T. Daughtry, W.P. Dulaney, and G.W. McCarty, "Watershed-scale sensing of subsurface flow pathways at OPE3 site", In: *Proceedings of the First Interagency Conference on Research in the Watersheds, October 27-30, 2003, Benson, Arizona*. p. 192-197, 2003.

GMS 6.0. Groundwater modeling system. Brigham Young University. 2002.

Grabow, G., D. C. Yoder, and C. R.Mote, "An empirically-based sequential ground water monitoring network design procedure, "*Journal of the American water resources association*, 36(3): 549-446, 2000.

Guber, A.K., T.J. Gish, Y.A. Pachepsky, M.T. van Genuchten, C.S.T. Daughtry, , T.J. Nicholson, and R.E. Cady, "Temporal Stability of Soil Water Content Patterns Across Agricultural Fields". *Catena,* 73: 125-133, 2007.

Guber, A.K., Y.A. Pachepsky, M.Th. van Genuchten, W.J. Rawls, D. Jacques, J. Simunek, R.E. Cady, and T.J. Nicholson, "Field-Scale Water Flow Simulations Using Ensembles of Pedotransfer Functions for Soil Water Retention," *Vadose Zone Journal*, 5:234–247, 2006.

Guber, A.K., Y.A. Pachepsky, M.Th. van Genuchten, J. Simunek, D. Jacques, A. Nemes, T.J. Nicholson, and R.E. Cady, "Multimodel Simulation of Water Flow in a Field Soil Using Pedotransfer Functions," *Vadose Zone Journal*, 8:1–10, 2008.

Guswa, A.J. and D.L. Freuberg, "On Using the Equivalent Conductivity to Characterize Solute Spreading in Environments with Low-Permeability Lenses," W*ater Resources Research*, 38(8): 1132, 2002.

Helton, J. C., C. W. Hansen, and C. J. Sallaberry, "Uncertainty and sensitivity analysis in performance assessment for the proposed repository for high-level radioactive waste at Yucca Mountain, Nevada," *Procedia - Social and Behavioral Science*, 2(6):7580-7582, 2010.

Harter, T, J., and W. Hopmans, "Role of vadose-zone flow processes in regional-scale hydrology: review, opportunities and challenges," In: R. A. Feddes, G. H. De Rooij, J.C. van Dam – Ed., *Unsaturated-zone Modeling: Progress, Challenges And Applications*, pp.179-208, Kluwer, 2004.

Hassan, A. and K. Hamed, "Prediction of Plume Migration in Heterogeneous Media Using Artificial Neural Networks," *Water Resources Research*, 37(3): 605-624, 2001.

Hassan, A. E., J. H. Cushman, and J. W. Delleur, "A Monte Carlo assessment of Eulerian flow and transport perturbation models", *Water Resources. Research*, 34(5), 1143–1163, 1998.
.

Hassan, A. E., "Long-term Monitoring Plan for the Central Nevada Test Area", DOE/NV/13609-30, U.S. Department of Energy, Office of Scientific and Technical Information, Oak Ridge, TN, 2003.

Harbaugh, A.W., E.R.Banta, M.C. Hill, and M.G. McDonald, "MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model—User guide to modularization   concepts and the ground-water flow processes," U.S. Geological Survey Open-File Report 00-92, 121 p., 2000.

Haykin, S., "Neural Networks and Learning Machines (3rd Edition)", Pearson, Prentice Hall, NY, 2008

Hill, M.C., "*Preconditioned Conjugate-Gradient 2 (PCG2), a computer program for solving ground-water flow equations*," U.S. Geological Survey Water-Resources Investigations Report 90-4048, 43 p., 1990.

Hill, M.C., "Methods and guidelines for effective model calibration: U.S. Geological Survey Water-Resources Investigations Report 98-4005", 90 p., 1998.

Hill, M. C., and C. R. Tiedeman, "*Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty*," Wiley, 2007.

Hill, M.C., "The Practical Use of Simplicity in Developing Ground Water Models," *Ground Water*, 44(6): 775-781, 2006.

Hill, M.C., D. M. Ely, C.R. Tiedeman, G. M. O'Brien, F.A. D'Agnese, and C.C. Faunt, "*Preliminary evaluation of the importance of existing hydraulic-head observation locations to advective-transport predictions*," Death Valley regional flow system, California and Nevada. Water-Resources Investigations Report 00-4282, U.S. GEOLOGICAL SURVEY, Denver, Colorado, 2001.

Himmelblau D.M., "*Process analysis by statistical methods*," John Wiley, New York, 463 p, 1970.

Hudak, P. F., and H. Loaiciga," **C**onjunctive vadose and saturated zone monitoring for subsurface contamination," Environmental *Monitoring and Assessment,* 59: 15–29, 1999.

Hudak, P. F., and H. A. Loaiciga , "A location modeling approach for groundwater monitoring network augmentation", *Water Resources Research*, 28(3), 643–649, 1992.

Hudak, P. F, and H. A. Loaiciga, "An optimization method for monitoring network design in multilayered groundwater flow systems," *Water Resources Research*, 29(8):2835–2845, 1993.

Hudak, P.F., "A method for monitoring ground water quality near waste storage facilities." *Environmental Monitoring and Assessment*, 30, 197-210. 1994.

Hunt, R. J., D. E. Prudic, J. F. Walke, and M. P.Anderson, "Importance of Unsaturated Zone Flow for Simulating Recharge in a Humid Climate", *Groundwater*, 46(4), 551–560, 2008.

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

Iooss, B., M. Ribatet, A. Marrel, "Global sensitivity analysis of stochastic computer models with generalized additive models", *Technometrics*, submitted, Available at URL: http://fr.arxiv.org/abs/0802.0443v1. 2008.

Jarvis, N.J. "Modeling the Impact of Preferential Flow on Nonpoint Source Pollution," pp. 195-221 in: H.M. Selim and L. Ma (eds*.), Physical Nonequilibrium in Soils: Modeling and Applications*, Ann Arbor Press, Chelsea, MI. 1999

Kim, K. H., K. K. Lee, "Optimization of groundwater-monitoring networks for identification of the distribution of a contaminant plume," *Stochastic Environmental Research and Risk Assessment*, 21:785–794, 2007

Kim, K., "The EPRI Groundwater Protection Program – Technology R&D for Groundwater Protection," available at the web as http://hps.ne.uiuc.edu/rets remp/PastWorkshops/2009/Presentations/monday/pm/EPRI_Nuclear_Groundwater-06 10-09.ppt, accessed on 6/111/2012, 2009.

Kitanidis, P. K. and E. G. Vomvoris, "A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations," *Water Resources Research.,* 19(3): 677-690, 1983.

Knopman, D. S. and C. I. Voss, "Multiobjective Sampling Design for Parameter Estimation and Model Discrimination in Groundwater Solute Transport", *Water Resources Research*, 25(10): 2245-2258, 1989.

Knopman, D. S., C. I. Voss, and S. P. Garabedian, "Sampling Design for Groundwater Solute Transport' Tests of Methods and Analysis of Cape Cod Tracer Test Data", *Water Resources Research*, 27(5):925-949, 1991.

Kollat, J.B., and P.M. Reed, "Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design", *Advances in Water Resources*, *29*(6):792-807, 2006

Kollat, ,J.B., P. M. Reed, and R. M. Maxwell, "Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics ," *Water Resources. Research*, 47, W02529, 2011.

Krause, P. and W.-A. Flugel, "Integrated research on the hydrological process dynamics from the Wilde Gera catchment in Germany", *Headwater Control VI: Hydrology, Ecology and Water Resources in Headwaters*, IAHS Conference, Bergen 2005.

Kullback, S., *Information Theory and Statistic*, John Wiley, New York, 1959.

Kullback S. and R.A. Leibler, "On information and sufficiency", *Annals of Mathematical Statistics*, 22, 79- 86, MR39968, 1951.

Kron T.D. and D. Rosberg. "*An Artificial Neural Network Based Groundwater Flow and Transport Simulator*," http://www.isva.dtu.dk/staff/tdk/nn.2.pdf ,1998.

Lauren, J.G., R.J. Wagenet, J. Bouma, and J.H.M. Wiisten, "Variability of saturated hydraulic conductivity in a Glossaquic Hapludalf with macropores", *Soil Science*, 145: 20-28, 1988.

Lawrie, J., and J.Hearne1, "Reducing model complexity via output sensitivity", *Ecological Modelling*, 207(2–4):137–144, 2007.

Lee, S. I., and P. K. Kitanidis, "Optimization of monitoring well installation time and location during the aquifer contamination," *Water Resources Management*, 10: 439-462, 1996.

Leistra, M., A.M.A. van der Linden, J.J.T.I. Boesten, A.Tiktak, and F. van den Berg, "*PEARL model for pesticide behaviour and emissions in soil-plant systems. Description of processes,*" Alterra report 13, RIVM report 711401009, Alterra, Wageningen, 107 pp., 2001.

Lenda, A. and A. Zuber, "Tracer dispersion in groundwater experiments," *Proceedings of a Symposium on the Use of Isotopes in Hydrology*. I.A.E.A., Vienna, pp. 619– 641. March 9–13, 1970.

Li, L., C. A. Peters, and M. A. Celia, "Upscaling geochemical reaction rates using pore-scale network modeling," *Advances in Water Resources*, 29:1351–1370, 2006.

Ling, M., H. S. Rifai, J. J. Aziz, J. R. Gonzales, and J. M. Santillan, "Strategies and decision-support tools for optimizing long-term groundwater monitoring plans-MAROS 2.0," *Bioremediation Journal*, 8(3-4):190–128, 2004.

Liang, X., Z. Xie, and M. Huang, "A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model," *Journal of Geophysical Research,* 108(D16), 8613, 2003.

Loaiciga, H. A, R. J. Charbeneau, L. G. Everett, G. E. Fogg, B. F. Hobbs,, and S. Rouhani S., "Review of ground-water quality monitoring network design," *Journal of Hydraulic Engineering,* 118(1):11–37, 1992.

Looms, M.C., K.H. Jensen, A. Binley, and L. Nielsen, "Monitoring Unsaturated Flow and Transport Using Cross-Borehole Geophysical Methods," *Vadose Zone Journal*, 7(1): 227 – 237, 2008.

Malmstrom, M. E., G. Destouni, S.A. Banwart, and B.H.E. Stromberg, "Resolving the scale-dependence of mineral weathering rates," *Environmental Science and Technology*, 34:1375–1378, 2000.

Marquardt, D.W., "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal of Applied Mathematics,* 11:431-44, 1963.

Masket, S, Dibike, Y.B, Jonoski, A and D.P. Solomatine, "Groundwater Model Approximation With ANN for Selecting Optimum Pumping Strategy for Plume Removal," pp. 67-80 in O. Schleider and A. Zijderveld, (eds.), *AI Methods in Civil Engineering Applications, 2nd Joint Workshop on AI Methods in Civil Engineering Applications*, Cottbus, 2000.

McKay, M. D. "Evaluating prediction uncertainty", Nuclear Regulatory Commission and Los Alamos National Laboratory Technical Report NUREG/CR-6311. 1995

McLaren, R.G., P.A. Forsyth, E.A. Sudicky, J.E. Vanderkwaak, F.W. Schwartz, and J.H. Kessler, "Flow and Transport in Fractured Tuff at Yucca Mountain : Numerical Experiments on Fast Preferential Flow Mechanisms," *Journal of Contaminant Hydrology,* 43(3-4): 211-238, 2000.

McLaughlin D. "Recent Developments in Hydrologic Data Assimilation" *Reviews of Geophysics. Supplement*: 977-984, US National report to International Union of Geodesy and Geophysics, Paper 95RG00740, American Geophysical Union. Washington, D.C., 1995.

Meisel, W.S. and D.C. Collins, "Repro-modeling: An Approach to Efficient Model Utilization and Interpretation'" *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*: 349-358, 1973.

Melles, S.J., G.B.M. Heuvelink, C.J.W Twenhöfel, A. van Dijk, P. H. Hiemstra, O. Baume, and U. Stöhlker, ".Optimizing the spatial pattern of networks for monitoring radioactive releases," *Computers & Geosciences*, 37(3): 280–288, 2011.

Meyer, D., A. J. Valocchi, and J. W. Eheart, "Monitoring network design to provide initial detection of groundwater contamination," *Water Resources Research, 30*(9): 2647–2659, 1994.

Meyer, P.D., M.L. Rockhold, G.W. Gee, and T.J. Nicholson, *Uncertainty Analyses of Infiltration and Subsurface Flow and Transport for SDMP Sites*, NUREG/CR-6565, PNNL-11705. U.S. Nuclear Regulatory Commission. Washington, DC 20555-0001, 1997.

Minsker, B –Ed.," Long-Term Groundwater Monitoring The State of the Art," *American Society of Civil Engineers,* 2003.

Mohanty, B.P., R.S. Bowman, J.M.H. Hendrickx, and M.Th. van Genuchten, "New Piecewise-Continuous Hydraulic Functions for Modeling Preferential Flow in an Intermittent Flood-Irrigated Field," *Water Resources Research*, 33: 2049-2063, 1997.

Montas, H., R. Mohtar, A. Hassan, and F. AlKhal, "Heuristic space-time design of monitoring wells for contaminant plume characterization in stochastic flow fields," *Journal of Contaminant Hydrology*, 43(3– 4): 271– 301, 2000.

Mualem, Y., "New model for predicting unsaturated hydraulic conductivity of unsaturated porous media," *Water Resources Research*, 12: 513-522, 1976.

Nash, J. E. and J.V. Sutcliffe, "River flow forecasting through conceptual models, Part I - A discussion of principles," *Journal of Hydrology*, 10: 282–290, 1970.

Neuman, S.P., "Analysis of Pumping Test Data From Anisotropic Unconfined Aquifers Considering Delayed Gravity Response," *Water Resources Research*, 11: 329-342, 1975.

Neuman, Shlomo P. "Universal scaling of hydraulic conductivities and dispersivities in geologic media." *Water Resources Research* 26(8): 1749-1758, 1990

Neuman, S. P., C.L. Winters, and C.N. Newman, "Stochastic theory of field-scale Fickian dispersion in anisotropic porous media," *Water Resources Research*, 23(3): 453-466, 1987.

Neuman, S.P., P.J. Wierenga, and T.J. Nicholson. *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, 2003.

Nordquist, R, and C. I. Voss, "A simulation-based approach for designing effective field-sampling programs to evaluate contamination risk of groundwater supplies," *Hydrogeology Journal* 4: 23-46, 1996.

Nunes, L. M., M. C. Cunha, and L. Ribeiro, "Groundwater Monitoring Network Optimization with Redundancy Reduction," *Journal of Water Resources Planning and Management*, *130*(1), 33-43, 2004

Ohi, T., H. Takase, M. Inagaki, K. Oyamada, T. Sone, M. Mihara, T. Ebashi, and K. Nakajima, "*Application of a Comprehensive Sensitivity Analysis Method on the Safety Assessment of TRU Waste Disposal in Japan. Materials Research Society*," Paper #: 0985-NN03-21. 2010, Available at: http://www.mrs.org/s_mrs/sec_subscribe.asp?CID=7392&DID=195195&action=detail

Pachepsky, Y., T. Gish, A. Guber, A. Yakirevich, M. Kouznetsov, M. Van Genuchten, T. Nicholson, and R. Cady, *Application of Model Abstraction Techniques to Simulate Transport in Soils,* NUREG/CR-7026, U.S. Nuclear Regulatory Commission, Washington, DC 20555-0001, 2011.

Pachepsky, Y.A., W.J. Rawls, and D.J. Timlin, "A one-parameter relationship between unsaturated hydraulic conductivity and water retention," *Soil Science*, 165: 911-919, 2000.

Pachepsky, Y.A., and W.J. Rawls, W.J. (eds.), "Development of Pedotransfer Functions in Soil Hydrology," *Developments in Soil Science*, 30, Amsterdam, Elsevier, 2004.

Pachepsky, Y.A., A.K. Guber, M.T. Van Genuchten, T.J. Nicholson , R.E. Cady, J. Simunek, and M.G. Schaap, *Model Abstraction Techniques for Soil-Water Flow and Transpor*t, NUREG/CR-6884, U.S. Nuclear Regulatory Commission, Washington, DC 20555-0001, 2006.

Pachepsky, Y., W. J. Rawls, W. J., and D. Giménez, "Comparison of soil water retention at field and laboratory scales," *Soil Science Society of America Journal, 65*(2), 460-462, 2001.

Pachepsky, Y.A., D.J. Timlin, and W.J. Rawls, "Soil Water Retention as Related to Topographic Variables," *Soil Science Society of America Journal*, 65: 1787-1795, 2001.

Page, A.L., R.H. Miller, and D.R. Keeney (Eds.), *Methods of soil analysis. Part 2. Chemical and microbiological properties.* American Society of Agronomy, Inc. and Soil Science Society of America, Inc., Madison, Wis., 1982.

Poeter, E. P. and D.R. Anderson, "Multi-model ranking and inference in ground water modeling", *Ground Water,* 43, 597–605, 2005.

Polubarinova-Kochina, P.Y., "Theory of groundwater motion" Nauka, Moscow, (in Russ.),1977

Pruess, K., Oldenburg, C. and Moridis, G., "TOUGH2 User's Guide, Version 2.0, LBNL-43134, Lawrence Berkeley National Laboratory," 1999.

Pruess, K. "*A General-Purpose Numerical Simulator for Multiphase Fluid and Heat Flow*," Report LBL-29400, Lawrence Berkeley Laboratory, Berkeley, CA, 103 p., 1991.

Quevauviller, P., A.-M. Fouillac, J. Grath., and B. Ward, *Groundwater Monitoring*. John Wiley and Sons, 2009.

Raick, C., K. Soetaert, and M. Gregoire, "Model complexity and performance: How far can we simplify?" *Progress in Oceanography,* 70:27–57, 2006.

Rawls, W.J., D. Gimenez, and Grossman, R., "Use of soil texture, bulk density, and slope of the water retention curve to predict saturated hydraulic conductivity," *Transactions of the ASAE*, 41(4): 983-988, 1998.

Rawls, W.J., and Y.A. Pachepsky, "Using Field Topographic Descriptors to Estimate Soil Water Retention," *Soil Science*, 167: 423-235, 2002.

Rawls, W.J., D.L. Brakensiek, and K.E. Saxton, "Estimation of Soil Water Properties," *Transactions of ASAE* 25: 1316-1320, 1982.

Reed, P.M., and J.B. Kollat, "Save now, pay later? Multi-period many-objective groundwater monitoring design given systematic model errors and uncertainty," *Advances in Water Resources*, 35, 55–68, 2012.

Refsgaard, J. H., S. Christensen, T. O. Sonnenborg, D. Seifert, A.L. Højberg, and L. Troldborg, "Review of strategies for handling geological uncertainty in groundwater flow and transport modeling," *Advances in Water Resources* 36:36–50, 2012

Reilly, T.E, and A.W. Harbaugh, A.W., Guidelines *for Evaluating Ground-Water Flow Models.* Scientific Investigation Report. 2004-5038, 30 pp., US Geological Survey, 2004

Reynolds, W.D. and D.E. Elrick, "Constant Head Soil Core Method", *In*: Dane, J.H. and G.C. Topp - ed.), *Methods of Soil Analysis. Part 4. Physical Methods*, pp. 802-808, ASA and SSSA, Madison, WI, 2002.

Rivest, M D. Marcotte, and P. Pasquier, "Sparse data integration for the interpolation of concentration measurements using kriging in natural coordinates.", *Journal of Hydrology*, 416–417: 72–82, 2012.

Rojas, R., O. Batelaan, L. Feyen, and A., Dassargues, "Assessment of conceptual model uncertainty for the regional aquifer Pampa del Tamarugal – North Chile," *Hydrology Earth System Science*, 14:171–92, 2010.

Romano, N., and M. Palladino, "Prediction of soil water retention using soil physical data and terrain attributes," *Journal of Hydrology,* 265(1–4): 56–75, 2002.

Rose, M. R. and Harmsen, R., "Using sensitivity analysis to simplify ecosystem models: a case study," Simulation 31: 15-26, 1978.

Rubin, Y., and S.S. Hubbard (eds.), *"Hydrogeophysics,"* Springer, 2005.

Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Computer Physics Communications*, 181(2): 259-270, 2010.

Saltelli, A., M. Ratto, S. Tarantola, and F. Campolongo, F., "Sensitivity Analysis for Chemical Models," *Chemical Reviews*, 105: 2811-2827, 2005.

Schaap, M.G. and F. J. Leij, "Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model," *Soil Science Society of America Journal*, 64: 843–851, 2000.

Schulze-Makuch, D., D.A. Carlson, D.S. Cherkauer, and P. Malik, "Scale dependency of Hydraulic Conductivity in Heterogeneous Media," *Ground Water*, 37:904-919, 1999.

Shen, C. and M.S. Phanikumar, "A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling," *Advances in Water Resources,* 33: 1524-1541, 2010.

Shlomi, S., and A.M. Michalak., "A geostatistical framework for incorporating transport information in estimating the distribution of a groundwater contaminant plume," *Water Resources Research*., 43, W03412, , 2007.

Šimůnek, J., Jarvis, N.J., van Genuchten, M. Th., and Gardenas, A., "Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone," *Journal of Hydrology* 272: 14–35, 2003.

Šimůnek, J., M.Th. van Genuchten, and M. Šejna, "*The HYDRUS software package for simulating two- and three-dimensional movement of water, heat, and multiple solutes in variably saturated media*", Technical manual, version 1.0, PC Progress, Prague, Czech Republic. 2006.

Smettem, K.R.J., Oliver, Y.M., Pracilio, G., and R.J. Harper, "Data Availability and Scale in Hydrologic Applications," pp. 253-272, in*:* Pachepsky, Y. and Rawls, W.J. (eds.), *Development of Pedotransfer Functions in Soil Hydrology.* Amsterdam, Elsevier, 2004.

Stoppelenburg, F.J., K. Kovar, M.J.H. Pastoors, and A. Tiktak, "*Modeling the interactions between transient saturated and unsaturated groundwater. Off -line coupling of LGM and SWAP*," RIVM Rep. 500026001/2005. RIVM, Bilthoven, the Netherlands, 2005.

Storck, P., J. W. Eheart, and A. J. Valocchi, "A method for the optimal location of monitoring wells for detection of groundwater contamination in three-dimensional heterogeneous aquifers," *Water Resources Research,* 33(9):2081–2088, doi:10.1029/97WR01704, 1997.

Straface, S., F. Chidichimo, E. Rizzo, M. Riva, W. Barrash, A. Revil, M. Cardiff, A. Guadagnini, "Joint inversion of steady-state hydrologic and self-potential data for 3D hydraulic conductivity distribution at the Boise Hydrogeophysical Research Site," *Journal of Hydrology*, 407(1–4): 115–128, 2011.

Szucs, P., F. Civan, and M. Virag, "Applicability of the most frequent value method in groundwater modeling", *Hydrogeology Journal,* 14, 31–43, 2006.

Therrien, R., and E.A. Sudicky, "Three-Dimensional Analysis of Variably-Saturated Flow and Solute Transport in Discretely-Fractured Porous Media," *Journal of Contaminant Hydrology*, 23: 1-44, 1996.

Thoms, R.B., R.L. Johnson, and R.W. Healy, "*User's guide to the variably saturated flow (VSF) process for MODFLOW,*" Techniques and Methods 6-A18. Available at http://pubs.usgs.gov/tm/2006/tm6a18/ (verified 13 May 2009). USGS, Reston, VA, 2006.

Timlin, D.J., Y.A. Pachepsky, and C.L. Walthall. "A Mix of Scales: Topography, Point scales, and Yield Maps," pp. 227-242, in Y.A. Pachepsky, D.E. Radcliffe, and H.M. Selim (eds.), *Scaling Methods in Soil Physics,* CRC Press, Boca Raton, FL, 2003.

Tonkin, M. J., C. R. Tiedeman, D. M. Ely, M.C. Hill, "*OPR-PPR, a Computer Program for Assessing Data importance to Model Predictions Using Linear Statistics. Techniques and Methods*", Report TM –6E2 U.S. Geological Survey, Reston, Virginia, 2007.

Topp, G.C., J.L. Davis, and A.P. Annan, "Electromagnetic determination of soil water content: Measurements in coaxial transmission lines," *Water Resources Research* 16:574–582, 1980.

Tsai, F. T.-C. and X. Li, "Inverse groundwater modeling for hydraulic conductivity estimation using Bayesian model averaging and variance window", *Water Resources Research*, 44, W09434, 2008.

Tsai, F.T-C., N-Z. Sun, and W.W-G. Yeh, "Global-Local Optimization Methods for the Identification of Three-Dimensional Parameter Structure in Groundwater Modeling," *Water Resources Research*, 39, 1043, 2003.

Twarakavi, N.K.C., J. Šimůnek, and S. Seo, "Evaluating interactions between groundwater and vadose zone using the HYDRUS-based flow package for MODFLOW," *Vadose Zone Journal,* 7: 757–768, 2008.

Twarakavi, N.K.C., J. Šimůnek, and S. Seo, Reply to "Comment on 'Evaluating Interactions between Groundwater and Vadose Zone Using the HYDRUS-based Flow Package for MODFLOW" by Navin Kumar C. Twarakavi, Jirka Šimůnek, and Sophia Seo, *Vadose Zone Journal* 8: 820–821, 2009.

U.S. EPA, "*Roadmap to long-term monitoring optimization*", EPA 542-R-05-003, EPA/National Service Center for Environmental Publications, Cincinnati, OH 45242, 2005

U.S. DOE, "*Ground Water Surveillance Monitoring Implementation Guide for Use with DOE O 450.1, Environmental Protection Program*," DOE G 450.1-6. U.S. Department of Energy Washington, D.C. 20585, 2004.

U.S. EPA "Guidance for monitoring at hazardous waste sites: framework for monitoring plan development and implementation," OSWER directive no. 9355.4-28, U. S. EPA Office 2004.

Usunoff, E., J. Carrera, and S. Mousavi, "An approach to the design of experiments for discriminating among alternative conceptual models," *Advances in Water Resources*, 15:199-214, 1992.

Vauclin, M., J. Khanji, and G. Vachaud, G., "Experimental and numerical study of a transient, two-dimensional unsaturated-saturated water table recharge problem," *Water Resources Research,* 15: 1089-1101, 1979.

Valocchi, A. J. "Hydrogeochemical models," In: Kitanidis, P.K. and P.L. McCarty, Ed.. *Delivery and Mixing in the Subsurface: Processes and Design Principles for In Situ Remediation*. Springer, 2012.

Van Genuchten, M. Th., F. Kaveh, W. B. Russel, and S.R. Yates, S. R., "Direct and indirect methods for estimating the hydraulic properties of unsaturated soils," In: Bouma, J., A.K. Bregt (Eds.) *Land Quality in Space and Time*. PUDOC, Wageningen, 61-72, 1989.

Van Genuchten, M.Th., F.J. Leij, and L.J. Lund (eds.), "*Proceedings of the International Workshop on Indirect methods for Estimating the Hydraulic Properties of Unsaturated Soils*", University of California, Riverside, CA, 1992.

Van Genuchten, M.Th., F.J. Leij, and L.Wu (Eds.). "Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media," *Proceedings of the International Workshop*, Riverside, California, October, 22-24, 1997, University of California, Riverside, 1997.

Van Genuchten, M.Th., "A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils," *Soil Science Society of America Journal,* 44:892–898, 1980.

Van Ness, E.H., and M. Scheffer, "A Strategy to Improve the Contribution of Complex Simulation Models to Ecological Theory", *Ecological Modelling*, 185: 153-164, 2005.

Ventrella, D., Mohanty, B.P., Simunek, J. Losavio, N., and M.Th. van Genuchten, "Water and Chloride Transport in a Fine-Textured Soil: Field Experiments and Modeling," *Soil Science*, 165(8): 624-631, 2000.

Vepraskas, M. J., and J.P. Williams, "Hydraulic Conductivity of Saprolite as a Function of Sample Dimensions and Measurement Technique," *Soil Science Society of America Journal*, 59: 975-981, 1995.

Vereecken, H., M. Weynants, M. Javaux, Y. Pachepsky, M.G. Schaap, M.G., and M.Th. van Genuchten, "Using Pedotransfer Functions to Estimate the Van Genuchten–Mualem Soil Hydraulic Properties: A Review," *Vadose Zone Journal* 9: 795-820, 2010.

Volkova E., B. Iooss, and F. Van Dorpe, "Global study analysis for a numerical model of radionuclide migration from the RRC "Kurchatov Institute" radwaste disposal site,". *Stochastic Environmental Research and Risk Assessment*. 22(1):17-31, 2008.

Wang, C.G., and D.G. Jamieson, "An Objective Approach to Regional Wastewater Treatment Planning," *Water Resource Research*, 38(3): 1022, 2002.

Wang, W, Neuman, S. P., Yao, T., and P. J. Wierenga, "Simulation of Large-Scale Field Infiltration Experiments Using a Hierarchy of Models Based on Public, Generic, and Site Data," *Vadose Zone Journal,* 2: 297–312, 2003.

Ward, A.L., "*Vadose Zone Transport Field Study Summary Report*", PNNL-15443, U. S. DOE, 2006.

White, A. F., and S.L. Brantley, "The effect of time on the weathering of silicate minerals: why do weathering rates differ in the laboratory and field?" *Chemical Geology*, 202:479–506, 2003.

Willmot, C. J., "On the validation of models," *Physical Geography*, 2: 184–194, 1981.

Wilson, G.V., J.P. Gwo, P.M. Jardine, and R.J. Luxmoore, "Hydraulic and Nonequilibrium Effects on Multiregion Flow," pp. 37-61 in: H.M. Selim and L.Ma (eds.), *Physical Nonequilibrium in Soils; Modeling and Application*, Ann Arbor Press, Chelsea, MI, 1998.

Wösten, J. H. M., and M.Th. van Genuchten, "Using Texture and Other Soil Properties to Predict the Unsaturated Soil Hydraulic Functions," *Soil Science Society of America Journal,* 52: 1762-1770, 1988

Wu, Y. S., C. F. Ahlers, P. Fraser, A. Simmons, and K. Pruess, *Software Qualification of Selected TOUGH2 Modules*, Research Report, Earth Sciences Division, Lawrence Berkeley National Laboratory, LBL-39490, UC-800, October, 1996.

Wu, J., C. Zheng, C. C. Chien, "Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions," *Journal of Contaminant Hydrology*, 77: 41– 65, 2005.

Ye, M., S.P. Neuman, and P.D. Meyer, "Maximum Likelihood Bayesian Averaging of Spatial Variability Models in Unsaturated Fractured Tuff", *Water Resources Research*, VOL. 40, W05113, 2004.

Ye, M., P. D. Meyer, P.D., and S.P. Neuman, "On model selection criteria in multimodel analysis", *Water Resources Research,* 44, W03428, 2008.

Xu, X., G. Huang, H. Zhan, Z. Y. Que, and Q. Huang, "Integration of SWAP and MODFLOW-2000 for modeling groundwater dynamics in shallow water table areas," *Journal of Hydrology*, 412-413: 170-181, 2012.

Yakirevich, A., V. Borisov, S. Sorek, "A quasi three-dimensional model for flow and transport in unsaturated and saturated zones: 1. Implementation of the quasi two-dimensional case," *Advances in Water Resources*, 21: 679–689, 1998.

Yates, S. R., M.Th van Genuchten, A.W. Warrick, and F.J. Leij, "Analysis of measured, predicted, and estimated hydraulic conductivity using the RETC computer program," *Soil Science Society of America Journal*, 56: 347 – 354, 1992.

Yu, C., A.J. Zielen, J.- J Cheng, D.J. LePoire, E. Gnanapragasam, S. Kamboj, J. Arnish, A. Wallo, W.A. Williams, H. Peterson, *"User's Manual for RESRAD Version 6,"* Environmental Assessment Division. Argonne National Laboratory, Argonne, Illinois, 2001.

Zhang, Z. F., M. Oostrom, and A.L. Ward, "Saturation-Dependent Hydraulic Conductivity Anisotropy for Multifluid Systems in Porous Media," *Vadose Zone Journal*, 6: 925 – 934, 2007

Zhu, J. and B.P. Mohanty, "Spatial Averaging of Van Genuchten Hydraulic Parameters for Steady State Flow in Heterogeneous Soils: A Numerical Study," *Vadose Zone Journal*, 1: 261-272, 2002.

Zlotnik, V. A., B.R. Zurbuchen, T. Ptak, G. Teusch, "Support volume and scale effect in hydraulic conductivity: experimental aspects," in Zhang, D., and Winter, C. L., eds, *Theory, Modeling, and Field Investigations in Hydrogeology: A special Volume in Honor of Shlomo P. Neuman's 60th Birthday*: Boulder, Colorado, Geological Society of America Special Paper 348, p. 215-231.

Zobeck, T. M., N.R. Fausey, and N.S. Al-Hamdan, Effect of sample cross-sectional area on saturated Hydraulic Conductivity in two structured clay soils," *Transactions of the American Society of Agricultural Engineers*, 28: 791-794, 1985.

Zurmühl, T. "Capability Of Convection-Dispersion Transport Models To Predict Transient Water and Solute Movement in Undisturbed Soil Columns," *Journal of Contaminant Hydrology*, 30: 101-110, 1998.

Zurmühl, T., and W. Durner, "Modeling Transient Water and Solute Transport in a Biporous Soil," *Water Resources Research*, 32: 819-829, 1996.

APPENDIX A   REVIEWING ASSUMPTIONS MADE DURING DEVELOPMENT OF THE BASE MODEL ANDJUSTIFICATIONS FOR THOSE ASSUMPTIONS

<u>Subsurface structural units</u>
The vadose zone typically consists of various horizons and connected or disconnected lenses. The profile may contain massive, cross-bedded, and horizontally-bedded units, both poorly and well sorted. All of these structural units may conduct water differently, with chemical transport and transformation similarly varying among the different units. Even with significant advances recently in hydrogeophysical techniques, precise delineation of the various units remains a challenge in many practical situations. Borehole-based stratification often cannot provide the type of exhaustive coverage generally needed for the simulation domain. Measurements of pollutant concentrations are point based and depend upon the interaction of the measurement device (e.g. suction lysimeters) with the soil at different levels of the water content.

Because of great uncertainty in characterization, many often question the feasibility of performing vadose zone flow and transport simulations. While uncertainty definitely is an issue, such simulations are likely to become increasingly more popular, and essential, because of the importance of the vadose zone in terms of storing, transforming and redirecting pollutants entering this zone from the soil surface. In such simulations, the pertinent challenge is not to obtain a full-fledged three-dimensional description of the locations and properties of structural units. The relevant challenges are answers to such questions as:

- Does a specific site have subsurface structural units and features that may drastically change the fate and transport of pollutants in the vadose zone along the projected trajectory of a contaminant plume?
- If a restrictive fine-material layer is expected, does it have dikes or faults; can the layer have gaps?
- If a restrictive layer is expected, can it have a topography causing flow and transport via preferential pathways along it upper surface?
- Are there natural capillary barriers (i.e., boundaries between finer material overlaying the coarse sediments)? If yes, are gaps in these barriers expected?
- Can funnel flow in coarse-textural soils develop due to presence of a layer of coarse materials between two fine-textured layers?
- Can geochemical conditions of saturated or perched zones in the vadose zone cause changes in pollutant transformations or retention?
- Can well-conducting layers contain fine-scale high-conductivity parts of the pore space that will facilitate transport through large pores during episodic infiltration events?
- Is the lateral conductivity of the capillary fringe large enough to allow substantial contribution of the capillary fringe to lateral transport above the water table?

An essential condition here is to obtain answers to these questions from an available database that is as broad as possible. The database has to include information from public and private sources, cover both quantitative and qualitative (expert) information, and encompass both site-specific and generic information. The geological, hydrogeologic, and hydrogeophysical assessment could benefit greatly from complementary sources such as types of stream hydrographs, the presence and hydrology of springs, the presence of paleosoils, documented

fragipans, experience during well construction, occurrence and concentrations of agricultural chemicals in well water, and crop yield variability in dry years.

Dimension of the problem

Most vadose zone flow and transport simulations are currently carried out in only one dimension (i.e., only vertical transport is considered). This is not sufficient for sites where unsaturated flow and transport pathways may be altered substantially due to the presence of redirecting structural units mentioned in the previous section.

A two-dimensional representation is needed when:
- lateral transport can be substantial
- restrictive layers, if present, are well defined
- flow in the vadose zone is controlled by infiltration

A three-dimensional representation is needed when:
- the contaminant release source is providing also a substantial amount of infiltration
- substantial lateral transport in the vadose zone is expected, both along and across the main groundwater flow direction

Simulation domain, initial and boundary conditions

The simulation domain must be selected such that it reflects the dimension of the problem and allows one to set justifiable boundary conditions along all boundaries. One-dimensional (1D) problems require boundary conditions along the soil surface and the bottom of the vadose zone. Either a free drainage condition, or a zero pressure head condition at the ground water level (if the latter is monitored in time), is generally sufficient to simulate flow in this case.

1D simulations likely will not provide correct simulations if substantial transport in the capillary fringe is expected, or if a restrictive layer exist that not only can cause perched water to develop but also may contain gaps that allows rising groundwater to move upward through gaps in the restrictive layer. In such cases 2D or 3D simulations are necessary. Multi-dimensional vadose zone flow and transport simulations are best simulated when:

- the simulation domain includes both a vadose zone and an aquifer
- flow and transport need to be considered jointly, without assuming some rule of fluid and mass exchange between the vadose zone and groundwater

Boundary conditions in the groundwater subdomain have to be established according to existing guidelines for groundwater modeling. Boundary conditions for flow in the vadose zone must be established either from vadose zone monitoring data, or from preliminary 1D simulations of the vadose zone at the lateral boundaries of the flow domain. Alternatively, they can be defined in a more arbitrary fashion to provide continuity in the pressure head along a vertical direction from the water table to the soil surface. In the latter case, however, the boundaries should be placed far enough from the contaminant plume, so that the boundary conditions would not affect flow and transport within the domain of interest. This has to be verified by preliminary simulations using realistic values for the flow and transport parameters.

Daylight surface boundary conditions have to be:

- harmonized with the time step accepted in the model
- set with proper attention to runoff from and run-on to the soil surface simulation domain
- reflect water and chemical uptake

The representation of initial conditions depends on the purpose of the modeling. For exploratory and forecast purposes it is customary to run the model with realistic boundary conditions for relatively long simulation periods (e.g., up to one year), and to use the results of this run as initial conditions for the flow model. For parameter estimation and understanding the current situation at the field site, measured initial values of the vadose zone state variables should be used.

Estimated model parameters

Estimated parameter values must be obtained and reviewed whether or now calibration is applied. The estimated parameters for flow and transport either:

- are used directly in the simulations
- serve as initial estimates for model calibration
- are used in evaluations of the results of a calibration

A list of estimated flow and transport parameters for each subsurface structural unit typically includes:

- soil water retention parameters
- unsaturated soil hydraulic conductivity parameters
- molecular diffusion coefficient of the contaminant
- dispersivity
- dual-porosity parameters
- chemical transformation parameters
- biological transformation parameters

Additional parameters are used if the pollutant experiences colloid-facilitated transport, volatilization, or is represented by nanoparticles.

General requirements of the parameter estimation include:

- using several sources,
- matching scales
- correcting for field conditions
- defining uncertainty

Using several sources is recommended since a similarity in soil type (e.g., soil texture) by no means guarantees a comparable similarity in the flow and transport parameters. For example, soils of similar texture may exhibit up to two orders of magnitude differences in hydraulic conductivity. A compendium of literature flow and transport parameters in soils is given by Pachepsky and Rawls (2004).

A match in the scale of an experiment is important when selecting literature values of the hydraulic conductivity and dispersivity parameters since they are known to depend on the support or extent of the experiment in which they have been measured. A rule of thumb recommendation is to use parameters from experiments in which the support or extent is close to the vertical size of the cell used in numerical simulations.

A correction is often needed when relying on laboratory-measurements of the hydraulic conductivity, the dispersivity, and dual-porosity parameters since they depend on the spatial scale of an experiment. Methods to account for this change in scale can be found in both the soils (Pachepsky et al., 2003) and groundwater (Neuman, 1990, 1995) literature. Corrections to field conditions are more difficult for chemical and biological transformation because they are specific to the type of pollutant. However, the anticipated differences in adsorption, dissolution, and transformation rates have to be articulated and taken into account in uncertainty-based modeling projects.

Defining uncertainty in the estimated parameters is necessary for:

- performing multiple simulations to evaluate the prediction uncertainty if no calibration is envisaged
- determining the prior distributions of parameters to determine the posterior distributions of calibrated parameters if Bayesian methods are used in the calibrations
- evaluating the calibration results

Obtaining parameter estimates from multiple sources provides the necessary information for uncertainty characterization. It is imperative to use estimates not only of average values found in the literature, but also of the uncertainty in the average values that are reported in many (but not all) literature sources. An alternative method is to use probability distributions developed for soils from a large international database (Meyer et al., 1997).

Data available for calibration

Data needed for calibration of a vadose zone flow and transport model may include, in particular, monitoring data of:

- the soil water content
- the soil pressure head
- soil water fluxes
- concentrations of pollutants in soils
- tracer concentrations if a tracer test has been run
- hydrogeophysical data from cross-borehole monitoring of infiltration events
- hydrogeophysical data from surface monitoring of soil water contents
- groundwater levels
- concentrations of pollutant or tracers in groundwater

In spite of the fact that flow and transport processes in the vadose zone are slow, tracer tests are desirable for evaluating the possibility of rapid pollutant transport in dual-porosity soils. Calibration data must be available if the purpose of modeling is to analyze results of current or future management scenarios of contingencies. One may argue that running a model with all parameters within their variability ranges may provide an exhaustive characterization of the

uncertainty in predictions. This is true provided the structural units are conceptualized in a correct way, but that cannot be guaranteed in absolute terms. The credibility of a model increases if the calibration returns physically meaningful parameter values.

All vadose zone measurements are to some extent indirect, and are affected by interactions between the measuring device and the medium being studied. Therefore, all data used for calibration have to be reported with expertly or directly estimated errors. This is needed to evaluate the calibration results in terms of their accuracy.

The type of data, number of data points, and the frequency of data collection for reliable calibration is site-specific and depends on the method of calibration (see below). An optimal methodology of defining the best monitoring strategy for calibrating a vadose zone model presents an avenue of future research.

Calibration procedure and results

Two calibration methods are used most often in vadose zone flow and transport modeling: trial-and-error (manual) calibration and automated calibration. Automated calibration does not change locations and dimensions of structural units; it only varies material properties within these units to match simulated and measured values. Trial-and-error calibrations usually change both the locations and the dimensions of the structural units until a better fit is obtained.

An advantage of automated calibration is the objectivity in which the parameters are obtained. A disadvantage is that it may end up supporting an incorrect conceptualization of the flow and transport processes embedded in the model. Trial-and-error methods mirror automated calibration in terms of these advantages and disadvantages. If automated calibration is used, the statistics of parameters should be analyzed in the same way as recommended in groundwater modeling projects (Hill and Tiedeman, 2007).

Vadose zone simulations are notoriously slow because of the nonlinearity in the governing flow equations. Therefore, automated calibration is usually implemented only for one-dimensional simulations. Calibration in two- and three-dimensional simulations is usually achieved using trial-and-error process.

The calibration process should include:

- normalization of measurements to exclude the effect of the measurement unit on the calibration results
- justified removal of non-sensitive parameters from calibration
- assignment of different and explicable weights to measurements of different types

Software properties

Because of the nonlinearities involved, numerical solutions of coupled vadose zone and groundwater problems often show unstable behavior, or produce unacceptable errors in the mass balances for water and solutes. This behavior, including the software used for the simulations, has to be documented and reported.

Model documentation

Reviewing the documentation of a model is an essential component of the base model review. Reilly and Harbaugh (2004) indicate in their guidelines for evaluation of groundwater models that "because models are embodiments of scientific hypotheses, a clear and complete documentation of the model development is required for individuals to understand the hypotheses, to understand the methods used to represent the actual system with a mathematical counterpart, and to determine if the model is sufficiently accurate for the objectives of the investigation". The same is true for vadose zone modeling. Clarifications on all of the topics addressed in this appendix have to be included in the documentation.

APPENDIX B   THE QUASI-3D CODE BENCHMARKING

B.1 Model Evaluation Procedure

The QUASI-3D model described in section 4.3.1 was tested and benchmarked against the corresponding full numerical code, FULL-3D. At a first stage, the performance of the FULL-3D and the QUASI-3D codes were assessed against measured and simulated water flow through a sandy soil using the laboratory setup of Vauclin (1979). This experiment was simulated earlier with the VSF model (Thoms et al., 2006) and the HYDRUS-1D-MODFLOW model (Twarakavi et al., 2008). At a second stage, the performance of FULL-3D and QUASI-3D was compared against simulations with synthetic data of 3D problems addressing infiltration, pumping and groundwater mound dissipation through soils with different hydraulic properties. Benchmarking against available analytical solutions was also performed. Comparisons were made with reference to the following metrics.

- The maximum difference between simulated groundwater levels with the QUASI-3D (Z_{Q3D}) and the FULL3D (Z_{F3D}) codes, at the end of the simulation period:

$$dZ_{\max} = \max(Z_{Q3D} - Z_{F3D}) \tag{B1}$$

- The maximum relative difference between simulated groundwater levels at the end of the simulation period

$$dZ_{\max}^{rel} = \max[(Z_{Q3D} - Z_{F3D})/Z_{F3D}] \tag{B2}$$

- Mass balance error, M_{err}, assessed by
$$M_{err} = (W_{fin} - W_{ini} - Q)/Q_s *100\% \tag{B3}$$

     where W_{ini} and W_{fin} are the initial and final water storage in the investigated domain, Q is the difference between total amounts of injected and pumped water, and Q_s is the sum of the absolute values of injected and pumped water volume.
- The number of time steps, N_t, required for the simulation.
- The total computation time, CPU.
- The average number of iterations (accounting for the nonlinearity) within a time step, $n_{iter.}$ This refers to the total number of iterations N_{iter} during a specific simulation divided by the number of time steps needed for this simulation. It characterizes the computations load needed to build the linearized matrix of the algebraic equations.
- The average number of iterations per solution of the linear system of equations at one iteration, $n_{lin,}$ commencing from the original nonlinear system. This addresses the total number of iterations needed to solve the linear system of (8), for which N_{lin}, is divided by the total number of iteration $N_{iter.}$ This characterizes the proportion of computations needed to solve the linearized system of the FD equations, (8), using the PCG2-MODFLOW routines based on the modified incomplete Cholesky Preconditioned **C**onjugate Gradients method (Hill, 1990)

B.1.1 2-D Variably Saturated Transient Infiltration Observed in Laboratory Experiments

Results obtained with several models are compared in this section, including against observed laboratory data by Vauclin et al. (1979), The laboratory experiment was conducted using a 6 m long, 2 m high box was filled with sandy soil. Two drainage outlets were set at the box walls at 0.65 m from its bottom. Each outlet was connected to a reservoir in which a level of 0.65 m above the bottom of the box was maintained during the entire experiment. At the beginning, the experimental system was maintained in hydraulic equilibrium with soil in full saturation at and below the levels of the drainage outlets. Water was applied for 8 hours at a 2x1 m^2 strip in the center of the soil surface at a rate Q of 3.55 m day^{-1}. The remainder of the soil surface was sealed to prevent evaporation. The simulation was carried out using the following soil hydraulic parameters: θ_s=0.30 m^3 m^{-3}, θ_r = 0.01 m^3 m^{-3}, α= 3. 3 m^{-1}, n = 4.1, and K_s=8.50 $_m$ day^{-1}.

To compare the performance of the FULL-3D code with VSF, The ARS team followed the same spatial discretization cited in Thoms et al. (2006), i.e. a grid spacing of 0.1 m laterally and 0.06 m vertically. The time step was 1 min. A domain of 3 m by 2 m was considered due to symmetry of the flow process.

Results of the water table simulations with the VSF and FULL-3D codes (both were developed by modifying MODFLOW) are compared in Fig. B1a, which also shows the experimentally measured water table. Simulation results for VSF and FULL-3D codes coincides and yielded an acceptable fit with the measured data. This provided confidence about the FULL-3D code to serve as appropriate benchmark for evaluating the performance of the QUASI-3D code. Fig. B1b shows results of the water table simulations conducted with two quasi 3D models: HYDRUS-1D - MODFLOW (Twarakavi et al. 2008) and our QUASI-3D code which did not consider the capillary fringe. Both codes underestimates the water table depth. The most significant discrepancies occurred below the water application area. HYDRUS-1D - MODFLOW created a smoother water table profile as compared with QUASI-3D. The latter produced a flatter groundwater table under the irrigated area and a sharp decrease in the water level outside of the irrigated area. Differences between the two codes relative to the groundwater level are attributed to the constant specific yield term in HYDRUS-1D - MODFLOW, which appears to be a fitting parameter. Using the constant specific yield provides a smoothing of the solution close to the infiltration source. Unlike HYDRUS-MODFLOW, the specific yield is not used in QUASI-3D where the water table is considered to be a moving boundary with its position estimated by the coupled solution of the Laplace equation in the saturated domain and a series of 1D equations for vertical flow in the unsaturated zone. Xu et al. (2012) simulated the same experiment using the SWAP-MODFLOW-2000 package. Results showed that the simulated water table elevations closely matched the observed levels, except during the first 2 hours when they were slightly higher than the observed levels, probably due to neglecting lateral flow in the unsaturated zone.

The effect of introducing a capillary fringe is shown in Fig B1c. The increase in the thickness of the capillary fringe (a decrease in the value of h_{cf}) leads to smaller differences between simulations with QUASI-3D and the observed elevation of the water table. Decreasing h_{cf} to -0.5 m, improves the agreement between the QUASI-3D and FULL-3D simulations. The position of the specific surface was not affected by h_{cf} for values of h_{cf}<-1 m. At the beginning of the simulation, the water table rise in the center of the box strongly depended upon lateral flow in the unsaturated zone. At later stages, accounting for the capillary fringe with h_{cf} < -0.5 m provided satisfactory correspondence between the benchmark solution using FULL-3D and the QUASI-3D codes across the entire simulation domain.
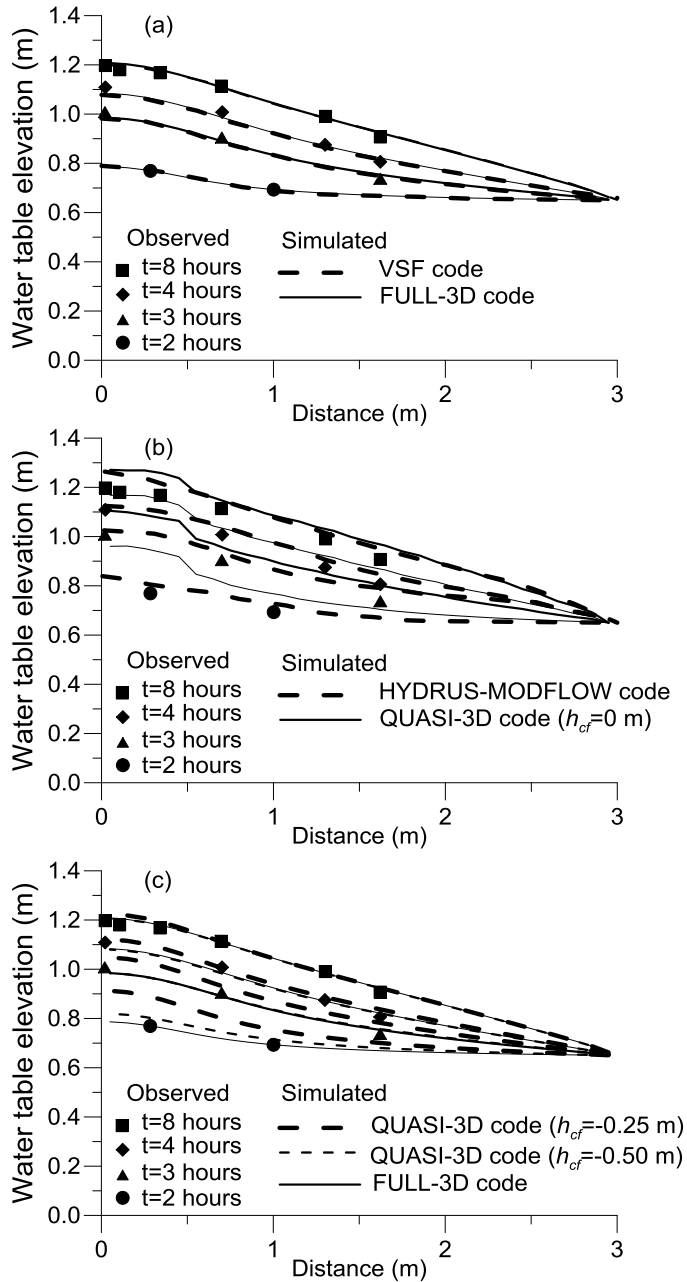
Figure B-1 A comparison of water table elevation calculated using the (a) VSF and FULL-3D codes, (b) HYDRUS-MODFLOW and QUASI-3D (with h_{cf}=0 m) codes, (c) QUASI-3D code (with h_{cf}=-0.25 m and -0.50 m), with the experimental data of Vauclin et al. (1979).

B.1.2 3-D Transient Infiltration, Pumping and Groundwater Mound Dissipation Simulations.

A comparison of the FULL-3D and QUASI-3D codes was carried out for simulations of infiltration/pumping events with synthetic data. The flow domain consisted of a rectangular parallelepiped with lengths L in both horizontal directions, and thickness D (Fig. B2). Simulations were performed for two sets of hydraulic parameters for a homogeneous porous medium at three scales: small (L=100 m, D=1 m), medium (L=1000 m, D=10 m), and large (L=10,000 m, D=100 m). The initial water table depth was D/2. A uniform grid was introduced in the simulation domain. The number of grid points was 100 in each direction, with the total number of nodes being 10^6. Hydraulic properties of the three sediments (sand, sandy loam and silt loam), expressed as in terms of the van Genuchten-Mualem model (Eqs. 4-3, 4-4), are presented in Table B1.

Four simulation setups were considered.

Simulation setup 1: Infiltration from the soil surface. Infiltration from square areas at the soil surface (10x10 m, 100x100 m, and 1000x1000 m for the small, medium and large spatial scales, respectively) located at the center of the surface domain would cause the water tables to rise. The bottom and lateral boundaries of the domain were subjected to no-flow conditions. An infiltration rate of 0.01 m day^{-1} during 15, 150 and 1500 days were ascribed to the spatial scales. Since simulations were performed for the aquifer composed of sandy material (Table B1), no capillary fringe was considered.

Simulation setup 2: Groundwater pumping. A vertical well was located in the center of the simulation domain, with the well screen reaching the bottom of the aquifer bottom. The lengths were 0.2, 2.0 and 20 m for the small, medium and large scales, respectively. Constant pumping rates of 0.5, 50, and 5000 m^{3}d^{-1} during 15, 150 and 1500 days were assumed for the three space scales, respectively. All boundaries were subject no-flow conditions. Similar simulations with the same soil (Table B1) were also performed for the case of having constant head at the lateral boundary of the domain. Results for this latter case could be compared with the analytical solution of Neuman (1975) for pumping from an isotropic unconfined aquifer. Again, no capillary fringe was considered.

Simulation setup 3: Sequential infiltration, redistribution (groundwater mound dissipation), and pumping. Infiltration over a square area 100x100 m at a rate of 0.01 m day^{-1} during 365 days, was followed by a pause of 365 days allowing redistribution of the water (all boundary fluxes and internal sinks were set to zero), after which pumping occurred during 365 days at a rate of 50 m^{3}d^{-1}. Simulations were performed for the medium scale with parameters for sediment 2 (Table B1) corresponding to sandy loam, and for two h_{cf} values (0 and -0.15 m).

Simulation setup 4: Groundwater mound dissipation
The mathematical formulation and analytical solution of the problem are presented in Poubarinova-Kochina (1977). The groundwater level at the initial time, t_0, is given as a rectangular parallelepiped (height Δh) on top of flat surface (elevation h_0)

$$h(x, y, t_0) = \begin{cases} h_0 \text{ if } |x| > R_1, \ |y| > R_2 \\ h_0 + \Delta h \text{ if } |x| \leq R_1, \ |y| \leq R_2 \end{cases} \tag{B4}$$
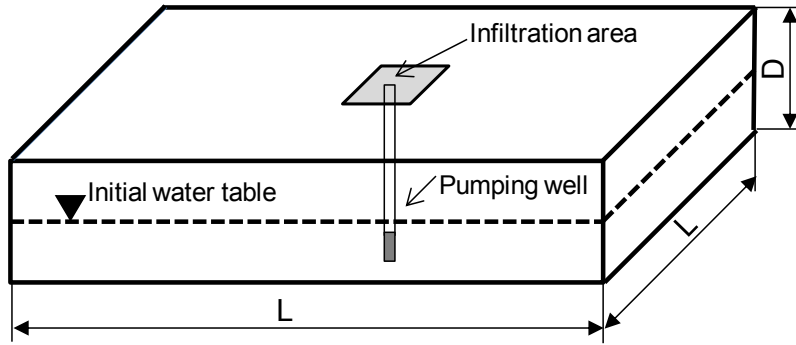
Figure B-2 A sketch of the domain for synthetic simulations of infiltration/pumping stresses.

Table B-1 Hydraulic properties of porous media used in numerical experiments

Setup	Texture	θ_r	θ_s	α, (m^{-1})	n	K_s, (m/day)
1 and 2	Sand	0.070	0.400	0.145	2.68	5.0
3	Sandy loam	0.065	0.410	0.075	1.89	1.06
4	Silt loam	0.131	0.369	0.423	2.06	0.0496

where R_1 and R_2 are initial extensions of the groundwater mound in *x* and *y* directions, respectively. The analytical solution of the linearized 2-D Boussinesq equation for this case is given by

$$h(x,t) = h_0 + \frac{\Delta h}{4}\left(\text{erf}\,\frac{R_1 - x}{2\sqrt{at}} + \text{erf}\,\frac{R_1 + x}{2\sqrt{at}}\right)\left(\text{erf}\,\frac{R_2 - y}{2\sqrt{at}} + \text{erf}\,\frac{R_2 + y}{2\sqrt{at}}\right) \quad\quad \text{(B5)}$$

where $a = K_s\tilde{h}/(S_y)$, S_y is the specific yield and $\tilde{h} = h_0 + \Delta h/2$ is the average groundwater level.

The ARS study team carried out simulations for a silty loam soil (Table B1) at the medium spatial scale. The size of the initial groundwater mound was: $R_1 = R_2 = 55$m, $\Delta h = 1$ m. Specific yield for the analytical solution was taken to be $S_y = 0.136$.

Fig. B3 compares results obtained with the QUASI-3D and FULL-3D models for the infiltration example (simulation setup 1). While Table B2 and Fig. B3 compares results for simulation setups 1 (infiltration) and 2 (pumping). Overall, results of the simulated water table elevation with the FULL-3D and QUASI-3D codes are very close, showing a maximum relative difference (dZ_{max}^{rel}) that did not exceed the value of 0.004 for the infiltration simulations and 0.053 for pumping simulations. The maximum relative difference in groundwater levels for the two codes was larger for the small-scale experiments as compared to the larger domain.

In the infiltration simulations, the QUASI-3D code formed a slightly taller and steeper groundwater mound as compared to the one simulated by the FULL-3D at all scales (Figure B3a, b, and c). The largest difference (dZ_{max}) in groundwater levels simulated by the two codes was 0.02 m at the large scale, while the smallest $dZ_{max} = 0.002$ m was obtained at the small scale.

For the second setup (pumping) with no flow boundary conditions along the lateral boundaries, the cone of depression simulated with QUASI-3D code was shallower than that simulated with FULL-3D (Fig. B3a, b, and c). The largest absolute difference (dZ_{max}) between the groundwater levels obtained with the two codes was 0.529 m for the larger scale. The smallest value at $dZ_{max} = 0.019$ m was for the small-scale simulated pumping with a constant head boundary condition at the lateral boundary. This exhibited practically identical solutions for the QUASI-3D and FULL-3D codes. The absolute maximum deviation in groundwater levels was 0.0016 m. Comparison of the QUASI-3D solution with the analytical solution of Neuman (1975), assuming a value of 0.3 for the specific yield of sand, provided a very good agreement between the simulated groundwater levels (not further shown here). Maximum absolute errors in the vicinity of the pumping well were 0.18, 0.08 and 0.002 m for the large, medium and small scales, respectively.

The mass balance of all simulations was satisfactory for all soil textures. For both models the mass balance errors did not exceed values of 0.21%, 0.06% and 0.04% for the small, medium and large scales, respectively.

Results of simulations for setup 3 (sequential infiltration, mound dissipation, and pumping) are presented in Fig. B4 and Table B2. Simulations with the FULL-3D and QUASI-3D codes for

groundwater levels with h_{cf}=-0.15 m were practically identical, with the maximum difference being less than 0.1 m. The maximum difference obtained for groundwater levels between the FULL-3D and QUASI-3D codes without a capillary fringe (h_{cf}=0) varied from 0.07 m during infiltration to 0.7 m during pumping.
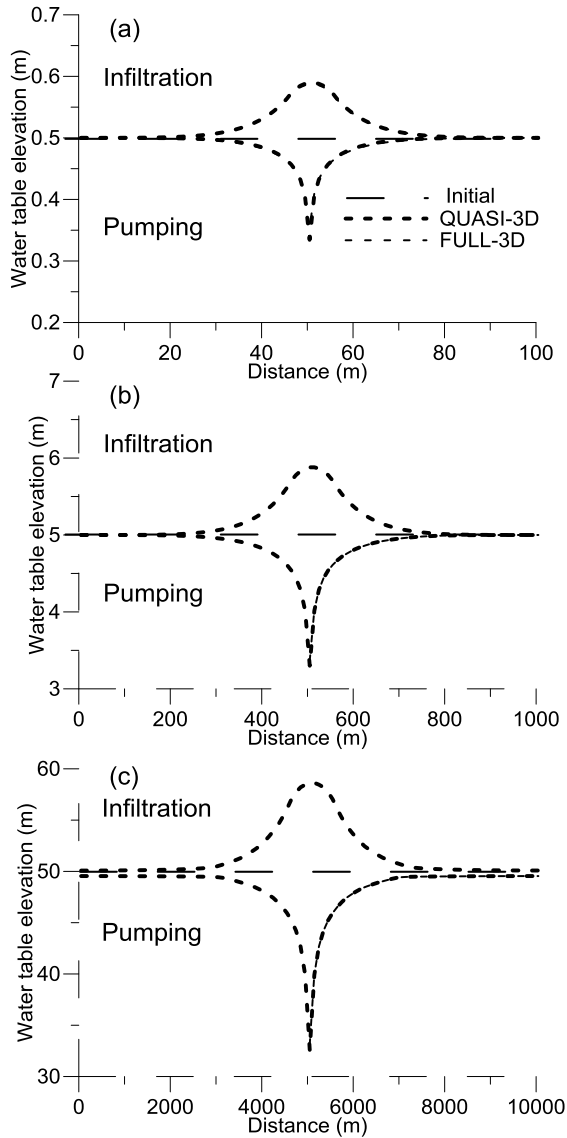

Figure B-3 Results of simulations of examples considering infiltration (setup 1) and pumping (setup 2) for (a) small, (b) middle and (c) large scales, with the QUASI-3D and FULL-3D codes.

The mass balance error was less than 0.01% in all simulations..

Results of the simulations of groundwater mound dissipation (setup 4) are shown in Fig. B5 and Table B2. The simulation with QUASI-3D for a groundwater level of hcf=0, was actually closer to the analytical solution of the linearized Boussinesq equation as compared to the FULL-3D solution. This is due to the fact that the QUASI-3D code accounts both for the specific features of the variably-saturated flow model and for the limitations of the Dupuit assumption. The maximum difference was 0.042 m at the groundwater mound summit, where QUASI-3D predicted higher groundwater levels than FULL-3D which did not account for horizontal flow in the capillary fringe. Introducing the capillary fringe into QUASI-3D with a value of -3.4 for hcf provided very good agreement between the QUASI-3D and FULL-3D models. The maximum deviation of the simulated groundwater levels was only 0.004 m.

B.2 Computational Efficiency Of The Quasi-3d Code

The numerical efficiency was expressed in terms of CPU time, the number of time steps N_t, the number of nonlinear iterations N_{iter}, and the number of iterations to solve the linearized set of finite difference (FD) equations N_{lin} (Table B2). In all simulations for setups 1 and 2 (infiltration and pumping, respectively), the CPU time for QUASI-3D code less than that of the FULL-3D code. However, only in two cases out of six simulations was the computational efficiency of QUASI-3D significantly higher (up to 300%) than that of the FULL-3D. For the infiltration simulations at the medium and large scales, the number of time steps and the number of nonlinear iterations were larger for QUASI-3D as compared FULL-3D (Table B2). For all other simulations, N_t and N_{iter} were approximately the same for both codes, which means a very similar rate of convergence (n_{iter}) when solving the set of non-linear FD equations. Yet, for setup 1 (infiltration), convergence was slower for the QUASI-3D code in comparison to FULL-3D, while the opposite was true for the setup 2 (pumping) simulations (Table B2). The number of iterations to solve the linearized set of the FD equations (N_{lin}) using the PCG2-MODFLOW code was smaller for QUASI-3D as compared to the FULL-3D code at the small scale. The number of iterations for QUADI-3D were, however, larger for the medium and large scales.

For setup 3, the computations with QUASI-3D in terms of CPU time and the number of iterations were about 30% faster than those of the FULL-3D code. The better numerical efficiency of the QUASI-3D code was not only due to a decrease in the computations to form the matrix of FD equations and its solution, but also because of a lower total number of iterations by about 30%. For setup 4, the computational efficiency of QUASI-3D was only slightly better than that of FULL-3D.

The total computation time was partitioned between (1) the time needed to compose the FD matrix equations, including calculations of the non-linear soil hydraulic properties and various terms in the FD equations, and (2) the time needed to solve the final set of matrix equations. To assess this partitioning, the ARS study team built a linear regression relation for the total CPU time (T_{CPU}) as a function of total number of nonlinear iterations (N_{iter}) and total number of iterations to solve the linearized set of equations using PCG2-MODFLOW (N_{lin}):

$$T_{CPU} = m_{iter} N_{iter} + m_{lin} N_{lin} \tag{B6}$$

Table B-2 Performance of the benchmark of FULL-3D and QUASI-3D models for the setups 1, 2[*], 3 and 4.

Scale	Setup	Model	M_{err} %	CPU min	N_t	N_{iter}	N_{lin}	dZ_{max} m	dZ_{max}^{rel}
Small	1. Infiltration	QUASI-3D	0.06	3.29	138	311	359	0.002	0.004
		FULL-3D	0.05	4.02	136	290	523		
	2. Pumping	QUASI-3D	0.21	3.11	136	272	407	0.019	0.053
		FULL-3D	0.05	3.39	136	272	408		
Medium	1. Infiltration	QUASI-3D	0.06	20.0	543	1743	6199	0.006	0.001
		FULL-3D	0.02	23.33	535	1735	5028		
	2. Pumping	QUASI-3D	0.04	6.58	183	486	3224	0.067	0.020
		FULL-3D	0.00	22.07	378	1315	7395		
Large	1. Infiltration	QUASI-3D	0.01	75.12	1220	4984	43388	0.020	0.000
		FULL-3D	0.02	78.15	1094	4329	29408		
	2. Pumping	QUASI-3D	0.00	24.5	521	1454	15585	0.529	0.016
		FULL-3D	0.04	63.05	887	2788	29387		
Medium	3.Infiltration, pause, pumping	QUASI-3D h_{cf}=0 m	0.01	27.16	694	2450	6824	0.7	0.029
		QUASI-3D h_{cf}=-0.15 m	0.01	27.35	697	2493	6535	0.01	0.002
		FULL-3D	0.01	43.11	930	3342	8151		
Medium	4.Groundwater mound dissipation	QUASI-3D h_{cf}=0 m	0.00	13.41	603	1206	1635	0.042	0.007
		QUASI-3D h_{cf}=-3.4 m	0.00	14.43	603	1206	1639	0.004	0
		FULL-3D	0.00	15.14	603	1206	1640		

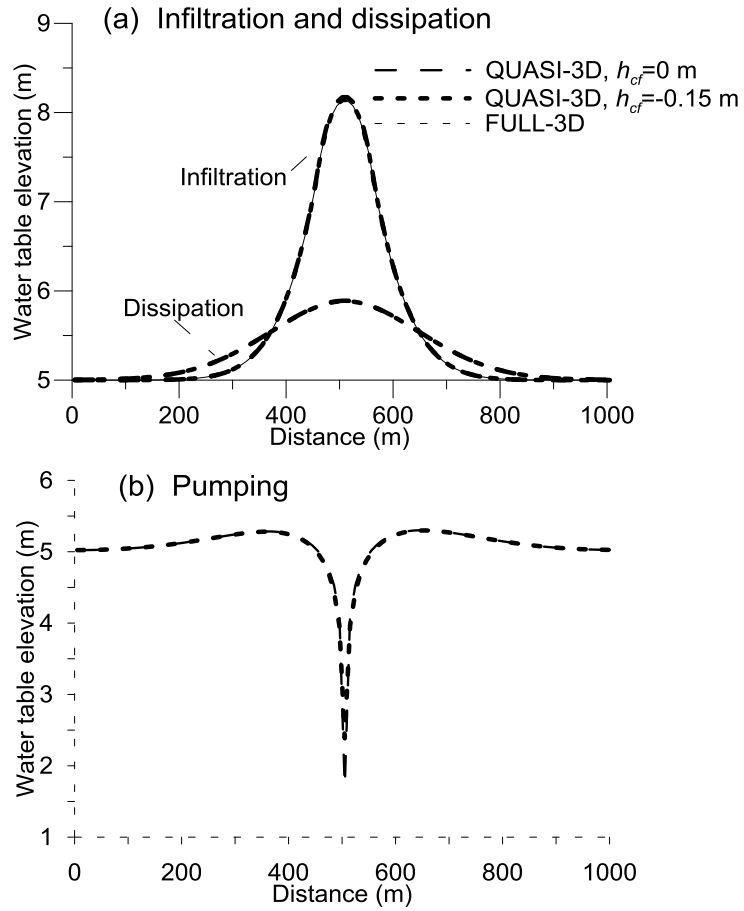[*]Pumping simulations with no flow boundary condition at the lateral boundary.

Figure B-4 Results of simulations of the setup 3 example: (a) infiltration and mound dissipation and (b) pumping, with the QUASI-3D and FULL-3D codes.
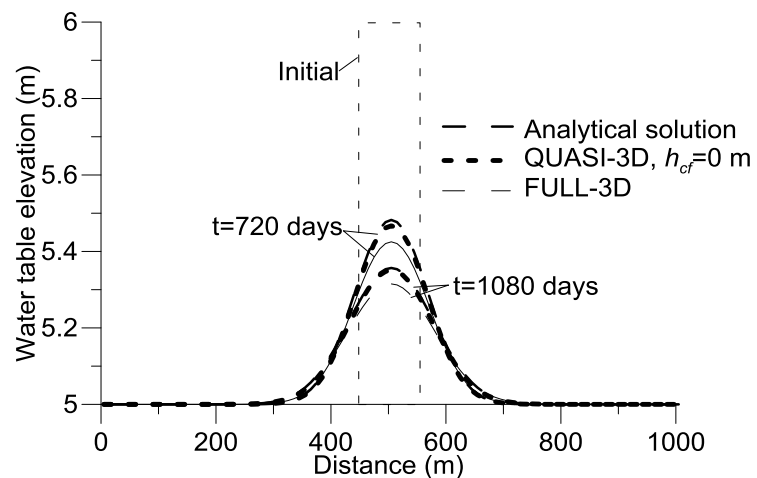
Figure B-5 Results of simulations with the analytical solution, QUASI-3D and FULL-3D codes of groundwater mound dissipation, for the setup 4 example.

where m_{iter} and m_{lin} are the specific CPU times required to build matrices and to solve the linearized set of equations for one iteration, respectively.

The ARS study team used the results presented in Table B2 to obtain estimated of the coefficients in Eq. (B6). Table B3 presents parameters of the linear regression. For setups 1 and 2, when using the QUASI-3D code, the number of nodes for the FD grid associated with 1D flow in the unsaturated zone was about half of the total number of nodes. When using FULL-3D, all FD grid nodes were associated with 3D flow process. The CPU time of the QUASI-3D code for one linear iteration (i.e., to solve linearized FD equation set using the PCG2-MODFLOW code) was found to be almost half of that of FULL-3D for both infiltration and pumping. This because of the significantly smaller number of nodes associated with the 3D flow process. Note that the computational expense to perform one iteration for solving the linear set of equations with positive definite matrix PGG2, is proportional to the total number of nodes. Most of the CPU time was spent to compute matrices of the FD equations, while the time required to solve this set of equations depended on the rate of convergence speed (10-40% of total CPU time).

Table B-3 Linear regression parameters of (B6) to assess time performance efficiency of the FULL-3D and QUASI-3D codes.

Code	QUASI-3D		FULL-3D	
Coefficients	m_{iter} min	m_{lin} min	m_{iter} min	m_{lin} min
Values	0.0086	0.00073	0.0097	0.0013
R^2	0.999		0.999	

B.3 Effect of Capillary Fringe

The QUASI-3D code was found to be efficient and accurate in comparisons with the benchmark FULL-3D code when the capillary fringe was considered. Numerical experiments presented in the sections 4.1 and 4.2 showed that horizontal flow of water in the unsaturated zone can have a noticeable effect on groundwater level in the vicinity of sources and sinks. Assigning a value of the pressure head (hcf) to define the position of the specific surface representing the capillary fringe, which is accounted for by a moving boundary between 1D flow in the unsaturated zone and 3D flow in groundwater, allows one to expand the zone of 3D flow to the capillary fringe and increase the accuracy of groundwater table simulations. The value of hcf can be estimated from the soil water curve or by comparing the QUASI-3D solution with the more accurate FULL-3D solution for prescribed conditions. The effect of the capillary fringe on water flow and solute transport is controversial. Experiments by Abit et al. (2008) showed that when Br was introduced in the capillary fringe, it was moving laterally above the water table. The role of the capillary fringe should depend on dependency of the hydraulic conductivity on the pressure head near saturation. The pore-connectivity parameter l in the van Genuchten-Mualem model (Eq. 4-4) and properties of the soil water retention curve both affect this dependency. The parameter l (median value of 0.5 derived by Mualem, 1976) describes the tortuosity of the pore space and can significantly affect the rate at which the hydraulic conductivity decreases with the pressure head. However, values of l from a range of -10 to 10 have been encountered (van Genuchten et al. 1989). Yates et al. (1992) and Schaap and Leij (2000) reported an even wider range of l values. When this

parameter is negative and has a large absolute value, the hydraulic conductivity of unsaturated soils in the capillary fringe remain close to the value of the saturated hydraulic conductivity as soil dries out (Pachepsky et al., 2000). On the other hand, if the parameter l is a large and positive, the hydraulic conductivity decreases fast as soil loses water, which happens most often in fine-textured soils or in soils with well-developed macroporosity (Wösten and van Genuchten, 1988). The effect of the parameter l on the efficiency of accounting for the capillary fringe presents an interesting avenue for further research.

The 1D approximation for flow in the unsaturated zone to estimating risks of groundwater contamination has been employed for various contaminants, e.g. PEARL (Leistra et al, 2001) for pesticides and RESRAD for radionuclides (Yu et al., 2001). No capillary fringe is considered in these codes. The QUASI-3D code may be used to assess potential errors caused by ignoring the capillary fringe. The code may be useful also for assessing the role of vertical heterogeneities that can cause accumulation of perched water. Approximation with and without a capillary fringe can also be applied to the QUASI-3D and with FULL-3D codes. This can facilitate further evaluation of the conceptual model behind the QUASI-3D approximations.

NRC FORM 335 (12-2010) NRCMD 3.7	U.S. NUCLEAR REGULATORY COMMISSION	1. REPORT NUMBER (Assigned by NRC, Add Vol., Supp., Rev., and Addendum Numbers, if any.)
	BIBLIOGRAPHIC DATA SHEET (See instructions on the reverse)	NUREG/CR-7221

2. TITLE AND SUBTITLE	3. DATE REPORT PUBLISHED	
Integrating Model Abstraction into Subsurface Monitoring Strategies	MONTH **March**	YEAR **2017**
	4. FIN OR GRANT NUMBER	

5. AUTHOR(S)	6. TYPE OF REPORT
Y. Pachepsky[1], A. Guber[1], A. Yakirevich[2], F. Pan[1], T. Gish[1], M. Kouznetsov[2], M. Van Genuchten[3], R. Cady[4], T. Nicholson[4]	Technical
	7. PERIOD COVERED (Inclusive Dates)

8. PERFORMING ORGANIZATION - NAME AND ADDRESS (If NRC, provide Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address; if contractor, provide name and mailing address.)

[1]United States Department of Agriculture Agricultural Research Service Environmental Microbial and Food Safety Laboratory Hydrology and Remote Sensing Laboratory Beltsville, MD 20705

[2]Department of Environmental Hydrology & Microbiology Zuckerberg Institute for Water Research Blaustein Institutes for Desert Research Ben-Gurion University of the Negev Sede Boqer Campus, 84990, Israel

[3]Department of Mechanical Engineering, COPPE/LTTC Federal University of Rio de Janeiro, UFRJ Rio de Janeiro, RJ CEP 21945-970, Brazil

[4]Division of Risk Analysis Office of Nuclear Regulatory Research U.S.Nuclear Regulatory Commission Washington, DC 20555

9. SPONSORING ORGANIZATION - NAME AND ADDRESS (If NRC, type "Same as above", if contractor, provide NRC Division, Office or Region, U. S. Nuclear Regulatory Commission, and mailing address.)

Division of Risk Analysis
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, DC 20555

10. SUPPLEMENTARY NOTES
T.J. Nicholson, NRC Project Manager

11. ABSTRACT (200 words or less)

NUREG/CR-7221 reports on integrating modeling abstraction techniques into subsurface monitoring strategies. This research is the culmination of many field and modeling studies conducted by the USDA/Agricultural Research Service (ARS) at their Beltsville Area Research Center. The research design was to identify and examine near-surface water flow pathways by monitoring performance indicators within the unsaturated zone and local water-table system. The peak tracer concentration and the time to peak concentration at several monitoring locations served as the performance indicators. The objective was to apply model abstraction techniques in designing monitoring networks such as those used at nuclear waste and decommissioned facilities. The level of spatial and temporal detail in characterizing soil properties (e.g. water contents and hydraulic conductivities) is based upon the model abstraction considerations. Simplifications may omit significant processes and conditions that control the water and contaminant migration. The ARS field studies provided detailed databases for modeling water and chemical tracer movement in 2 and 3 dimensions to facilitate understanding of what processes and properties could be simplified (abstracted). Model abstractions included using pedotransfer functions for hydraulic conductivity, soil profile homogenization, and unsaturated zone omission. This latter abstraction proved to be the most accurate in generating a monitoring network that reflected the calibrated model. A comprehensive sensitivity analysis was performed to identify possible directions of model simplification in the model abstraction process. The integration of model abstraction into monitoring strategies based upon the ARS field and modeling findings was documented and reviewed by international soil scientists. The studies were jointly funded by NRC and USDA/ARS.

12. KEY WORDS/DESCRIPTORS (List words or phrases that will assist researchers in locating the report.)	13. AVAILABILITY STATEMENT
groundwater flow and transport modeling model abstraction model simplification monitoring strategies pedotransfer function sensitivity analysis soil water subsurface hydrology unsaturated zone	unlimited
	14. SECURITY CLASSIFICATION
	(This Page) unclassified
	(This Report) unclassified
	15. NUMBER OF PAGES
	16. PRICE

NUREG/CR-7221

Integrating Model Abstraction into Subsurface Monitoring Strategies

March 2017