Inverse Modeling of Unsaturated Flow Using Clusters of Soil Texture and Pedotransfer Functions

Yonggen Zhang[1,2], Marcel G. Schaap[3], Alberto Guadagnini[1,4] , and Shlomo P. Neuman[1]

[1]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, 85721

[2]School of Water Resources and Environment, China University of Geosciences, Beijing, 100083

[3]Department of Soil, Water and Environmental Science, University of Arizona, Tucson, AZ, 85721

[4]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

Corresponding author:

Marcel G. Schaap (mschaap@email.arizona.edu)

Telephone number: +1-520-626-4532

Fax Number: +1-520-621-1647

Key points:

- Two new ways to parameterize vadose zone hydraulic properties based on soil texture are proposed and analyzed.
- One of these preserves heterogeneity with only a few adjustable parameters.
- The two approaches are compared through application to deep vadose zone experimental data.

Abstract

Characterization of heterogeneous soil hydraulic parameters of deep vadose zones is often difficult and expensive, making it necessary to rely on other sources of information. Pedotransfer functions (PTFs) based on soil texture data constitute a simple alternative to inverse hydraulic parameter estimation, but their accuracy is often modest. Inverse modeling entails a compromise between detailed description of subsurface heterogeneity and the need to restrict the number of parameters. We propose two methods of parameterizing vadose zone hydraulic properties using a combination of *k*-means clustering of kriged soil texture data, PTFs and model inversion. One approach entails homogeneous and the other heterogeneous clusters. Clusters may include subdomains of the computational grid that need not be contiguous in space. The first approach homogenizes within-cluster variability into initial hydraulic parameter estimates that are subsequently optimized by inversion. The second approach maintains heterogeneity through multiplication of each spatially varying initial hydraulic parameter by a scale factor, estimated *a posteriori* through inversion. This allows preserving heterogeneity without introducing a large number of adjustable parameters. We use each approach to simulate a 95-day infiltration experiment in unsaturated layered sediments at a semiarid site near Phoenix, Arizona, over an area of 50 × 50 m^2 down to a depth of 14.5 m. Results show that both clustering approaches improve simulated moisture contents considerably in comparison to those based solely on PTF estimates. Our calibrated models are validated against data from a subsequent 295-day infiltration experiment at the site.

1. Introduction

Modeling of vadose zone flow and transport processes requires characterization of subsurface architecture and hydraulic properties. Information about lithotype distribution can often be obtained from well logs, ground penetrating radar (GPR) [*Kowalsky et al.*, 2005], electrical resistance tomography (ERT) [*Yeh*, 2002; *Liu and Yeh*, 2004] or seismic tomography [*Nolet*, 1987; *Tromp et al.*, 2005]. In deep vadose zones, hydraulic parameters are difficult if not impossible to measure in situ. Acquiring undisturbed samples and determining their hydraulic properties in the laboratory is likewise difficult and expensive. A common alternative is to characterize hydraulic properties indirectly, on the basis of soil composition and/or flow data, through pedotransfer functions (PTFs) and/or inverse modeling.

PTFs allow the estimation of soil hydraulic properties using empirical correlations between hydraulic characteristics, soil texture and quantities such as soil bulk density. Soil texture and bulk density are generally easier and less expensive to assess than hydraulic properties [e.g., *Rawls et al.*, 1982; *Pachepsky et al.*, 2006]. The combination of a PTF with high-density sampling and/or geospatial modeling of soil texture and bulk density should, in principle, allow one to resolve subsurface hydraulic properties in detail. The accuracy of PTFs is, however, often modest when applied to data collected independently of those employed for PTF calibration [*Schaap and Leij*, 1998]. Calibration of PTFs against site-specific data brings about improvement [*Ye et al.*, 2007] but also requires considerable effort and cost. As shown by *Wang et al.* [2003] and suggested further by our study, PTF-derived estimates of hydraulic properties tend to result in systematic errors that can (and we believe should) be reduced by calibrating these properties further against observed state variables, such as moisture content and/or pressure head, via inverse modeling.

Inverse modeling in hydrogeology typically entails the following steps [e.g., *Neuman*, 1973; *Carrera and Neuman* 1986b, 1986c; *Neuman*, 2003; *Franssen et al.*, 2009]: 1) proposition of a conceptual model (or a set of alternative conceptual models) of the system under study in terms of geological/sedimentological structure, mathematical rendition of (mass, energy, momentum) conservation principles, initial and boundary conditions, as well as forcing terms; 2) parametrization of the system model with the objective of characterizing (either in a stochastic or deterministic fashion) spatial variability of critical parameters throughout the domain; and 3) estimation of model parameters by minimizing a suitable measure of mismatch between observed and simulated state variables (e.g., moisture content or pressure heads). In some cases, this measure includes prior information about model structure and parameters. The latter is typically embedded in a regularization (or plausibility) term, which penalizes deviations of estimated parameters from prior values and helps stabilize the inverse solution [*Neuman*, 1973; *Carrera and Neuman* 1986a, 1986b; *Vrugt and Bouten* 2002]. Inverse methods can be combined with numerically intensive Monte Carlo techniques to cope with propagation of uncertainty associated with spatial parameter distributions (and/or initial conditions and forcing terms) to state variables of interest [e.g., *Zimmerman et al.*, 1998; *Chen and Zhang*, 2006].

Inverse modeling of unsaturated flow is more challenging than that of saturated flow. Notably, functional dependence of soil moisture content and hydraulic conductivity on matric potential leads to high-dimensionality issues in the parameter space, even under conditions where closed-form expressions of these models such as the Brooks-Corey-Burdine [*Burdine*, 1953; *Brooks and Corey*, 1964] or van Genuchten-Mualem formulations are used [*Mualem* 1976; *van Genuchten*, 1980]. Reviews of inverse methods in the context of vadose zone hydrology are found in *Hopmans and Simunek* [1999], *Hopmans et al.* [2002], and *Vrugt et al.* [2008]. Inverse methods based on zonation (subdivision of the flow domain into uniform subdomains [*Emsellem and De Marsily*, 1971; *Neuman and Yakowitz*, 1979; *Wildenschild and Jensen*, 1999; *Wang et al.*, 2003; *Vrugt et al.*, 2008]) are used among practitioners due to their relatively straightforward implementation and its flexibility to accommodate geological information. However, the number of zones (and so the resolution of heterogeneity) is limited by the need to avoid overparameterization. A variation on the zonation method to resolve subsurface heterogeneity is to use similar media scaling [*Miller and Miller*, 1956; *Vogel et al.,* 1991], which relies on the dependence of hydraulic properties on pore size and pore geometry descriptors. This allows scaling of hydraulic water retention and unsaturated hydraulic conductivity functions of multiple soils to unique reference functions [e.g. *Tuli et al.*, 2001; *Das et al.*, 2005; *Nasta et al.*, 2013]. *Zhang et al.* [2004] used a Combined Parameter Scaling and Inversion Technique (CPSIT) to estimate the hydraulic properties of Equivalent Hydraulic Media (EHMs). Their method requires that soil hydraulic parameters at the local scale be determined using the same method as that used for the experimental site and at the same spatial scale, i.e., core size. In their approach, ratios of hydraulic properties in different EHMs relative to hydraulic properties in reference EHMs remain fixed during inversion. Therefore, field scale hydraulic parameters of reference EHMs can be estimated during inversion from local scale values. *Zhang et al.* [2004] recognized this to be a limitation in that any local-scale parameter estimation error transfers to the field scale.

Inverse methods combined with geostatistical methods constitute an alternative approach to estimate soil hydraulic parameters. The Pilot Point Method (PPM), introduced by *de Marsily* [1984], is one such approach and consists of calibrating an initial kriged parameter field, generated from the measured values of hydraulic parameters and a set of additional parameter values (which are unknown prior to calibration) at selected unmeasured locations in the simulation domain, called “pilot points”. The location of pilot points can also be incorporated into the inverse problem to find the optimal position using a couple of adjoint sensitivity analysis and kriging [*LaVenue and Pickens*, 1992; *RamaRao et al.*, 1995; *Franssen et al.*, 2009]. The PPM method has mostly been applied to saturated problems and has found little application in vadose zone systems. *Kowalsky et al*. [2005] used the PPM to derive the distribution of permeability using GPR and hydrological measurements collected during a transient flow experiment. *Morales-Casique et al*. [2010] calibrated log permeability and porosity at selected pilot points against observed pressures in two pneumatic injection tests of unsaturated fractured tuff in Arizona.

In this study, we introduce two ways of parameterizing deep vadose zone hydraulic properties based on *k*-means clustering of kriged soil textural data, a pedotransfer function (PTF) and numerical inversion of a vadose zone flow model. In contrast to traditional zonation often employed in vadose zone inverse modeling, a cluster in our model may (and generally does) consist of noncontiguous subdomains. The initial hydraulic parameters at each grid point in a cluster are estimated with a PTF. Our approach admits that these initial PTF estimates entail systematic errors [*Schaap and Leij*, 1998] which are not known *a priori* [*Romano*, 2004; *Chirico et al.*, 2007; *Assouline and Or*, 2013]. The purpose of our inverse analysis is precisely to minimize these, and ancillary random, errors in parameter estimates. Our approach is predicated on the belief that it is better to rely on reasonably well founded (if not entirely accurate) PTF-derived initial parameter estimates than on other, less robust alternatives. In the first homogeneous cluster approach, each hydraulic parameter (or its logarithm) is averaged over all grid points in a cluster to yield prior hydraulic parameter estimates; posterior estimates (which are uniform within each cluster but differ between clusters) are then estimated by optimizing the fit between computed and observed moisture contents. All prior and posterior parameter estimates within a cluster are homogeneous. In the heterogeneous cluster approach, prior hydraulic parameter estimates vary from one grid point to another. Posterior estimates in a cluster are expressed as products of corresponding prior estimates and a cluster-specific scaling factor. Scaling factors of all parameters in all clusters are then estimated by the same criteria. Both of our approaches are evaluated by model quality measures. We use our approach to simulate a 95-day infiltration experiment in unsaturated layered sediments at a semiarid site near Phoenix, Arizona, over an area of 50×50 m^2 down to a depth of 14.5 m. We then validate our calibrated models against data from a subsequent 295-day infiltration experiment at the site.

1.1. Description of Site and Infiltration Experiment

The data used in this study were collected at the University of Arizona Maricopa Agricultural Center (latitude 33.069478 N, longitude 111.973667 W), Arizona, USA, between 1997 and 2004. Four deep vadose zone infiltration experiments were conducted at this site to test the effectiveness of several vadose zone monitoring instruments and modeling techniques. The site was nominally 60 × 60 m^2 in the horizontal direction and 15 m in the vertical direction and situated in alluvial valley deposits with a textural composition ranging from gravel to clay. An impermeable pond liner was used to eliminate evaporation and rainfall and an inner area of 50 × 50 m^2 (Figure 1) was outfitted with 164 irrigation driplines containing emitters spaced 30 cm apart. Major instrumentation included nine neutron thermalization wells with depths down to 14.25 m (numbered 402…445 in Figure 1) and tensiometers placed one meter south of each well at depths of 3, 5 and 10 m. A perched groundwater table was observed at a depth of about 13 m. Detailed descriptions of the site and instrument calibration are provided by *Young et al.* [1999], *Wang et al.* [2002] and *Schaap* [2013].

Here we focus on data from Experiment 3 and 4 used, respectively, for model calibration and validation. Experiment 3 started on 17 January 2001 (Day-of-Year 17, hereafter termed as DOY 17) and ended on 28 January 2002 (corresponding to DOY 393). An extensive 800 day drainage period preceded Experiment 3, resulting in a nearly constant soil moisture content profile, as verified by neutron thermalization measurements on DOY 17.5, 47.5, 67.5, and 108.5. Drip irrigation (and associated infiltration) started at noon on 24 April 2001 (DOY 114.5) and ended 28 days later at noon on 22 May 2001 (DOY 142.5). With minor interruptions, metered irrigation was applied six times a day at a mean rate of 27.2 mm/day; about 16 mm of water was applied before DOY114.5 to test the irrigation system. Tensiometer readings indicated that full saturation conditions did not occur at any of the monitored locations.

Neutron thermalization was conducted on 42 dates with 0.25 m increments from a depth of 0.25 m down to 12.5 m; neutron count ratios were converted to soil moisture contents using a texture-dependent calibration model presented in *Schaap* [2013]. Sparse data at depths greater than 12.5 m were also available at some wells, but were not used in this study. Data from well 442 (Figure 1) were not considered due to evidence of lateral flow from a flood irrigated field immediately to the north of the site. Data from well 405 at depths of 5.0 - 10.0 m were likewise not used because of anomalously dry readings, presumably due to large air pockets around the PVC well casing. For model calibration data collected between DOY 67.5 and DOY 163.5 were used, which included 29 dates with 11,020 individual moisture content observations. There were two observation dates before the start of irrigation (DOY 114.5), while sampling took place every one to three days during irrigation and at approximately weekly intervals during subsequent redistribution. As explained in Section 2.2.1, the relatively sparse data before infiltration caused some problems with obtaining physically realistic initial moisture contents.

Experiment 4 started on 26 March 2002 (DOY 450) and ended 295 days later on 14 January 2003 (DOY 744). Irrigation started on 26 March 2002 (DOY 450) ended 230 days later on 11 November 2002 (DOY 680). It was followed by a 65-day drainage period that ended on 14 January 2003 (DOY 744). The site was irrigated for 5 minutes, 12 times a day, at a mean rate of 26.8 mm/day. Neutron counts (and, correspondingly, water contents) were measured on 32 dates; 26 dates were measured during the infiltration period and 6 dates during the drainage period. They were conducted at vertical increments of 0.25 m from depth 0.25 m down to 12.5 m in 9 boreholes. Neutron depth coverage was less consistent than in Experiment 3 in that measurements were taken preferentially near the infiltration front, less frequently below it during the infiltration period. Discarding unreliable observations in well 442 due to the same reason as in Experiment 3 left us with a total of 9,297 neutron count values for validation.

1.2. Geospatial Analysis of Soil Texture and Bulk Density

No reliable measurements of site's hydraulic characteristics are available. For this reason, we estimate these characteristics using PTFs based on a geospatial analysis of texture and bulk density data. *Wang* [2002] performed a three-dimensional geostatistical analysis of 520 texture samples collected at depths down to 5 m. *Schaap* [2013] reanalyzed an extended dataset of 1042 soil texture and 250 bulk density samples down to a depth of 15 m and identified two principal components (PC1 and PC2) extracted from measured sand, silt, and clay percentages, with PC1 accounting for 92% of textural triangle variance and PC2 accounting for the remaining 8% of this variance. Residual variograms were obtained upon subtracting spatially-averaged vertical trends from PC1, PC2 and bulk density data (see *Schaap* [2013] for details).

Using the same data as *Schaap* [2013], *Guadagnini et al.* [2013] showed that the univariate distribution of texture is non-Gaussian, rendering texture amenable to representation as a sub-Gaussian random field, key statistics of which vary with scale. The same is true for hydraulic parameters estimated from textural data [*Guadagnini et al.*, 2014] using the pedotransfer function Rosetta of *Schaap et al.* [2001]. Based on results obtained by *Guadagnini et al.* [2013] and [*Guadagnini et al.*, 2014] we recognize that the current kriging approach (see below) has the potential to produce a bias in the prior estimates, which can be reduced, but not eliminated [*Grondona and Cressie*, 1991], through an iterative approach proposed by *Neuman and Jacobson* [1984]. As a result, the approach followed may yield somewhat biased prior hydraulic parameter estimates because kriging of texture and bulk density provides a smooth estimate of actual spatial variability while sub-Gaussian fields may present a more accurate description of the prior parameter estimate.

Because the development of conditional simulation and kriging-based estimation of sub-Gaussian fields of the kind found by *Guadagnini et al.* [2013, 2014] is still in its infancy [*Riva et al.,* 2015; *Panzeri et al*., 2016], we cannot yet simulate conditional sub-Gaussian random fields. Quantifying the potential bias of the kriging approach is impossible. However, we expect it to impact our prior parameter estimates to a greater extent than our posterior estimates, which depend strongly on additional data (in our case, water contents) and are known to be generally less biased.

Our study is thus confined to the kriged three-dimensional (3D) texture and bulk density models that were obtained by *Schaap* [2013] for prior soil hydraulic parameter estimates. The work by *Schaap* [2013] relied on anisotropic Gaussian models with horizontal variogram ranges of 13.1, 5.6 and 7.7 m for PC1, PC2 and bulk density, respectively. Vertical range estimates were 0.28 and 0.85 m for PC1 and PC2, respectively. No reliable estimate of vertical range was found for bulk density and we (as did *Schaap* [2013]) set bulk density below 5 m depth equal to 1.85 g/cm^3. Point kriging was used to obtain PC1, PC2 for the entire $60 \times 60 \times 15$ m domain as well as bulk density for the domain above 5 m depth. We found that a grid with a resolution of $5 \times 5 \times 0.25$ m produced a variability in PC1 and PC2 that was nearly identical to the observations; higher resolution grids did not yield meaningful improvements. The vertical resolution is further consistent with resolution of moisture content measurements. Additional details of the experimental setting and geospatial analysis are given by *Schaap* [2013]. Once kriging was completed, PC1 and PC2 were back-transformed into sand, silt and clay percentages. Point values of these kriged results formed prior estimates for purposes of inverse modeling.

1.3. Soil Hydraulic Properties

The Rosetta-H3 pedotransfer function model [*Schaap et al.*, 2001] was applied to the three-dimensional (3D) kriged fields of sand, silt, clay, and bulk density determined in Section 1.2 to obtain prior estimates of parameters entering into the Mualem-van Genuchten model (*van Genuchten* [1980]; *Mualem* [1976], abbreviated here as VGM) for water retention (1) and unsaturated hydraulic conductivity (2):

$$\theta(h) = \begin{cases} \theta_r + \dfrac{\theta_s - \theta_r}{\left[1 + |\alpha h|^n\right]^m} & h \le 0 \\ \theta_s & h > 0 \end{cases} \tag{1}$$

$$K(S_e) = \begin{cases} K_s S_e^L [1 - (1 - S_e^{1/m})^m]^2 & h \le 0 \\ K_s & h > 0 \end{cases} \tag{2}$$

where

$$S_e = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} \tag{3}$$

is effective saturation; θ is volumetric moisture content (cm^3 cm^{-3}) at matric potential h (cm, < 0 for unsaturated conditions); θ_r (cm^3 cm^{-3}) and θ_s (cm^3 cm^{-3}) are residual and saturated moisture contents, respectively; α (> 0, in cm^{-1}) and n (> 1) are curve-shape parameters; and $m = 1-1/n$. In (2), K_s is the saturated hydraulic conductivity (cm d^{-1}) while L is an empirical parameter with a value of 0.5 (*Mualem* [1976]). The application of Rosetta-H3 to the kriged field of texture and bulk density thus yields 3D distributions of each of the five VGM parameters θ_r, θ_s, α, n and K_s. The purpose of our work is not to compare one PTF with another but to introduce and illustrate two methods of parameterizing vadose zone hydraulic properties based on a (in principle any) PTF and clustering, followed by inversion. Therefore, only Rosetta-H3 PTF is reported in this study.

2. Inverse Modeling Approach

To implement our homogeneous and heterogeneous cluster approaches we subdivide the flow domain into clusters using the method of *k*-means clustering. In the homogeneous cluster approach, initial estimates of the VGM parameters are obtained by averaging over all grid points belonging to a cluster. In the heterogeneous cluster approach, initial VGM estimates vary from one grid point to another, each of which is associated with a cluster-specific scaling factor. We optimize hydraulic parameters during the inversion of the homogeneous approach, while we optimize scale factors in the case of the heterogeneous approach.

2.1. Definition of Clusters and Inversion Method

Various ways to decompose a domain into clusters are available, such as grouping spatially varying kriged values of target quantities into stratigraphic [e.g., *Wang et al.*, 2003] or USDA textural classes. Here we adopt *k-means* clustering according to which kriged values are grouped into k clusters by minimizing

$$\xi(k) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \| \boldsymbol{x}_{ij} - \mu_i \|^2 \tag{4}$$

where n_i is the number of data points belonging to cluster i; $\boldsymbol{x}_{ij}$ is a vector of attributes (i.e., kriged sand, silt, clay percentages in this study) of the j^{th} data point in cluster i; μ_i is a vector of attribute mean values in cluster i; and || || denotes Euclidean norm (in data units). To perform *k-means* clustering we used the algorithm of *Hartigan and Wong* [1979] in the statistical package R (version 3.0.2, *Ihaka and Gentleman*, 1996; *R development core team*, 2005, http://www.R-project.org). A series of preliminary analyses suggested that classification relying on *k-means* clustering of kriged soil texture yielded the most satisfactory results (in terms of root mean square error between computed and observed water contents at all times through the domain of interest). Simulations based on *k-means* clustering of initial soil hydraulic parameters determined in Section 1.3 yielded results of distinctly inferior quality. A reason for this might be that initial soil hydraulic parameter estimates are not very accurate, and posterior estimates are not available prior to inversion. We therefore rely on *k-means* clustering of soil texture. Clustering associates each kriged point with a unique cluster without requiring that points defining a cluster be contiguous in space.

As the first principal component (PC1) represents closely the overall soil texture at the site, we illustrate in Figure 2a its spatial variability along an east-west vertical section at $y = 30$ m (see Figure 1), which passes through wells 422, 423 and 425. This is to be compared with $k = 2, 3, \ldots 6$ soil texture clusters in Figures 2b, 2c, ... 2f defined by the *k-means* method. As expected, increasing the number k of clusters renders their distribution closer and closer to that of PC1 in Figure 2a. Note that the clusters are generally not contiguous in space. Three-dimensional versions of these clusters form the basis for our definition of homogeneous and heterogeneous clusters below.

In the homogeneous cluster approach, each VGM parameter within a cluster is a constant. Table 1 lists arithmetic mean sand, silt and clay percentages for clusters associated with various numbers k of clusters and arithmetic mean values of PTF-derived (using Rosetta-H3) hydraulic parameter estimates in each cluster. Parameters α, n and K_s in Table 1 are antilogs of average $\log_{10}(\alpha)$, $\log_{10}(n)$ and $\log_{10}(K_s)$ Rosetta estimates. These represent prior hydraulic parameters for the homogeneous cluster approach. Posteriors are estimated by the simulation procedure defined in Section 2.2.

In the heterogeneous cluster approach parameters are expressed as

$$\mathbf{p}_i'(\boldsymbol{x}) = \mathbf{B}_i \times \mathbf{p}_i(\boldsymbol{x}) \tag{5}$$

where $\mathbf{p}_i(\boldsymbol{x})$ is a vector whose entries are the five VGM parameters θ_r, θ_s, $\log_{10}(\alpha)$, $\log_{10}(n)$, and $\log_{10}(K_s)$ estimated from texture and bulk density data using Rosetta-H3 at location $\boldsymbol{x} = (x, y, z)$ in cluster i ($i = 1, \ldots, k$). The square matrix $\mathbf{B}_i$ in (5) is taken to be diagonal, containing scaling factors initialized to 1 and then optimized by inversion. Vector $\mathbf{p}_i'(\boldsymbol{x})$ thus represents posterior (inverse) estimates of the five VGM parameters. We designate forward runs with homogeneous clusters HoC and those with heterogeneous clusters HeC. Inverse runs are designated by prefix I and suffix k where

k denotes number of clusters.

2.2. Vadose Zone Flow Simulation and Estimation Criterion

2.2.1. Numerical Simulation of Flow

Water flow is simulated with the Subsurface Transport Over Multiple Phases (STOMP) code of *White and Oostrom* [*2006*] that solves the Richards equation using finite differences with Newton-Raphson iteration. Consistent with the geospatial grid, the simulation grid cells measure 5×5 m^2 horizontally covering an area of 60×60 m^2 (Figure 1), and 0.25 m vertically, extending down to depth 14.5 m. In both Experiment 3 and 4, vertical flux at the top boundary is prescribed to be zero except during irrigation when it is set equal to the daily irrigation rate across the inner 50×50 m^2 area (Figure 1). Pressure head at the bottom boundary, at a depth of 14.5 m, is set equal to positive 1.5 m to reflect the presence of a perched water table at a depth of 13 m [*Wang et al.*, 2003]. As the irrigated area was surrounded by a tarp-covered collar which helped render flow to be predominantly vertical [*Schaap*, 2013], no flow is allowed to take place across the four lateral sides of the grid during flow simulations. Experiment 3 is used for the model calibrations and flow is simulated over a period of 95 days including a 47-day pre-irrigation period from DOY 67.5 to DOY 114.5, 28 days of irrigation from DOY 114.5 to DOY 142.5, and 20 redistribution days from DOY 142.5 to DOY 162.5. Observation data in Experiment 4 is used for model validations, which includes 230-day irrigation period from DOY 450 to DOY 680, and 65 redistribution period from DOY 680 to DOY 744.

Initial moisture contents on DOY 67.5 at locations other than the neutron wells [e.g., *Schaap*, 2013] are not available. Due to the prolonged drainage period prior to DOY 67.5, moisture contents at the neutron wells were nearly constant with the zeroth moment of moisture content (see Appendix A) varying only by a few millimeters over a total profile length of 12.5 m between DOY 17.5 and 114.5 (See Figure 8.5 in *Schaap* [2013]). Tensiometer pressure head readings at depths 3, 5 and 10 m, ranging between negative 2 and negative 3.5 m, were also nearly constant over time. This implies that total head is not constant throughout the profile (but rather varies with depth) and the system is thus not under static equilibrium. Ambient flow prior to infiltration, and following redistribution, is nevertheless negligibly small due to the low hydraulic conductivity of the profile at its ambient water content and pressure head values. We therefore assumed that there should be a strong correlation between texture and neutron thermalization count ratio (CR) [for details see *Schaap*, 2013]. A stepwise regression between observed CR for DOY 67.5 and observed texture yielded the following expression:

$$CR = a \times sand + b \times silt + c \times sand^2 + d \times silt^2 + e \times sand \times silt + f \tag{6}$$

where *sand* and *silt* are expressed as percentages, respectively, while a = 0.3210; b = 0.4038; c = -0.0017; d = -0.0030; e = -0.0039; f = -13.8815; the Pearson correlation coefficient (R) was 0.71. Equation (6) was subsequently used to estimate CR for all non-well grid points in the flow domain. Initial moisture contents for all grid points were subsequently estimated using a site-specific neutron thermalization model [*Schaap*, 2013] :

$$\theta = a \times CR + b \times PC1 + c \times PC1^2 + d \times PC1 \times CR + e \quad (7)$$

where a = 0.520; b = 0.0119; c = -7.545×e^{-5}; d = -5.097×e^{-3}; e = -0.544.

Regardless of inversion method, preliminary simulations consistently resulted in substantial drainage prior to the start of infiltration on DOY 114.5, contradicting observations of nearly constant moisture contents. This may be the consequence of inaccurate initial moisture content derivation from (6) and (7) or inaccurate initial hydraulic parameter estimation by PTF (the estimates being inconsistent with water retention at high pressure or hydraulic conductivity). This problem persisted even after inversion, mainly because of the sparse observations on dates with presumably constant moisture content (2 dates) before the start of infiltration compared to the 26 observations with dynamic moisture content during and after the infiltration period. Ad-hoc approaches, such as assigning large weights to observations before DOY 114.5 did not alleviate the problem.

To eliminate the inconsistency we adopted a three-step approach to inversion, applied in each case. The 3-step approach is not applicable to models HoC and HeC, which do not entail inversion; initial water contents for these two models were based on regression results derived from (6) and (7). In Step I, inversion was conducted by simulating the entire 95-day period of the experiment starting with initial moisture contents determined in the above manner. The same initial moisture contents coupled with parameter estimates obtained in Step I were then used to predict, through forward simulation (Step II), moisture contents at the end of a 50-day continued drainage period (i.e. this period did not have any infiltration). The final moisture contents of Step II, and parameters obtained in Step I, were then assigned as initial values on DOY 67.5 in a final 95-day inversion run during Step III.

Validation was carried out by running the models forward in time from DOY 67.5 till DOY 744 by using initial moisture contents and final parameters from Step III. Simulation results between DOY 450 and 744 were subsequently compared with moisture content observations in Experiment 4.

2.2.2. Model Quality Measures

Model inversion is conducted with PEST [*Doherty et al.*, 1994; *Doherty*, 2003] using Python and Unix-style shell scripts to facilitate data interchange with STOMP. Average soil hydraulic parameters in IHoC*k* and scaling factors in IHeC*k* are estimated by minimizing the sum of squared differences between observed and simulated moisture contents at all times through the domain of interest. Estimates are constrained to ensure that $0 < \theta_r < 0.2$, $0.2 < \theta_s < 0.6$, $0.001 < \alpha < 0.1$ (1/cm), $1.1 < n < 5.0$, $0 < K_s < 50{,}000$ (cm/day). As primary measures of model fit we use the root mean square error

$$RMSE = \sqrt{\frac{1}{N_z}\sum_{i=1}^{N_z}(\theta_i - \theta_i')^2} \tag{8}$$

and coefficient of determination

$$R^2 = 1 - \frac{\sum_{i=1}^{N_z}(\theta_i - \theta')^2}{\sum_{i=1}^{N_z}(\theta_i - \bar{\theta})^2} \tag{9}$$

where N_z is the number of (space-time varying) moisture content observations, N_z = 11,020 in model calibration and 9,297 in model validation; θ_i and θ_i' are the i^{th} observed and simulated moisture contents, respectively; and $\overline{\theta}$ is the average of θ_i. *RMSE* is dimensionless (cm^3 water per cm^3 sediment). Other measures of model fit we use include zeroth, first and second temporal moments of observed and simulated moisture contents (Appendix A).

3. Results and Discussion

3.1. Forward and Inverse Modeling Results

Figure 3 shows how *RMSE* varies with number of clusters for various forward and inverse schemes. As one might expect, *RMSE* is largest (0.0688) in the forward homogeneous single cluster case HoC1, dropping down to below 0.045 as the number of clusters increases. No clustering is required for the heterogeneous case HeC to yield a similarly small *RMSE* of 0.0429 without inversion.

Inversion is seen to reduce *RMSE* considerably in all cases. In the case of IHoC1 *RMSE* decreases from 0.0688 to 0.0619, declining further to 0.0316 as the number of clusters is increased to 2 (IHoC2) down to 0.0260 when this number reaches 4 (IHoC4). Inversion renders the heterogeneous scheme better than the homogeneous scheme: *RMSE* = 0.0309 in the single cluster case IHeC1 and 0.0224 in the four-cluster case IHeC4. It is noted that whereas varying the initial hydraulic parameters of forward and inverse models may change the *RMSE* values, it would not affect our overall conclusions in any significant way.

Figure 3 suggests that, in all cases, increasing the number of clusters beyond 4 fails to reduce *RMSE* further. Like *Neuman* [1973], we attribute this to overparameterization and adopt four clusters as optimum subdivision of our domain. More sophisticated performance metrics such as *AIC* [*Akaike*, 1974; *Ye et al.*, 2008], *AICc* [*Hurvich and Tsai*, 1989] and *BIC* [*Schwarz*, 1978] yielded similar results (not reported).

3.2. Interpretation of Results

Table 2 lists estimated scale factors (ratios between posterior and prior values) associated with IHoC4 and IHeC4 inversion and corresponding standard errors. A standard error is calculated in PEST as the square root of parameter estimation variance; the latter constitute diagonal entries of the parameter covariance matrix, computed to lead order of approximation. We note that these scale factors correspond to $\log_{10}$ transformations of α, n, and K_s, as described in Section 2.1. All standard errors are low, suggesting that so is parameter estimation uncertainty. These models also yielded similar patterns in the resulting optimized scale factors, i.e., if one method adjusts a scale factor upward or downward from an initial ratio of 1.0, so does the other method (with only limited exceptions). The ranges of estimated scale factors within each of the two methods are more substantial for some optimized parameters than for others. The largest range is found for θ_r (from 0.64 to 3.00, across both models and clusters), n (from 0.64 to 1.61), and K_s (from 1.06 to 1.66) and moderate for α (from 0.72 to 1.13) and θ_s (from 0.66 to 0.99). Actual VGM parameter values (not shown in Table 2) were consistent with limited laboratory measurements on disturbed cores.

A visual comparison of simulated moisture contents and observed values corresponding to HeC, IHoC4 and IHeC4 for model calibration and validation is provided in Figure 4. Inversion is seen to improve the quality of this visual comparison markedly in both the homogeneous and heterogeneous clustering cases. Whereas heterogeneous clusters yield better results than do homogeneous clusters, the improvement does not appear to be dramatic for both model calibration and validation. Quantitatively, in model calibration, inversion reduces the *RMSE* (in comparison to HeC of calibration) by 40.8% in the homogeneous and 47.8% in the heterogeneous cases, bringing about an increase in the coefficient of determination (R^2) from 0.66 for HeC through 0.88 for IHoC4 to 0.91 for IHeC4. *RMSE* values associated with models IHoC4 and IHeC4 were lower during the validation period (as they had been during the calibration period) by 43.7% and 49.7%, respectively, than that associated with model HeC; correspondingly, R^2 increased from 0.66 in the case of HeC to 0.83 and 0.87 in the respective cases of IHoC4 and IHeC4. The poor performance of HeC results is likely due to (a) uncertainty in the model used to convert neutron thermalization CR and texture into moisture contents [*Schaap*, 2013], (b) the approximation of initial moisture contents in forward simulations, as discussed in Section 2.2.1, and (c) the assumption of the Gaussian nature of univariate and spatial distributions which is not entirely consistent with findings by *Guadagnini et al.* [2013].

The validation results strengthen our conclusion that both clustering approaches improve parameter estimates considerably in comparison to those based solely on PTF estimates from soil texture. The improvement achieved with heterogeneous clusters is slightly better than that obtained with homogeneous clusters.

We end by comparing in Figures 5a, 5b, and 5c the ways in which the first three temporal moments of moisture content $M0(t)$, $M1(t)$ and $M2(t)$, defined in Appendix A, evolve when computed on the basis of observations and simulations corresponding to HeC, IHoC4 and IHeC4. Only results for Experiment 3 are shown because the number of neutron count measurements during Experiment 4 was too small to allow computing spatial moments. The depicted moments are averages over seven wells as explained in Appendix A. $M0(t)$ represents incremental moisture content between depth 0.25 and 12.5 m, multiplied by this depth (hence given in meters); $M1(t)$ corresponds to mean depth (in meters) of the center of mass of infiltrated water (given in meters); and $M2(t)$ measures the vertical spread of moisture content about its center of mass (in square meters). Because $M1(t)$ and $M2(t)$ are normalized by $M0(t)$, which is small prior to DOY 114.5, their values during this period are unstable and therefore not plotted.

Zeroth moments computed on the basis of observations and simulations correspond closely to that of cumulative infiltration in Figure 5a until DOY 130. Following this date, they first drop below the latter and, following the end of the infiltration period on DOY 142.5, decline with time. This, as explained in Appendix A, is due to the infiltration front's arrival at depth 12.5 m on DOY 130. It also explains the stabilization of $M1(t)$ and gradual decrease in $M2(t)$ seen, respectively, in Figures 5b and 5c. Whereas HeC simulations underestimate observation-based $M1(t)$ significantly at all times, results based on IHoC4 and IHeC4 represent the latter closely and consistently. The poor performance of HeC results can be attributed in part, as noted previously, to the poor definition of initial moisture contents in this forward simulation. Whereas IHoC4 results overestimates observation-based $M2(t)$ significantly at all times, IHeC4 results underestimate the latter at all but intermediate time. It is difficult to tell on the basis of Figure 5 which of these two inverse approach represent observation-based moments better.

4. Conclusions

Our work leads to four major conclusions:

1. Whereas it is possible to estimate deep vadose zone hydraulic parameters on the basis of soil texture data with the aid of a pedotransfer function (PTF), as many soil and climate modelers tend to do, we find it necessary to improve upon these estimates by conditioning them on observed system variables such as moisture content through the adoption of a suitable inverse method. The same conclusion was reached previously by *Wang et al.* [2003].

2. We proposed two ways of parameterizing vadose zone hydraulic properties on the basis of soil texture data by utilizing PTF and *k*-means clustering. In contrast to traditional zonation often employed in hydrologic inverse modeling, a cluster in our model may (and generally does) consist of noncontiguous subdomains. In both of our two approaches hydraulic parameters at each grid point in a cluster are estimated initially with the aid of a PTF. The heterogeneous cluster approach preserves heterogeneity without introducing more adjustable parameters.

3. Upon applying our approach to experimental data from a deep vadose zone site near Maricopa, Arizona, we found clustering combined with inversion improved estimates of moisture contents considerably in comparison to those based solely on soil texture data. The optimum number of clusters in both cases was found to be the same (four). In terms of root mean square errors, the

improvement achieved with heterogeneous clusters was slightly better than that obtained with homogeneous clusters. Moment analysis revealed little differences between the two methods.

4. The calibrated model was validated against an independent infiltration experiment, producing results of essentially the same quality as those obtained during calibration.

Acknowledgments

Site construction and field observations were supported by the US Nuclear Regulatory Commission under contract number NRC-04-97-056 (1997-2002). The data collection for and the development of the geospatial model was supported in part by NSF-EAR grant 0737945 (2005-2009). The current work was supported in part through a contract between the University of Arizona and Vanderbilt University under the Consortium for Risk Evaluation with Stakeholder Participation (CRESP) III, funded by the U.S. Department of Energy. The authors acknowledge previous related work by Mr. Zhufeng Fang.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 19(6), 716-723.

Assouline, S., and D. Or (2013), Conceptual and Parametric Representation of Soil Hydraulic Properties: A Review, *Vadose Zo. J.*, *12*, 1-20.

Brooks, R., and A. Corey (1964), Hydraulic properties of porous media, *Hydrol. Pap. Color. State*

Univ., 3.

Burdine, N. T. (1953), Relative Permeability Calculations From Pore Size Distribution Data, *J. Pet. Technol.*, 5, 71-78.

Carrera, J., and S. P. Neuman (1986a), Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 1. Maximum Likelihood Method Incorporating Prior Information, *Water Resour. Res.*, 22(2), 199-210.

Carrera, J., and S. P. Neuman (1986b), Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 2. Uniqueness, Stability, and Solution Algorithms, *Water Resour. Res.*, 22(2), 211-227.

Carrera, J., and S. P. Neuman (1986c), Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 3. Application to Synthetic and Field Data, *Water Resour. Res.*, 22(2), 228-242.

Chen, Y., and D. Zhang (2006), Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Adv. Water Resour.*, 29(8), 1107-1122.

Chirico, G. B., H. Medina, and N. Romano (2007), Uncertainty in predicting soil hydraulic properties at the hillslope scale with indirect methods, *J. Hydrol.*, *334*(3), 405-22.

Das, B. S., N. W. Haws, and P. S. C. Rao (2005), Defining geometric similarity in soils, *Vadose Zo. J.*, 4(2), 264-270.

de Marsily, G., G. Lavedan, M. Boucher, and G. Fasanino (1984), Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, *Geostatistics Nat. Resour. Charact. Part*, *2*, 831-849.

Doherty, J. (2003), Ground water model calibration using pilot points and regularization, *Ground Water*, 41(2), 170-177.

Doherty, J., L. Brebber, and P. Whyte (1994), PEST: Model-independent parameter estimation, *Watermark Comput. Corinda, Aust*., 122.

Emsellem, Y., and G. de Marsily (1971), An automatic solution for the inverse problem, *Water Resour. Res*., 7(5), 1264-1283.

Franssen, H. J. H., A. Alcolea, M. Riva, M. Bakr, N. Van der Wiel, F. Stauffer, and A. Guadagnini (2009), A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments, *Adv. Water Resour*., 32(6), 851-872.

Grondona, M. O., and N. Cressie (1991), Using Spatial Considerations in the Analysis of Experiments, *Technometrics*, 33(4), 381-392.

Guadagnini, A., S. P. Neuman, M. G. Schaap, and M. Riva (2013), Anisotropic statistical scaling of vadose zone hydraulic property estimates near Maricopa, Arizona, *Water Resour. Res*., 49(12), 8463-8479.

Guadagnini, A., S. P. Neuman, M. G. Schaap, and M. Riva (2014), Anisotropic statistical scaling of soil and sediment texture in a stratified deep vadose zone near Maricopa, Arizona, *Geoderma*, 214-215, 217-227.

Hartigan, J. A., and M. A. Wong (1979), Algorithm AS 136: A K-Means Clustering Algorithm, *J. R. Stat. Soc. C*, 28(1), 100-108.

Hopmans, J. W., and J. Šimůnek (1999), Review of inverse estimation of soil hydraulic properties, Proc. Int. Workshop Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media, University of California, Riverside, CA.

Hopmans, J. W., J. Simunek, N. Romano, and W. Durner (2002), Inverse modeling of transient water flow, *Methods Soil Anal. Part 1, Phys. Methods*, 2.

Hurvich, C. M., and C. L. Tsai (1989), Regression and time series model selection in small samples, *Biometrika*, 76, 297-307.

Ihaka, R., and R. Gentleman (1996), R: A Language for Data Analysis and Graphics, *J. Comput. Graph. Stat.*, 5(3), 299-314.

Kowalsky, M. B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee (2005), Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data, *Water Resour. Res.*, 41(11), 1-19.

LaVenue, A. M., and J. F. Pickens (1992), Application of a coupled adjoint sensitivity and kriging approach to calibrate a groundwater flow model, *Water Resour. Res.*, *28*(6), 1543-1569.

Liu, S., and T.-C. J. Yeh (2004), An Integrative Approach for Monitoring Water Movement in the

Vadose Zone, *Vadose Zo. J.*, 3, 681-692.

Miller, E. E., and R. D. Miller (1956), Physical theory for capillary flow phenomena, *J. Appl. Phys.*, 27(4), 324-332.

Morales-Casique, E., S. P. Neuman, and V. V Vesselinov (2010), Maximum likelihood Bayesian averaging of airflow models in unsaturated fractured tuff using Occam and variance windows, *Stoch. Environ. Res. risk Assess.*, *24*(6), 863-880.

Mualem, Y. (1976), A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water Resour. Res.*, 12(3), 513-522.

Nasta, P., N. Romano, S. Assouline, J. A. Vrugt, and J. W. Hopmans (2013), Prediction of spatially variable unsaturated hydraulic conductivity using scaled particle-size distribution functions, *Water Resour. Res.*, 49(7), 4219-4229.

Neuman, S. P. (1973), Calibration of distributed parameter groundwater flow models viewed as a multiple - objective decision process under uncertainty, *Water Resour. Res.*, 9(4), 1006-1021.

Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of uncertain model predictions, Stoch. *Environ. Res. Risk Assess.*, 17(5), 291-305.

Neuman, S. P., and E. A. Jacobson (1984), Analysis of nonintrinsic spatial variability by residual kriging with application to regional groundwater levels, *J. Int. Assoc. Math. Geol.*, 16(5), 499-521.

Neuman, S. P., and S. Yakowitz (1979), A statistical approach to the inverse problem of aquifer hydrology, 1. theory, *Water Resour. Res.*, 15(4), 845-860.

Nolet, G. (1987), Seismic wave propagation and seismic tomography, in *Seismic tomography*, pp. 1-23, Springer.

Pachepsky, Y. A., W. J. Rawls, and H. S. Lin (2006), Hydropedology and pedotransfer functions, *Geoderma*, 131(3-4), 308-316.

Panzeri, M., M. Riva, A. Guadagnini, and S.P. Neuman (2016), Theory and generation of conditional, scalable sub-Gaussian random fields, *Water Resour. Res.*, 52.

RamaRao, B. S., A. M. LaVenue, G. De Marsily, and M. G. Marietta (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 1. Theory and computational experiments, *Water Resour. Res.*, *31*(3), 475-493.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton (1982), Estimation of Soil Water Properties, *Trans. ASAE*, 25(5), 1316-1320 & 1328.

Riva, M., S. P. Neuman, and A. Guadagnini (2015), New scaling model for variables and increments with heavy-tailed distributions, *Water Resour. Res.*, *51*(6), 4623-4634.

Romano, N. (2004), Spatial structure of PTF estimates, in *Development of pedotransfer functions in soil hydrology*, vol. 30, edited by Y. A. Pachepsky and W. J. Rawls, pp. 295-319, Elsevier, New York.

Schaap, M. G. (2013), Description, Analysis, and Interpretation of an Infiltration Experiment in a Semiarid Deep Vadose Zone, in *Advances in Hydrogeology*, pp. 159-183, Springer.

Schaap, M. G., and F. J. Leij (1998), Using neural networks to predict soil water retention and soil hydraulic conductivity, *Soil Tillage Res*., 47(1-2), 37-42.

Schaap, M. G., F. J. Leij, and M. T. Van Genuchten (2001), Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, *J. Hydrol*., 251(3-4), 163-176.

Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat*., 6(2), 461-464.

Tromp, J., C. Tape, and Q. Liu (2005), Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int*., 160(1), 195-216.

Tuli, A., K. Kosugi, and J. W. Hopmans (2001), Simultaneous scaling of soil water retention and unsaturated hydraulic conductivity functions assuming lognormal pore-size distribution, *Adv. Water Resour*., 24(6), 677-688.

van Genuchten, M. T. (1980), A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Soil Sci. Soc. Am. J*., 44(5), 892.

Vogel, T., M. Cislerova, and J. W. Hopmans (1991), Porous Media With Linearly Variable Hydraulic Properties, *Water Resour. Res*., 27(10), 2735-2741.

Vrugt, J. A., W. Bouten, H. V. Gupta, and S. Sorooshian (2002), Toward improved identifiability of hydrologic model parameters: The information content of experimental data, *Water Resour. Res.*, 38(12), 48-1-48-13.

Vrugt, J. A., P. H. Stauffer, T. Wöhling, B. A. Robinson, and V. V. Vesselinov (2008), Inverse Modeling of Subsurface Flow and Transport Properties: A Review with New Developments, *Vadose Zo. J.*, 7(2), 843.

Wang, W. (2002), Uncertainty analysis of groundwater flow and solute transport in unsaturated-saturated porous medium: Maricopa case, Ph.D diss, The University of Arizona.

Wang, W., S. P. Neuman, T. Yao, and P. J. Wierenga (2003), Simulation of Large-Scale Field Infiltration Experiments Using a Hierarchy of Models Based on Public, Generic, and Site Data, Vadose Zo. J., 2, 297-312.

White, M. D., and M. Oostrom (2006), STOMP Subsurface Transport Over Multiple Phases Version 4.0 User's Guide, Richland, Washington.

Wildenschild, D., and K. H. Jensen (1999), Numerical modeling of observed effective flow behavior in unsaturated heterogeneous sands, *Water Resour. Res.*, *35*(1), 29-42.

Ye, M., R. Khaleel, M. G. Schaap, and J. Zhu (2007), Simulation of field injection experiments in heterogeneous unsaturated media using cokriging and artificial neural network, *Water Resour. Res.*, 43(7).

Ye, M., P. D. Meyer, and S. P. Neuman (2008), On model selection criteria in multimodel analysis, *Water Resour. Res.*, 44(3).

Yeh, T.-C. J. (2002), A geostatistically based inverse model for electrical resistivity surveys and its applications to vadose zone hydrology, *Water Resour. Res.*, 38(12), 1-13.

Young, M. H. et al. (1999), Results of Field Studies at the Maricopa Environmental Monitoring Site, Arizona, University of Arizona.

Zhang, Z. F., A. L. Ward, and G. W. Gee (2004), A combined parameter scaling and inverse technique to upscale the unsaturated hydraulic parameters for heterogeneous soils, *Water Resour. Res.*, 40(8).

Zimmerman, D. A. et al. (1998), A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, *Water Resour. Res.*, 34(6), 1373.

Appendix A: Temporal moment analysis

For a given borehole the zeroth order temporal moment of moisture contents is defined as

$$M0(t) = \sum_{z=0.25}^{12.5} \theta_{diff}(t,z) \times \Delta z \tag{A1}$$

where θ_{diff} is the difference between observed or simulated moisture contents at time t and their initial values, z being depth and Δz = 0.25 m a depth increment. $M0(t)$ represents the incremental moisture content between depth 0.25 and 12.5 m, multiplied by this depth (hence given in meters). In this study we calculate $M0(t)$ at each of seven neutron wells and average the results (some values measured in well 405, and all values measured in well 442, are considered to be unreliable; we therefore exclude these two wells from our analysis of moments). The temporal evolution of this average $M0$ should follow closely the actual cumulative amount of irrigation water in the absence of horizontal and vertical drainage losses.

The first temporal moment is calculated as

$$M1(t) = \frac{1}{M0} \sum_{z=0.25}^{12.5} \theta_{diff}(t,z) \times z \times \Delta z \tag{A2}$$

and given in square meters. $M1(t)$ represents the mean depth (in meters) of the center of mass of infiltrated water in a borehole; as in the case of $M0(t)$, we average it over seven wells. When water drainage below a depth of 12.5 m is not negligible (in which case $M0(t)$ does not coincide with the total volume injected), $M1(t)$ provides information only about the center of mass of infiltrating water above this depth.

The second temporal moment,

$$M2(t) = \frac{1}{M0}\sum_{z=0.25}^{12.5}\theta_{diff}(t,z)\times z^2 \times \Delta z - M1^2 \,, \tag{A3}$$

measures the vertical spread of moisture content about its center of mass.

Tables

Table 1. *Soil clusters based on textural data, corresponding mean sand, silt and clay percentages and PTF derived mean hydraulic parameters.*

Number k of clusters	Cluster	Sand %	Silt %	Clay %	θ_r, cm^3cm^{-3}	θ_s, cm^3cm^{-3}	α, 1/cm	n	K_s, cm/day
1	a	81.034	12.633	6.333	0.041	0.316	0.040	1.867	64.845
2	a	71.279	19.135	9.587	0.040	0.328	0.043	1.429	24.912
	b	88.488	7.665	3.846	0.042	0.308	0.038	2.292	134.694
3	a	68.026	21.280	10.694	0.041	0.333	0.042	1.376	20.858
	b	89.901	6.850	3.249	0.042	0.306	0.037	2.440	165.878
	c	79.063	13.751	7.186	0.041	0.318	0.044	1.609	40.933
4	a	92.769	5.149	2.083	0.040	0.302	0.036	2.814	261.011
	b	66.602	22.072	11.327	0.041	0.336	0.040	1.363	19.750
	c	85.501	9.401	5.098	0.043	0.312	0.040	1.981	84.388
	d	75.766	16.291	7.944	0.040	0.321	0.046	1.498	31.308
5	a	77.282	14.861	7.857	0.040	0.321	0.045	1.539	35.178
	b	92.829	5.125	2.046	0.040	0.302	0.036	2.823	263.575
	c	65.888	19.041	15.071	0.049	0.371	0.029	1.403	24.892
	d	85.859	9.211	4.931	0.043	0.311	0.040	2.008	88.397
	e	69.443	23.309	7.248	0.033	0.302	0.055	1.355	18.064
6	a	92.927	5.116	1.957	0.040	0.301	0.036	2.840	267.898
	b	62.320	23.681	14.000	0.045	0.353	0.032	1.357	18.383
	c	85.233	6.696	8.071	0.047	0.328	0.036	1.881	81.146
	d	70.488	20.305	9.207	0.038	0.324	0.047	1.387	22.141
	e	86.240	10.699	3.061	0.041	0.302	0.043	2.094	94.359
	f	77.581	14.791	7.628	0.040	0.320	0.045	1.546	35.484

Table 2. *Optimized values of scale factors and standard errors for IHoC4 and IHeC4 models. Scale factors for IHoC4 were obtained upon dividing optimal VGM parameters by their initial estimates in Table 1 (4 clusters).*

Cluster	Parameters	IHoC4		IHeC4	
		Scale factor	Standard error	Scale factor	Standard error
a	θ_r	2.565	0.0023	2.8816	0.0064
	θ_s	0.9868	0.0025	0.6649	0.0052
	α	1.043	0.0098	0.9766	0.0023
	n	1.5454	0.003	1.196	0.0025
	K_s	1.3122	0.0226	1.0555	0.002
b	θ_r	0.8878	0.0004	0.6438	0.0053
	θ_s	0.8708	0.0005	0.825	0.0028
	α	0.7153	0.0004	0.9405	0.0017
	n	0.6492	0.0006	0.7354	0.0018
	K_s	1.2449	0.0069	1.5657	0.0018
c	θ_r	2.6837	0.0004	3.0034	0.0048
	θ_s	0.9385	0.001	0.7127	0.0051
	α	1.0039	0.005	0.9133	0.0027
	n	1.4952	0.0033	1.0051	0.0025
	K_s	1.3257	0.0018	1.6599	0.0017
d	θ_r	2.6325	0.0006	1.0241	0.0093
	θ_s	0.9153	0.0007	0.9259	0.0025
	α	1.1303	0.0051	0.8413	0.0021
	n	1.6109	0.0024	0.6401	0.0026
	K_s	1.242	0.0067	1.4387	0.0033

Figures

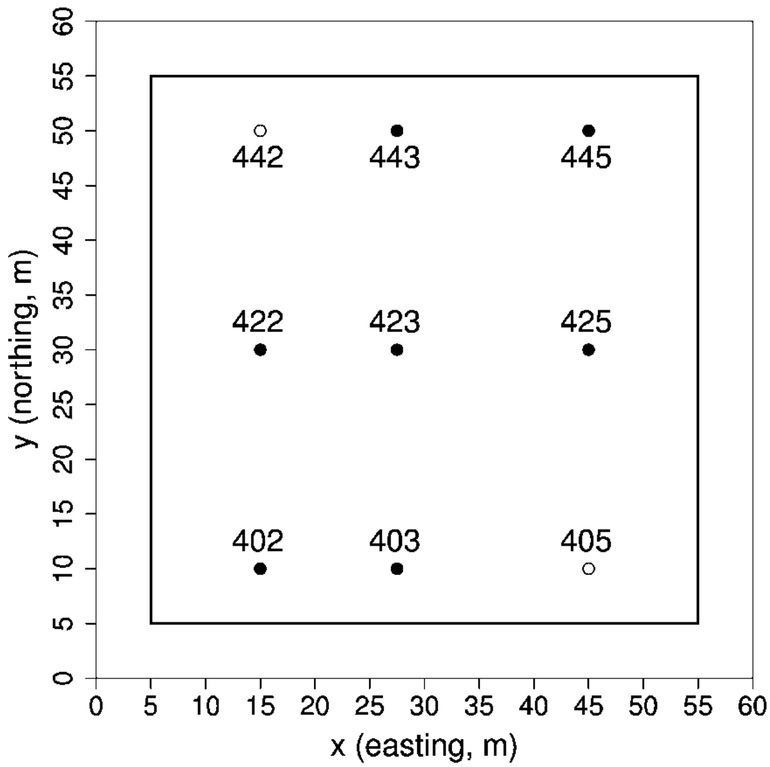

Figure 1. *Location of nine monitoring boreholes at Maricopa site. All moisture content data from wells designated by solid circles were employed during inversion; all or some such data in wells designated by open circles were considered unreliable and omitted (see text). The 60×60 meter outer solid square was covered by tarp to prevent evaporation; the inner 50×50 meter square was drip irrigated.*

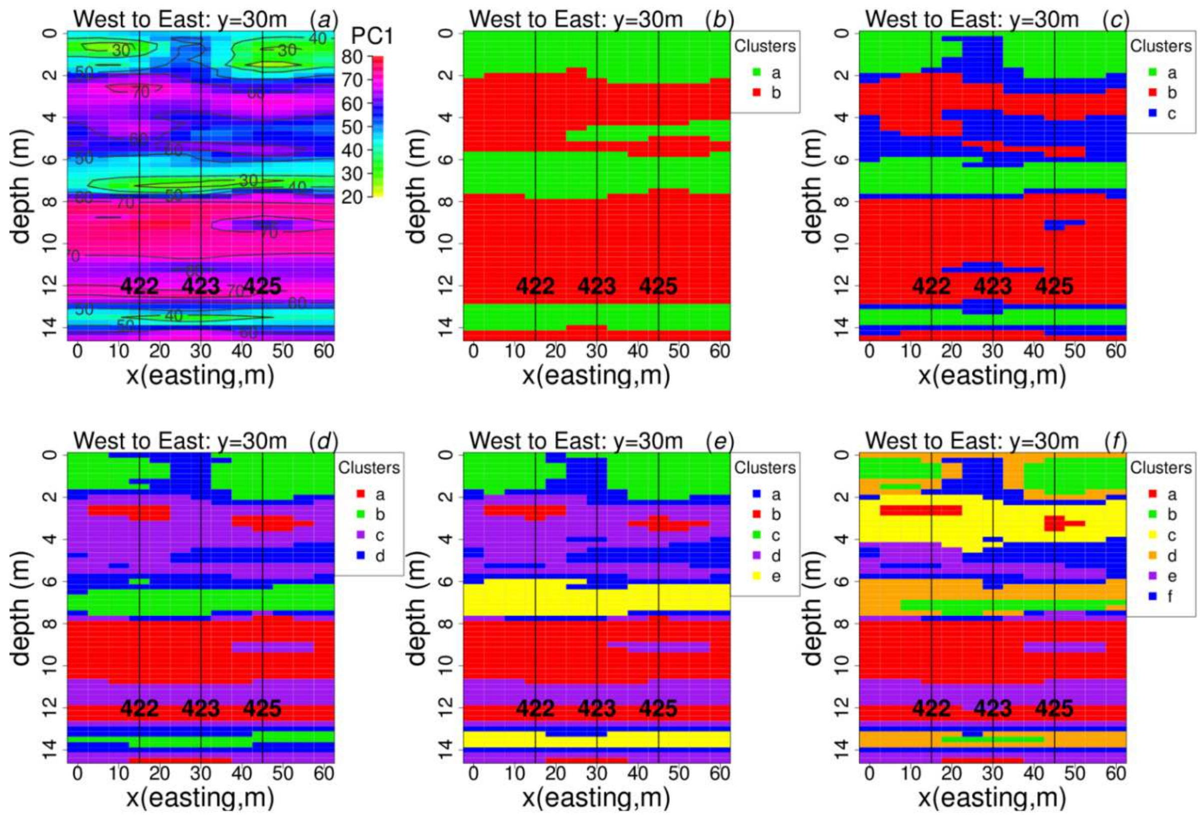

Figure 2. *(a) cross sectional depth profile (at y = 30m in Figure 1) of first principal component (PC1) extracted from soil texture data with labeled contours of PC1 values. PC1 measures the coarseness of soil similar to that of sand ; (b-f) clusters of kriged soil texture data with k = 2-6. Numbers at bottom designate well numbers in Figure 1.*

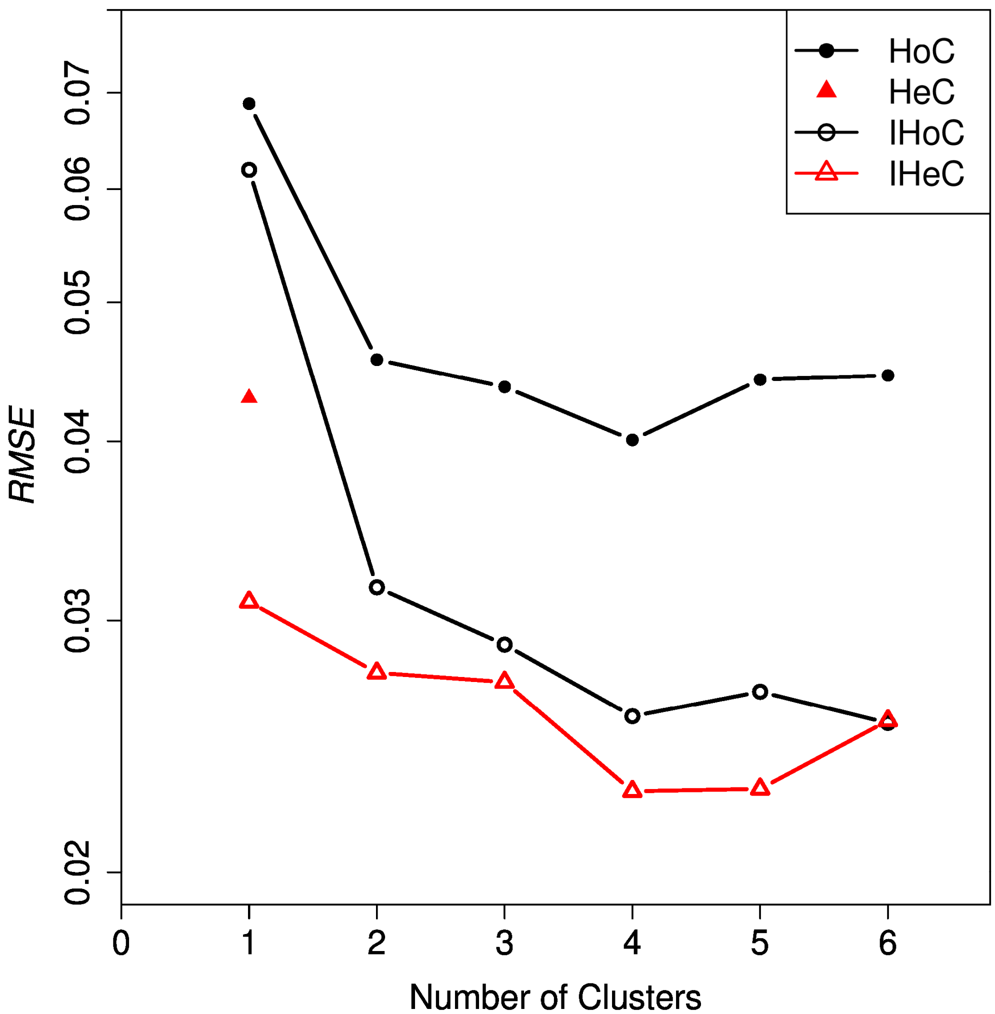

Figure 3. *RMSE versus number of clusters and model type (vertical axis is in logarithmic scale). HoC is initial simulation of IHoC inversion; HeC indicates simulation of full heterogeneous domain with hydraulic parameter estimates from Rosetta-H3 at all grid points; IHoC represents homogeneous cluster inversion; IHeC indicates heterogeneous cluster inversion.*

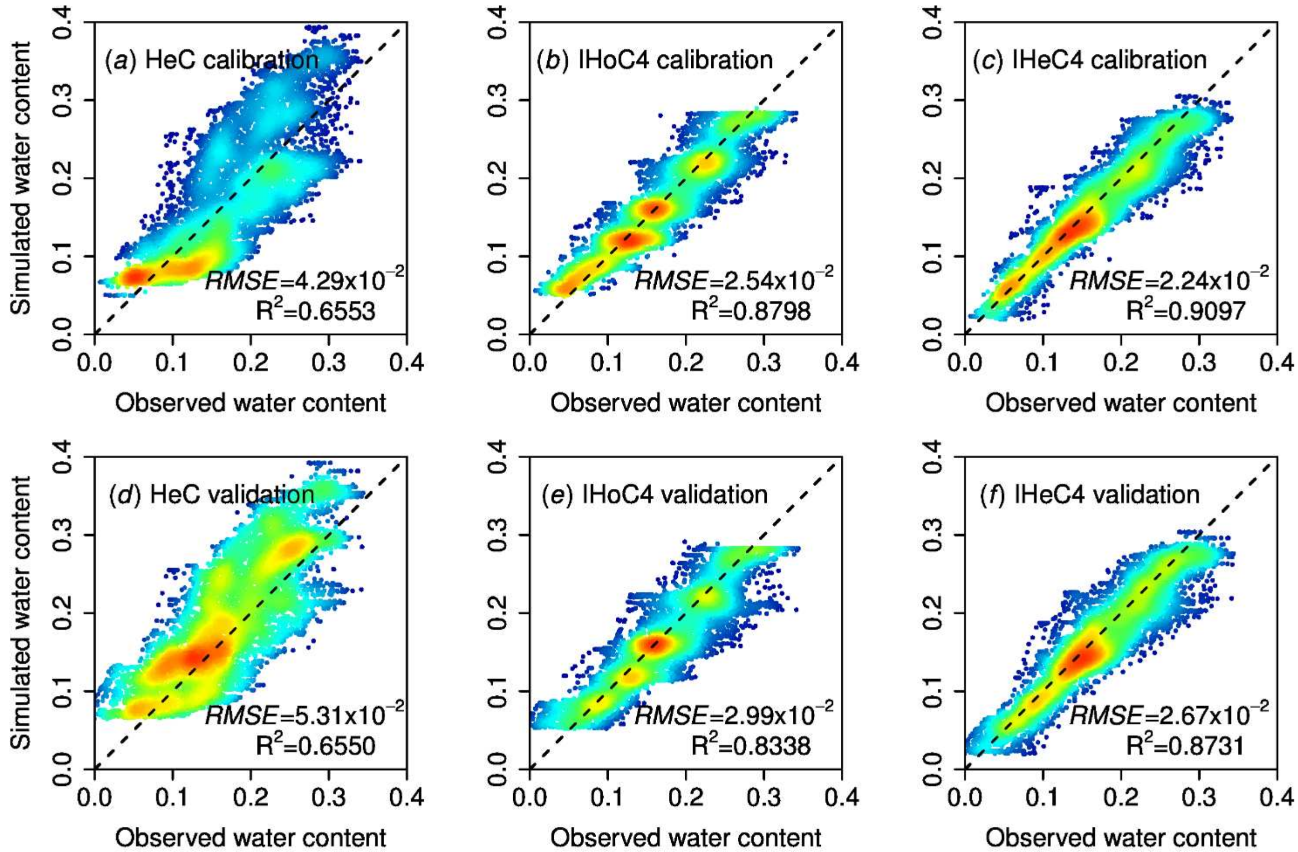

Figure 4. *Comparison between observed and simulated moisture contents using calibration data for (a) HeC, (b) IHoC4 and (c) IHeC4 and validation data for (d) HeC, (e) IHoC4 and (f) IHeC4. Red represents high data density, blue low density.*

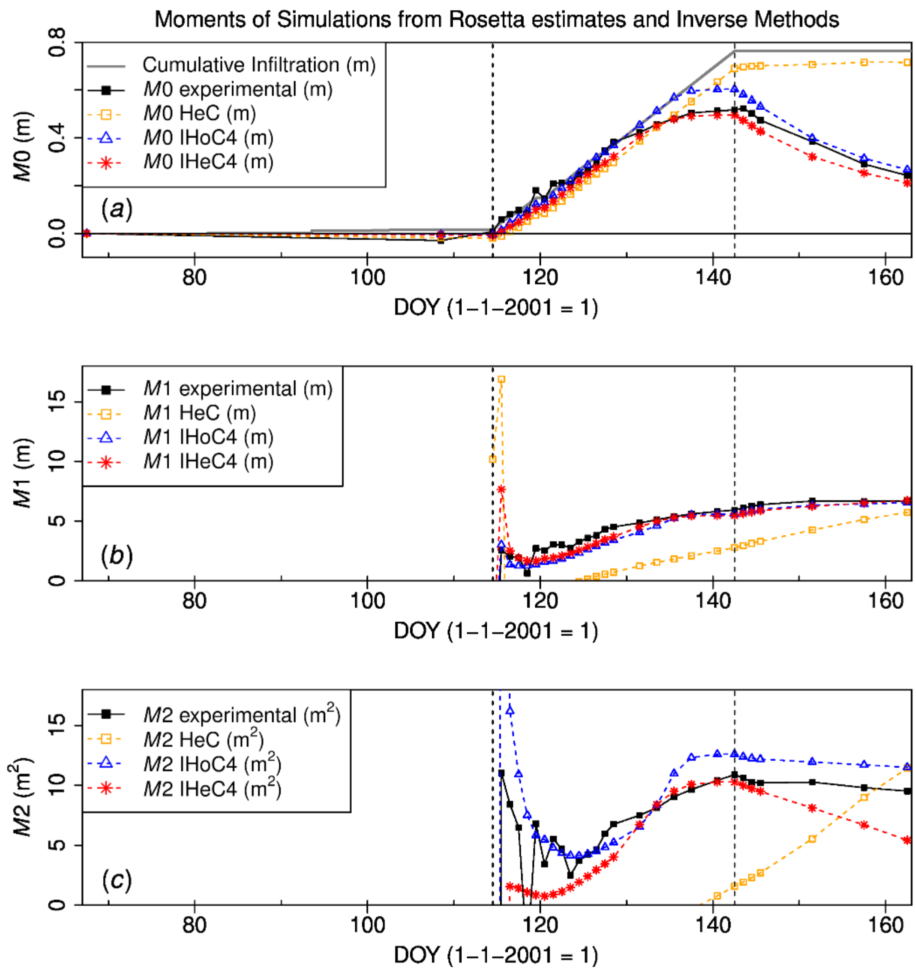

Figure 5. *Comparison of (a) zeroth (M0), (b) first (M1) and (c) second moment (M2) of measured and simulated moisture front in wells based on experimental data, HeC, IHoC4 and IHeC4. Vertical dashed lines indicate start (DOY 114.5) and end (DOY 142.5) dates of infiltration.*

Table Captions

Table 1. Soil clusters based on textural data, corresponding mean sand, silt and clay percentages and PTF derived mean hydraulic parameters.

Table 2. Optimized values of scale factors and standard errors for IHoC4 and IHeC4 models. Scale factors for IHoC4 were obtained upon dividing optimal VGM parameters by their initial estimates in Table 1 (4 clusters).

Figure Captions

Figure 1. Location of nine monitoring boreholes at Maricopa site. All moisture content data from wells designated by solid circles were employed during inversion; all or some such data in wells designated by open circles were considered unreliable and omitted (see text). The 60×60 meter outer solid square was covered by tarp to prevent evaporation; the inner 50×50 meter square was drip irrigated.

Figure 2. (a) cross sectional depth profile (at y = 30m in Figure 1) of first principal component (PC1) extracted from soil texture data with labeled contours of PC1 values. PC1 measures the coarseness of soil similar to that of sand ; (b-f) clusters of kriged soil texture data with k = 2 - 6. Numbers at bottom designate well numbers in Figure 1.

Figure 3. RMSE versus number of clusters and model type (vertical axis is in logarithmic scale). HoC is initial simulation of IHoC inversion; HeC indicates simulation of full heterogeneous domain with hydraulic parameter estimates from Rosetta-H3 at all grid points; IHoC represents homogeneous

cluster inversion; IHeC indicates heterogeneous cluster inversion.

Figure 4. Comparison between observed and simulated moisture contents using calibration data for (a) HeC, (b) IHoC4 and (c) IHeC4 and validation data for (d) HeC, (e) IHoC4 and (f) IHeC4. Red represents high data density, blue low density.

Figure 5. Comparison of (a) zeroth (M0), (b) first (M1) and (c) second moment (M2) of measured and simulated moisture front in wells based on experimental data, HeC, IHoM4 and IHeC4. Vertical dashed lines indicate start (DOY 114.5) and end (DOY 142.5) dates of infiltration.