Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen-Loève expansion with the Markov chain Monte Carlo technique

N. N. Das,[1] B. P. Mohanty,[1] and Y. Efendiev[2]

Received 18 April 2008; revised 23 November 2009; accepted 3 February 2010; published 25 June 2010.

[1] Process-based soil hydrologic models require input of saturated hydraulic conductivity (K_{sat}). However, model users often have limited access to measured data and thus use published or estimated values for many site-specific hydrologic and environmental applications. We proposed an algorithm that uses the Karhunen-Loève expansion (KLE) in conjunction with the Markov chain Monte Carlo (MCMC) technique, which employs measured soil moisture values to characterize the saturated hydraulic conductivity of an agricultural field at a 30 m resolution. The study domain is situated in the Walnut Creek watershed, Iowa, with soybean crop (in 2005) and well-defined top (atmospheric) and bottom (groundwater) boundary conditions. The KLE algorithm parameterizes and generates K_{sat} fields with random correlation lengths that are used in the SWMS_3D model for predicting the soil moisture dynamics for two different scenarios: (1) the van Genuchten soil hydraulic parameters (except K_{sat}) are constant and are based on the soil type of the grid block within the domain, and (2) K_{sat} is correlated with the van Genuchten parameter α as $K_{sat} \propto \alpha^2$. The predicted soil moisture fields for both the scenarios are evaluated with the measured soil moisture in the MCMC algorithm for acceptance (or rejection) of the K_{sat} fields. The accepted K_{sat} fields are evaluated against the laboratory-measured K_{sat} at specific locations as well as with a large K_{sat} data set measured in situ in a nearby field with similar hydrologic conditions, and the comparisons show reasonably good agreement. The KLE-MCMC algorithm was further tested in the same study domain for another year (2002) having different vegetation (corn) and local forcings. The algorithm shows potential to characterize the effective saturated hydraulic conductivity fields at 30 m resolution using inexpensive and more regularly measured soil moisture measurements. Further studies are required to incorporate variability in different hydroclimatic regions and diverse topography to extend the application of this algorithm.

Citation: Das, N. N., B. P. Mohanty, and Y. Efendiev (2010), Characterization of effective saturated hydraulic conductivity in an agricultural field using Karhunen-Loève expansion with the Markov chain Monte Carlo technique, *Water Resour. Res.*, *46*, W06521, doi:10.1029/2008WR007100.

1. Introduction

[2] Characterization of infiltration is important for better understanding of overland and subsurface water flow, and chemical transport in the vadose zone. A high degree of spatial variability in local infiltration is observed due to the random nature of saturated hydraulic conductivity (K_{sat}) [*Nielsen et al.*, 1973; *Sharma et al.*, 1987; *Warrick and Nielsen*, 1980; *Mohanty et al.*, 1996, 1997; *Mohanty*, 1999]. Saturated hydraulic conductivity is a complex soil property to describe because it can change many orders of magnitude over small distance. Numerous studies [*Loague and Gander*, 1990; *Mohanty et al.*, 1994a; *Mohanty and Mousli*, 2000; *Nielsen et al.*, 1973; *Sharma et al.*, 1987] have shown large variability exhibited by K_{sat}. In an agricultural field, spatial variability of K_{sat} is influenced by soil texture, cropping/tillage practice, landscape positions, and growing seasons [*Azevedo et al.*, 1998; *Mohanty et al.*, 1994a; *Mohanty and Kanwar*, 1994; *Mohanty and Mousli*, 2000]. *Carsel and Parrish* [1988] reported soil texture–based descriptive statistics of K_{sat}. They found the highest coefficient of variation for K_{sat} (ranged from 453.3 for silty clays to 52.4 for sands), and observed a large difference in the value of K_{sat} between sand and clay. Various other databases are also available that highlight the variability present in K_{sat}. Prominent among them is the UNSODA database [*Nemes et al.*, 2001] that provides wide range of K_{sat} values for uniform single-grain size sand, loam, and clay soils to mixtures thereof. Decayed root channels lead to the formation of macropores that ultimately enhance infiltration in soil. *Meek et al.* [1990] showed that alfalfa root system generates an extensive macropore flow network, consequently increasing K_{sat}. Wetting/drying, freezing/thawing, and earthworm activities in agricultural

[1]Department of Biological and Agricultural Engineering, Texas A&M University, College Station, Texas, USA.
[2]Department of Mathematics, Texas A&M University, College Station, Texas, USA.

0043-1397/10/2008WR007100

fields may also contribute to macropore formation that leads to a high infiltration rate. However, *Wager and Denton* [1989] and *Mohanty et al.* [1994a] also found that interrow wheel tracks by farm machinery compacted the underlying soil, resulting in low infiltration capacity and leading to a great reduction in K_{sat} in tracked area as compared to untracked interrow areas.

[3] Several laboratory and in situ techniques are available to measure K_{sat} and the results of these techniques often vary significantly [*Gupta et al.*, 1993; *Klute and Dirksen*, 1986; *Mohanty et al.*, 1994a; *Paige and Hillel*, 1993] based on different measurement support sizes and the governing principles. The spatial support of the measurement may be representative of few square centimeters to several square meters. Macropores also tend to increase the variability of K_{sat} measurements when small sample sizes are used [*Mohanty et al.*, 1994b, 1996, 1997; *Mohanty*, 1999]. Variability of K_{sat} in soils has also been viewed with respect to the numerous independent processes operating at different spatial and temporal scales [*McBratney*, 1998]. Recognizing that K_{sat} represents an end result of a number of independent processes, whose complex interactions makes it impractical to apply deterministic methods for describing the spatial continuity of K_{sat}. Studies have been conducted to characterize stochasticity of K_{sat} by autocorreleograms and variograms [*Hillel*, 1980; *Wierenga*, 1985]. *Mohanty et al.* [1991] used a simplified, split-window median-polishing technique in conjunction with a robust semivariogram estimator to examine the spatial structure of K_{sat} in a glacial till material. *Gupta et al.* [1992] used Fourier series analysis along with autoregressive methods to model hydraulic conductivity as a stochastic process. Due to such high spatial heterogeneity, K_{sat} is often modeled as a random field characterized by a lognormal probability density function with relevant spatial correlation. Thus, the characterization of K_{sat} generally requires experimental data acquisition, which is time consuming and costly.

[4] In the last two decades, many stochastic theories have been developed to study the spatial variability of hydrologic processes in the saturated zone [e.g., *Roy and Grilli*, 1997; *Yang et al.* [2004]. *Roy and Grilli* [1997] solved the steady state saturated flow equation by using the Karhunen-Loève expansion (KLE) combined with the perturbation method and obtained the mean head and the head variance in first and second order. *Yang et al.* [2004] worked on the stochastic saturated–unsaturated flow equation with spatially random hydraulic conductivity and solved for a pore size distribution parameter based on the combination of Karhunen-Loève expansion of the random inputs of the media properties and a perturbation method. In another relevant study, *Lu and Zhang* [2004] extended a moment equation approach based on KLE to a saturated heterogeneous porous medium. In that study, the K_{sat} field was decomposed using KLE conditioned on direct measurements that led to the determination of the higher moments of pressure head. They suggested that the KLE-based moment equation approach with higher-order corrections is superior to the conventional first-order approximations. *McLaughlin and Townley* [1996] have shown parameterization of large-scale log hydraulic conductivity using KLE for inverse problem in a groundwater system.

[5] As most of these past efforts focused on saturated heterogeneous porous media, in this study, our primary objective is to parameterize the effective K_{sat} for the vadose zone at a field scale with inexpensive and more regularly observed surface soil moisture measurements. We employed a physically based hydrologic modeling in conjunction with the Karhunen-Loève expansion (KLE) and Markov chain Monte Carlo (MCMC) method to characterize the effective K_{sat} distribution in the vadose zone for an agricultural field using available surface soil moisture data. With KLE, we can represent the high-dimensional K_{sat} field by a small number of parameters. Furthermore, proposed K_{sat} at some sparse locations can be incorporated into the KLE to further reduce the dimensionality of the parameter space. Imposing the known values of K_{sat} at certain locations within the field restricts the parameter space to a subspace (hyperplane). The effective K_{sat} fields from the KLE algorithm can be used in a three-dimensional hydrologic model with well-defined boundary conditions to obtain the evolution of profile soil moisture in the field site. The simulated profile soil moisture and the soil moisture measurements at the field site are used in a simple MCMC-based random walk sampler to accept or reject the K_{sat} realizations from the KLE algorithm. The posterior distribution of K_{sat} fields are subjected to statistical analysis and tested against the measured K_{sat} values at specific field locations. Additionally, the algorithm is further tested in the same study domain for a separate year having different vegetative cover (i.e., corn) and local forcings.

2. Study Area and Data Collection

[6] During the Soil Moisture Experiments in 2002 (SMEX02) and 2005 (SMEX05), hydrological sampling campaigns (details available at http://ars.usda.gov/Research/docs.htm) at field, watershed, and regional scale were conducted in the Walnut Creek watershed (Figure 1) and regional sites near Ames, Iowa. Our study uses the field data of SMEX02 and SMEX05 from the Walnut Creek watershed. Approximately 95% of the watershed is used for row crop agriculture (corn and soybean). The climate of the region is humid and average annual rainfall is 835 mm. The topography has low relief and poor surface drainage characteristics, resulting from prairie potholes that are water-holding depressions of glacial origin. In the watershed, ground sampling was conducted in 32 fields, each approximately 800 m × 800 m, for aircraft based soil moisture remote sensing validation. For our study, WC11 field (Figure 1) within the watershed was selected that was used for high-resolution soil moisture sampling during the SMEX02 and SMEX05 experiments. The geographical location of WC11 is 41.97°N and 93.69°W. The WC11 had corn during the SMEX02, and soybean during the SMEX05 with a small area of corn planted near the western edge of the field. The portion of western edge was eliminated from the study domain to obtain uniform vegetation cover that rendered it to a size of 600 m × 600 m. Average rooting depth of 30–60 cm for corn and 30–50 cm for soybean during the vegetative growth and development stage [*Smajstrla*, 1990; also observed during the field campaigns] was used. The leaf area index (LAI) measured during the SMEX02 and the SMEX05 field campaigns were used in the study to compute the plant water uptake rate (as described by *Feddes et al.* [1978]). The topography of the WC11 is characterized by low relief (Figure 1). The high-resolution digital elevation model data from SRTM reveals a low of

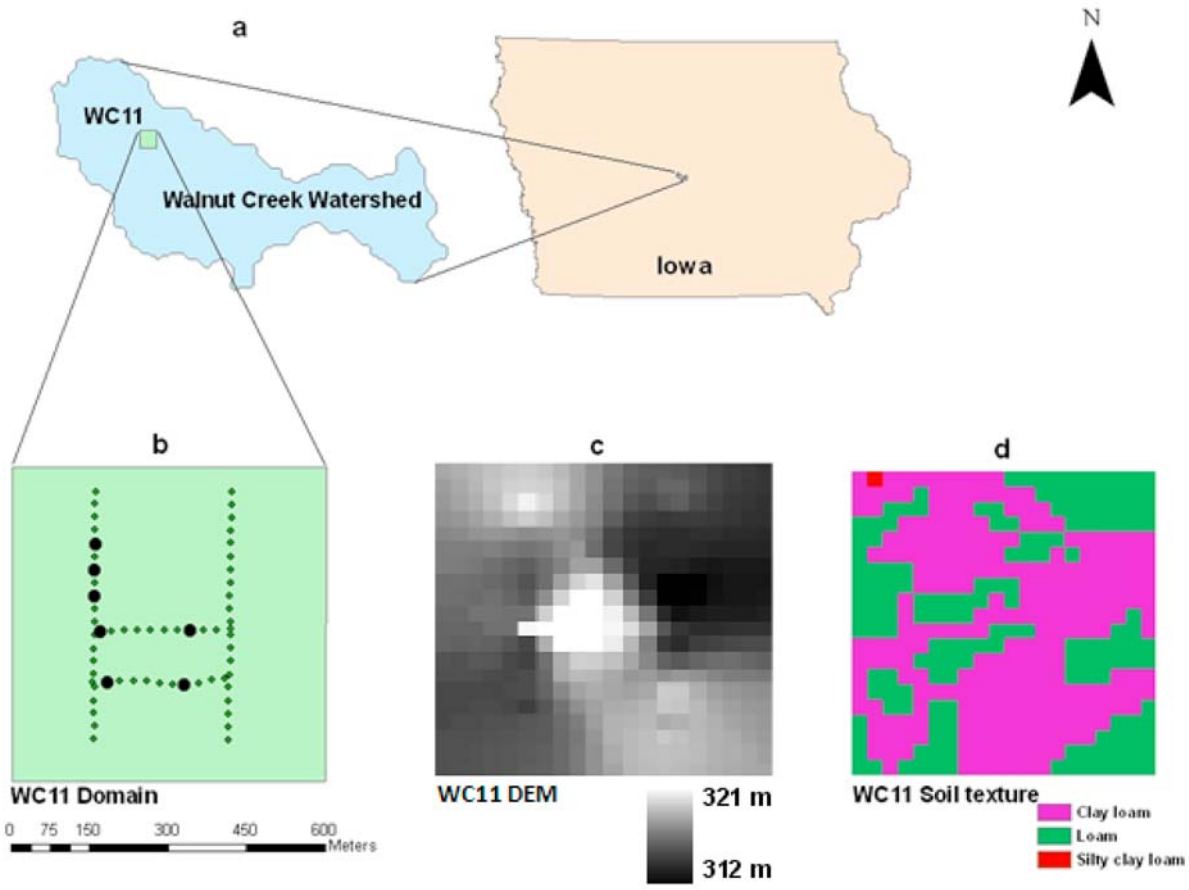

Figure 1. (a) Study area (41.97322° latitude, 93.6935° longitude). (b) Study domain with soil moisture sampling transect. Large dots show the locations of soil core extraction for soil hydraulic conductivity measurement in the laboratory. (c) Digital elevation model and (d) soil texture in the WC-11 field at the Walnut Creek watershed, Iowa.

312 m and high of 321 m with ±2 m errors. The visual observation of the terrain looks almost similar with the SRTM data set. This makes the topography of low relief within the domain as typically defined in this region. The representative soils of the WC11 as illustrated in Figure 1, are clay loam and silty clay loam of glacial origin. Meteorological data was obtained from USDA-NRCS managed SCAN site (42.00°N and 93.74°W) near Ames, Iowa.

[7] Primarily the study was conducted based on the data obtained and conditions observed in the WC11 field during the SMEX05 campaign. Data from the SMEX02 campaign was further used to test the algorithm. Volumetric soil moisture content was measured during the SMEX02 (for 12 days) and the SMEX05 (for 10 days) campaigns at the same 62 point locations in the WC11 field. Measurements were conducted between 1100 and 1500 local time (CDST). During SMEX05, along with these 62 surface soil moisture observation points, measurements were also made in the soil profile at 7 locations (highlighted points in Figure 1) up to a depth of 30 cm. Sampling points were located at nearly 30 m intervals along four transects oriented east–west and north–south within the WC11 field (as shown in Figure 1). During the experiment, volumetric soil moisture contents were measured using theta probe (HH2 device, Delta-T, Inc.) which is well established portable soil moisture measuring device. At a particular location, three measurements, two to three meters apart were taken and averaged to get a mean soil moisture measurement. The vertical support scale of theta probe is 5 cm. The horizontal support was ensured by selecting a more representative location by visual inspection within the 30 m × 30 m area. For the purpose of validating K_{sat} fields, soil cores were also collected at seven locations (highlighted) as illustrated in Figure 1 during the SMEX05. The soil cores were extracted in the vadose zone between the land surface and 30 cm depth. The dimensions of cylindrical soil cores were 7.62 cm diameter and 7.62 cm height (volume 347.5 cm^3). The surface and 30 cm depth soil cores were collocated at seven highlighted locations (Figure 1) in x and y coordinates but differs in z (vertical) direction by 30 cm in the soil profile. First, the sample was extracted from the soil surface and then a hole was dug up to 30 cm in the soil profile to extract sample from 30 to 37 cm depth. K_{sat} measurements were made on these soil cores using constant head permeameter in the soil hydrology laboratory at the Texas A&M University. Additionally, in situ K_{sat} data at 66 spatial locations collected using Guelph permeameter at 15 and 30 cm depths by *Mohanty et al.* [1991] in a nearby field (within 3 km) with similar terrain, land use, and glacial till soil were used to further test the performance of the proposed algorithm (described in section 3) in a stochastic framework. Guelph permeameter

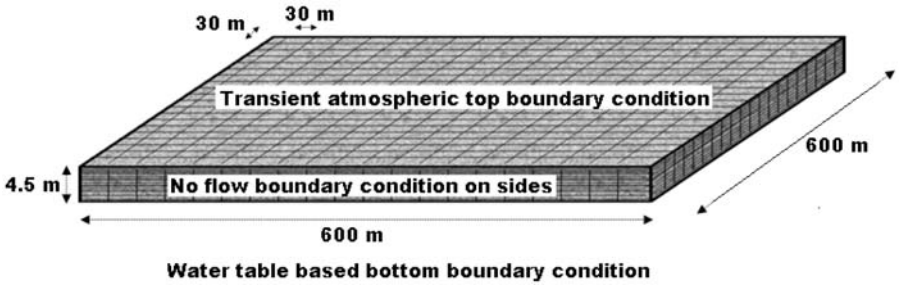

Figure 2. A 600 m × 600 m × 4.5 m three-dimensional domain configuration with horizontal discretization of 30 m × 30 m (in x and y directions) and vertical discretization comprising 55 elements (in the z direction) up to the depth of 4.5 m (not to scale) for hydrologic modeling.

typically measures saturated hydraulic conductivity of a large support volume of the soil profile (elliptical wetting bulb) of approximately 20–30 cm wide and 40–50 cm deep.

3. Methodology

[8] First, we present the parameterization of K_{sat} using Karhunen-Loève expansion. This parameterization reduces dimensionality of the uncertainty space by eliminating the modes with lower energy. Second, we present the sampling algorithm, which uses Markov chain Monte Carlo scheme.

3.1. Karhunen-Loève Expansion (KLE)

[9] Suppose the vadose zone effective K_{sat} (x, y) field is defined in the study domain $\Omega = 600$ m × 600 m of the WC11 field (Figure 2) with a depth of 4.5 m. We assume that the effective K_{sat} is known at some random spatial locations (x, y) within the domain, and the covariance of $\log(K_{sat})$ is also known. Thus, by discretizing the domain Ω into a rectangular mesh (30 m × 30 m) effective K_{sat} (x, y) is represented by a matrix, making it a high-dimensional vector. The discretization of the domain surface at 30 m × 30 m is based on the soil moisture sampling pattern in WC11 and the availability of 30 m resolution soil types and soil hydraulic parameters in the SSURGO database. Following *Loève* [1977], the KLE was used to generate effective K_{sat} field in term of an optimal L^2 basis. By truncating the expansion, we can represent the K_{sat} matrix by a small number of random parameters. To impose the hard constraints (the proposed values of the K_{sat} at prescribed locations), we find a linear subspace of the parameter space (a hyperplane) which yields the corresponding values of the K_{sat} field. For better readability, we briefly elaborate the KLE algorithm. The Karhunen-Loève expansion of a random space function $Y(x, \omega)$ is based on the spectral expansion of its covariance function $R(x, y)$, where, x and y denote spatial coordinates at different spatial points, and ω varies in probability space. Denote $Y(x, \omega) = \log[K_{sat}(x, \omega)]$ (note that hereafter "log" is used for natural logarithm), assuming $K_{sat}(x, y)$ as lognormal, where the random element ω is included for the uncertainty in K_{sat} (x, y). Suppose $Y(x, \omega)$ is a second-order stochastic process with

$$E\int_{\Omega} Y^2(x,\omega)dx < \infty, \tag{1}$$

where E is the expectation operator. Given an orthonormal basis $\{\varphi_k\}$ in $L^2(\Omega)$, we can expand $Y(x, \omega)$ as a general Fourier series

$$Y(x,\omega) = \sum_{k=1}^{\infty} Y_k(\omega)\varphi_k(x), \tag{2}$$

where

$$Y_k(\omega) = \int_{\Omega} Y(x,\omega)\varphi_k(x)dx. \tag{3}$$

[10] We are interested in the special L^2 basis $\{\varphi_k\}$ which makes the random variables Y_k uncorrelated. That is $E(Y_iY_j) = 0$ for all $i \neq j$. Denote the covariance function of Y as $R(x, y) = E[Y(x)Y(y)]$. The covariance function, being symmetrical and positive definite, has mutually orthogonal eigenfunctions that form a complete set spanning the function space to which $Y(x, \omega)$ belongs. Then such basis functions $\{\varphi_k\}$ satisfy

$$E[Y_iY_j] = \int_{\Omega}\varphi_i(x)dx\int_{\Omega} R(x,y)\varphi_j(y)dy = 0 \quad i \neq j.$$

Since $\{\varphi_k\}$ is a complete basis in $L^2(\Omega)$, it follows that are eigenfunctions of $R(x, y)$:

$$\int_{\Omega} R(x,y)\varphi_k(\mathrm{y})\mathrm{dy} = \lambda_k\varphi_k(x),\ k = 1,2,\ldots, \tag{4}$$

where $\lambda_k = E[Y_k^2] > 0$. Furthermore, we have

$$R(x,y) = \sum_{k=1}^{\infty}\lambda_k\varphi_k(x)\varphi_k(y). \tag{5}$$

Denote $\pi_k = Y_k/\sqrt{\lambda_k}$, then π_k satisfies $E(\pi_k) = 0$ and $E(\pi_i\pi_j) = \delta_{ij}$. It follows that

$$Y(x,\omega) = \sum_{k=1}^{\infty}\sqrt{\lambda_k}\pi_k(\omega)\varphi_k(x), \tag{6}$$

where φ_k and λ_k satisfy (4). We assume that the eigenvalues λ_k are ordered as $\lambda_1 \geq \lambda_2 \geq \ldots$. The expansion (equation (6)) is called the KLE. In equation (6), the L^2 basis functions $\varphi_k(x)$ are deterministic and resolve the spatial dependence of the K_{sat} field. The randomness is represented by the scalar random variables. After we discretize the domain by a rectangular mesh, the continuous KLE is reduced to finite terms. Generally, we only need to keep the leading-order terms (quantified by the magnitude of λ_k) and still capture most of

the energy of the stochastic process $Y(x, \omega)$. For an N term KLE approximation $Y_N = \sum_{k=1}^{\infty} \sqrt{\lambda_k} \pi_k \varphi_k$, define the energy ratio of the approximation as

$$e(N) = \frac{E \parallel Y_N \parallel^2}{E \parallel Y \parallel^2} = \frac{\sum_{k=1}^{N} \lambda_k}{\sum_{k=1}^{\infty} \lambda_k}. \quad (7)$$

If λ_k ($k = 1, 2, \ldots$) decays very fast, then the truncated KLE would be a good approximation of the stochastic process in the L^2 sense. Suppose the saturated hydraulic conductivity field K_{sat} is a lognormal homogeneous stochastic process, then $Y(x, \omega)$ is a Gaussian process and are independent standard Gaussian random variables. We assume that the covariance function of $Y(x, \omega)$ has the form

$$R(x, y) = \sigma^2 \exp\left(- \frac{|x_1 - y_1|^2}{2L_1^2} - \frac{|x_2 - y_2|^2}{2L_2^2} \right). \quad (8)$$

In the above formula, L_1 and L_2 are the correlation lengths in each dimension, and $\sigma^2 = E(Y^2)$ is a constant representing the variability in K_{sat}. The nature of covariance function $R(x, y)$ is based on the decay of correlation with increasing distance in x and y directions. This phenomenon is exhibited by most of the soil hydraulic parameters because they are influenced by local topography, vegetation, soil type, soil structure, and anthropogenic intervention. Gaussian decay function is favored over the other decay functions because of easy of computation and the generality involved with it. The rate of decay is subject to the covariance matrix of the intrinsic K_{sat} field; for example, the shorter the correlation length, the more terms are required in the expansion. The eigenvalues are ordered in decreasing fashion. We first solve the eigenvalue problem (equation (4)) numerically on the rectangular mesh and obtain the eigenpairs $\{\lambda_k, \varphi_k\}$. For smooth Gaussian fields used here, we can sample $Y(x, \omega)$ from the truncated KLE (6) by generating Gaussian random variables. The most important aspect of this spectral representation is that the spatial random fluctuations have been decomposed into a set of deterministic functions by multiplying random variables. If the random function $Y(x, \omega)$ is Gaussian, then the random variables form an orthogonal Gaussian vector. The Karhunen-Loève expansion is mean square convergent irrespective of the probabilistic structure of the process being expanded, provided that it has a finite variance [*Loève*, 1977]. In the simulations, we first generate a reference effective K_{sat} field using the KLE and obtain the corresponding soil moisture using a three-dimensional vadose zone mode (described in section 3.2). In our agricultural field case, to represent the discrete effective K_{sat} fields from the prior (unconditioned) distribution, we keep 16 terms in the KLE, which captures more than 90% of the energy of $Y(x, \omega)$. We propose K_{sat} at 11 distinct points i.e., randomly selecting 11 distinct points on the rectangular mesh and proposing the scalar random variables. This condition is imposed by setting

$$\sum_{k=1}^{16} \sqrt{\lambda_k} \pi_k(\omega) \varphi_k(\mathrm{x_j}) = \alpha_j, \quad (9)$$

where α_j (j = 1, …, 11) are prescribed values and $\mathrm{x_j}$ are prescribed locations. These values of α_j are used only for one realization of K_{sat} field from one instance of KLE expansion. During simulations, we propose eleven and calculate the rest five of by solving the linear system (9). Note that, L_1 and L_2 are assumed to be random in our simulations. For this study, KLE is implemented on 30 m × 30 m grid; however, coarsening or refining the grid resolution will impact the results by influencing the degree of freedom of the KLE-based parameterization of K_{sat} fields, and subsequently affecting the MCMC based sampling. The analysis after coarsening or refining of grid is not considered in this work. However, a numerical study that focuses on coarsening or refining of grid and related computational cost is planned as a follow-up study.

3.2. SWMS_3D Model and Domain Characteristics

[11] The generated K_{sat} fields from KLE algorithm were used in the SWMS_3D model. The SWMS_3D is a Fortran-based finite element code for simulating water flow in three-dimensional, variably saturated media [*Simunek et al.*, 1995]. This model is widely used by vadose zone hydrology community [*Javaux and Vanclooster*, 2006; *Javaux et al.*, 2006; *Lewandowska et al.*, 2004]. Water flow is solved using Richards' equation, which can be written in the conventional form as

$$\frac{\partial s}{\partial t} = \frac{\partial}{\partial x_i} \left[K \left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A \right) \right] - S, \quad (10)$$

where s is the volumetric water content [$L^3\ L^{-3}$], h is the pressure head [L], S is a sink term [T^{-1}] (root water uptake), x_i ($i = 1, 2, 3$) are the spatial coordinates [L], t is time [T], K_{ij}^A are components of a dimensionless tensor K^A representing the possible anisotropic nature of the medium, and K is the unsaturated hydraulic conductivity function [LT^{-1}] given by

$$K(h, x, y, z) = K_{sat}(x_1, x_2, x_3)\, K_r(x_1, x_2, x_3), \quad (11)$$

where K_r is the relative hydraulic conductivity (dimensionless) and K_{sat} is the principal saturated hydraulic conductivity [$L\ T^{-1}$]. According to this definition, the value of K_{ij}^A in equation (10) must be positive and less than or equal to unity. The diagonal entries of K_{ij}^A equal one and the off-diagonal entries zero for an isotropic medium. Finite element schemes are used for the discretization of the flow and transport equations and the resulting equations are solved in an iterative fashion, by linearization. Additional measures are taken to improve solution efficiency in transient problems, including automatic time step. The water content term is evaluated using the mass conservation method proposed by *Celia and Bouloutas* [1990]. For any additional information, readers are referred to *Simunek et al.* [1995].

[12] A 600 m × 600 m section of WC11 with soybean crop was set for model domain, as illustrated in Figure 2. The domain was further discretized horizontally into grids of 30 m × 30 m (in x and y directions) resulting in 400 grid blocks. The vertical discretization comprises 55 elements (in z direction) up to the depth of 4.5 m for all the 400 grid blocks. The 3-D domain configuration had 24,255 nodal points, 21,600 tetrahedral elements, and 882 boundary nodes. Nodal spacings were made relatively small at the soil surface where highest head gradients and flow velocities were expected. Atmospheric boundary condition was set for the topsoil surface of the domain. Soil surface boundary conditions may change from prescribed flux to prescribed head type conditions (and vice versa). In the absence of surface ponding, the numerical solution was obtained by

limiting the absolute value of the flux. SWMS_3D assumes that any excess water on the soil surface was immediately removed as surface runoff. Considering the small study domain size and low relief, the removal of excess water on the soil surface was a reasonable assumption with no obvious surface runoff routing. There were no prominent depressions in the study domain for ponding water, and during the precipitation event the runoff drains out off the field. However, to compensate the effect of micro-topography due to furrows in the field, a very small ponding depth was included in the numerical modeling to accommodate small water holding after a precipitation event. Such appropriate modeling measures take care of unconformities. The initial position of the groundwater was set at 3.5 m and a specified flux (Neumann) type boundary condition set at the bottom boundary. No flow boundary condition was imposed on the sides. A rooting depth of 2 cm to 50 cm from the top boundary was set with a uniform distribution of the potential water uptake. The unsaturated soil hydraulic properties including soil water retention and hydraulic conductivity were described by a set of closed-form equations [*van Genuchten*, 1980]. The surface soil texture information from SSURGO database was extended in vertical direction to form a homogeneous soil profile in each of the grid block. For this study, the van Genuchten parameters i.e., residual water content s_{res}, and saturated water content s_{sat} were assigned based on soil texture information from the SSURGO database. No stochasticity was introduced in these parameters with the assumption that they do not vary significantly at field scale. However, there are some conflicting reports in literature about the correlation among K_{sat}, α (cm^{-1}), and n (dimensionless) parameters. *Smith and Diekkruger* [1996] concluded that no significant correlation was observed among any of the characteristic parameters and suggested that most random variation in soil characteristic parameters could be treated as independent. However, *Wang and Narasimhan* [1992] indicated that K_{sat} and α were correlated with $K_{sat} \propto \alpha^2$. In this research, we study two different cases including case 1 (C1), with no correlation among K_{sat}, α, and n, and case 2 (C2), with correlation of $K_{sat} \propto \alpha^2$ having n as constant. A study conducted by *Hills et al.* [1992] demonstrated that the water retention characteristics could be adequately modeled using either a variable α with a constant n or a variable n with constant α, with a better result when α was variable. The KLE algorithm was used to obtain the stochastic K_{sat} fields, and for C2, the α fields were obtained from the corresponding K_{sat} fields. Analysis of existing soil hydraulic properties database (e.g., UNSODA) reveals a range of values for the constant of proportionality (P) in the correlation of $K_{sat} \propto \alpha^2$. The values of P between 5,000 and 75,000 cover a wide spectrum of soil types, e.g., coarse sand, loam, sandy loam, clayey loam, and clay. During simulation involving C2, for constant of proportionality, specific values between 5,000 and 75,000 were used with an increment of 10,000 to evaluate α. The simulation for both cases was conducted using SWMS_3D for 211 Julian days during 2005. The soil moisture evolution of the top 5 cm from the SWMS_3D model was subjected to the MCMC algorithm (described in section 3.3) to evaluate the vadose zone effective K_{sat} fields (at 30 m × 30 m × 4.5 m resolution).

[13] The study was conducted in a three-dimensional setup that consisted of 24,255 nodal points, 21,600 tetrahedral elements, and 882 boundary nodes of domain size 600 m × 600 m × 4.5 m that models the water fluxes and states. It was obvious from the domain characteristics that there was a lack of depthwise variation in soil hydraulic parameters because the algorithm determines effective K_{sat} for the grid blocks of 30 m × 30 m × 4.5 m that led to the determination of 400 effective K_{sat} values within the study domain. In this approach, the variability in horizontal directions was obvious but the variability in vertical direction was phased out by effectively integrating the soil hydraulic parameters. However, the algorithm compensated the lack of vertical variability in each grid block by determining a PDF of effective K_{sat} that encompassed the values expected in the vertical direction, discussed in sections 3.3 and 5.

3.3. MCMC Algorithm

[14] Bayesian methods provide a rigorous framework within which preexisting knowledge about parameters of a model can be combined with the observed data and the model outputs. Here the preexisting knowledge (priors) was the soil hydraulic properties except K_{sat} that were derived from the SSURGO database, and the effective K_{sat} realizations of the discretized field (i.e., study area) were obtained from the KLE algorithm. The observed data (likelihood) were the measured soil moisture values from WC11, and the model outputs were soil moisture predictions from SWMS_3D. The result from the Bayesian method is a probability distribution of the parameter space (posterior distribution), that summarizes uncertainty about the parameters based on the combination of preexisting knowledge and the sampled data values. The marginalization could be an intractable task in Bayesian method because of high dimensionality. We used the Metropolis algorithm within Markov chain Monte Carlo (MCMC) method with an independent sampler to sample the posterior distribution. The Metropolis algorithm [*Metropolis and Ulam*, 1949] has been widely used in Bayesian applications, because of its simplicity. Its principle can be summarized as follows: starting from a vector generated at iteration $i-1$, a new candidate vector is generated based on instrumental probability distribution. If this new vector leads to an increasing value of the target distribution (assuming symmetric probability distribution), it is accepted as the generated value at iteration i. Otherwise, the ratio between the new and the previous value of the target distribution is computed, and used as the acceptance probability of the candidate vector. In case of rejection, the generated vector at iteration i remains equal to that of iteration $i-1$. We used the Bayes rule $P(K_{sat}|s) \propto P(s|K_{sat})\,P(K_{sat})$ with a goal to sample $P(K_{sat}|s)$. Here s denotes the state variable (e.g., soil moisture) which we seek by using K_{sat}. The MCMC algorithm used in the study is described below.

[15] 1. Choose a starting point $K_{sat}(0)$ (where 0 refer to the number of iteration) and a variance σ^2.

2. Iterate $i = 1, \ldots, N_{iter}$.

2.1. Generate a proposal candidate vector K^*_{sat} from $q(\cdot \mid K_{sat}(i-1))$ through KLE that is based on $q(\cdot \mid (i-1))$, $q(\cdot \mid \varphi_k(i-1))$, and randomly selected correlation length L_1 and L_2.

2.2. Accept K^*_{sat} with probability $r = \frac{p\left(K^*_{sat}|s\right)q\left(K_{sat}(i-1)|K^*_{sat}\right)}{p(K_{sat}(i-1)|s)q\left(K^*_{sat}(i-1)\right)}$ or reject it with probability $(1-r)$ where

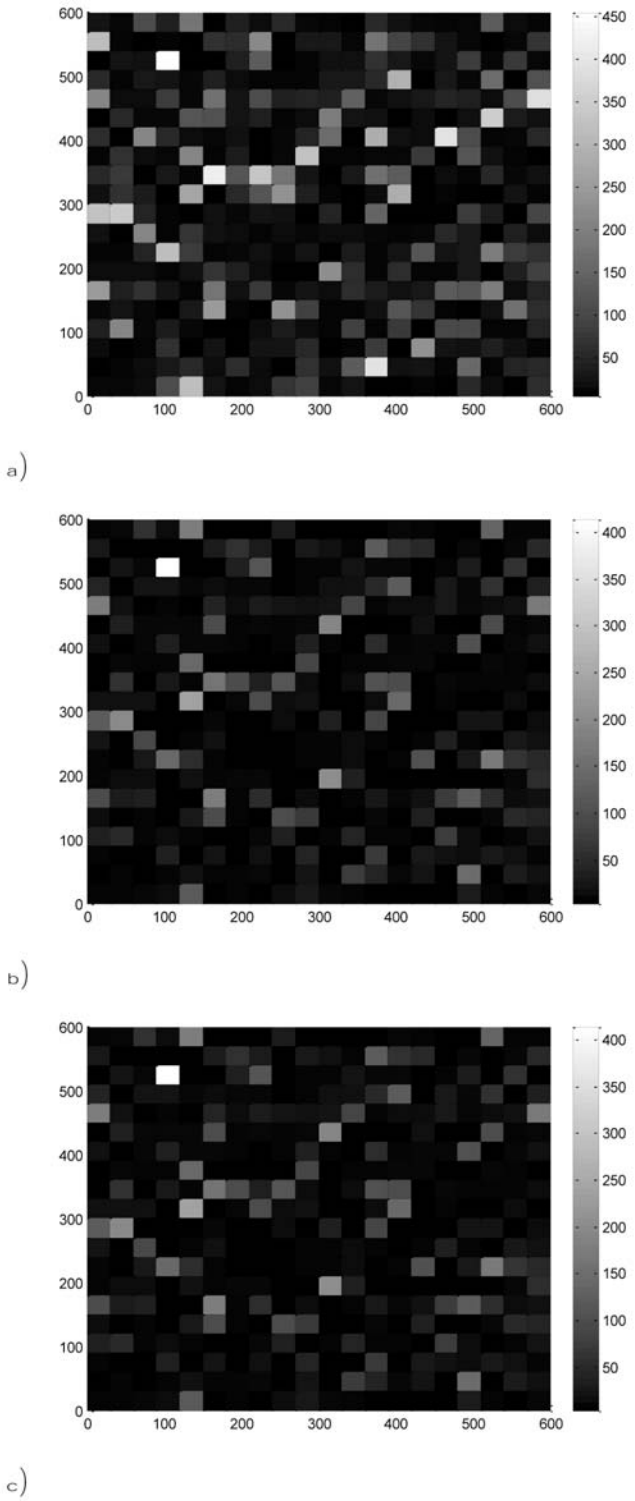

Figure 3. Generated K_{sat} (cm/d) fields of (a) one synthetic experiment and (b and c) a few accepted realizations of K_{sat} fields for case 2 (C2).

q is a realization of the saturated hydraulic conductivity field $q(\cdot \mid K_{sat}\ (i-1))$ based on proposed correlation lengths. The likelihood here is described as $P(s \mid K_{sat}) \propto \exp\left\{\frac{\left(-\sum_{i,j}\left(s_{K_{sat}}(x_i,t_j)-s_{obs}(x_i,t_j)\right)^2\right)}{\sigma^2}\right\}$, s_{Ksat} is the simulated soil moisture based on proposed K_{sat} field, s_{obs} is the measured soil moisture, and t is the time. In order to avoid numerical overflows, it is useful to consider the logarithm of the posterior distribution, and to compute the posterior ratio as $r = \exp(\log(p(K^*_{sat} \mid s)) - \log(p(K_{sat}\ (i-1) \mid s)))$. The MCMC algorithm generates a Markov chain consisting of K_{sat} whose stationary distribution is $P(K_{sat} \mid s)$. In our numerical simulations, $q(K^*_{sat} \mid K_{sat}\ (i-1))$ is a symmetric distribution (chosen by specifying the correlation lengths), $q(K^*_{sat} \mid K^*_{sat}\ (i-1)) = q(K_{sat}\ (i-1) \mid K^*_{sat})$, and thus the acceptance probability reduces to $\frac{P\left(K^*_{sat} \mid s\right)}{P(K_{sat}(i-1) \mid s)}$.

4. Synthetic Experiment

[16] To test the proposed methodology, two synthetic experiments were conducted. A well-established random field generator for hydrology i.e., HYDRO_GEN [*Bellin and Rubin*, 1996] package was used to generate K_{sat} fields with assigned covariance function for the synthetic experiments. A mean of the K_{sat} field is 200 cm/d chosen with 20% of mean as 1 standard deviation. Two realizations of K_{sat} fields having correlation lengths of 100 m from HYDRO_GEN were selected, and one of them is illustrated in Figure 3 for the case C2. The selection of 100 m is arbitrary to see whether the algorithm captures the stochasticity and correlation length of K_{sat}. The domain size of synthetic K_{sat} fields was kept same as the study area for the sake of consistency. For the synthetic fields, other soil hydraulic properties (i.e., saturated water content, residual water content, α, and n) were also assigned from *Carsel and Parrish* [1988] corresponding to the synthetic K_{sat} fields with a random white noise of 20% standard deviation from the mean values of the respective soil parameters. Profile volumetric soil moisture was simulated for these synthetic domains based on artificially created meteorological data (i.e., created by introducing additional noise in the meteorological data of SCAN site used for the study), vegetation and boundary conditions (as defined in section 3.2). Profile soil moisture values were sampled from the surface by adding noise (10% of mean as one standard deviation) to represent the field measurements. The boundary conditions were kept similar as mentioned in section 3.2.

[17] The KLE-MCMC algorithm mentioned in sections 3.1 and 3.3 was applied on the synthetic setup for each case. Only case C2 is discussed here. Figures 3b–3c show some of the K_{sat} realizations from the KLE-MCMC algorithm for the synthetic field shown in Figure 3a. For the two synthetic experiments, nearly 2000 samples of K_{sat} field were tested. The sum-of-square-root errors (SSQ) stabilized after nearly 500 samples. A total of 100 K_{sat} fields were selected based on acceptance probability after accounting for MCMC burn-in period. An acceptance rate of nearly 10% was observed in the MCMC algorithm. From the 100 selected K_{sat} fields from each experiment, probability distribution functions (PDF) were obtained and compared with the PDF of the synthetically generated truth (K_{sat} field). In most of the grid blocks of the synthetic domain an average error of nearly 5–6% was observed when compared for mean, variance and the correlation lengths. It is clearly visible from the scatterplot (Figure 4) that for most of the grid blocks simulated mean K_{sat} is very close to the reference. However, some anomalies are also observed especially for the grid blocks having high K_{sat}.

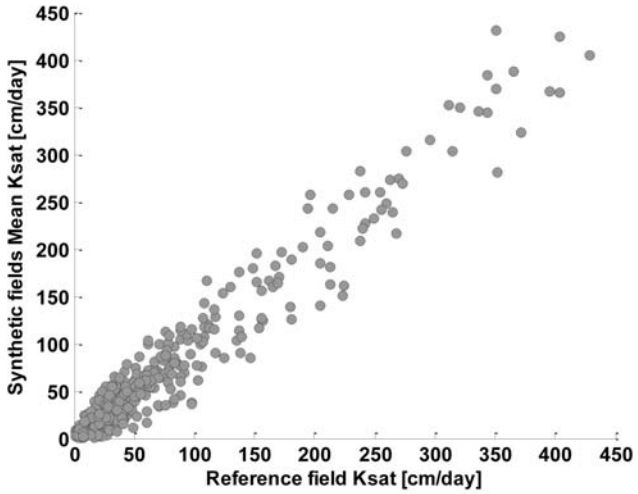

Figure 4. Scatterplot of simulated mean K_{sat} versus the reference (truth) for the synthetic experiment for C2.

[18] Synthetic experiments were conducted to assess the efficacy of the KLE-MCMC algorithm. The results were favorable that encouraged us to implement the algorithm to real field conditions. However, a caveat associated with the synthetic experiment is the controlled environment that may lead to favorable outcome.

5. Results and Discussion

[19] In this study, we tested for 10,000 samples of K_{sat} field for each case (C1 and C2) for the 600 m × 600 m × 4.5 m model domain of WC11 field in Walnut Creek watershed (Figure 2). The range of $\log(K_{sat})$ in these realizations are about $-5.0 \sim +5.0$, which means that the difference between the hydraulic conductivities are up to 4 orders of magnitude. The statistics (mean, variance, and directional correlation length) of the generated realizations are also evaluated for the study. Each realization of the conductivity field is then used to estimate pressure head and soil moisture fields from the SWMS_3D model with governing equations and boundary/initial conditions, i.e., (10) and (11) for both of the cases (C1 and C2). The soil moisture evolutions of these K_{sat} realizations are compared with the measured soil moisture values within the MCMC algorithm. The acceptance for these K_{sat} samples was nearly 2% (i.e., 200 from 10,000 samples). Taking into consideration of the number of parameters involved in the MCMC, stochasticity involved in K_{sat} fields, and field-scale physical process dynamics, a low acceptance rate is expected. Note that in the real field conditions greater number of sample K_{sat} fields are required in MCMC than the synthetic experiments. The synthetic experiments are controlled, therefore, the presence of field dynamics, heterogeneity in hydrologic processes are not captured during the synthetic experiments. In our simulations, we propose a realization of the saturated hydraulic conductivity field $q(\cdot \mid K_{sat}(i-1))$ based on proposed correlation lengths. The low success rate was expected because of the possibility of infinite number of combination of K_{sat} fields that could be generated from KLE algorithm. In our study, the Geweke test [*Geweke*, 1992] and the required sample size (200 fields to adequately represent the spatial variability) of K_{sat} field are the criteria used to determine length of the MCMC chain. The first hundred accepted samples are discarded to eliminate influence of starting proposal distribution and account for burn-in. The burn-in, i.e., discarding first 100 samples, is based on standard MCMC practice. The burn-in also helps to eliminate any dependence on starting proposal distribution. Figure 5 illustrates the residuals of two hundred selected samples for C1. Similar plot for C2 was also observed (not shown here). During the simulation, residuals are computed using $(\sum_{ij}(s_{K_{sat}}(x_i, t_j) - s_{obs}(x_i, t_j))^2)$. Note that $s_{obs}(x_i, t_j)$ is the soil moisture observation on the soil surface or at the 30 cm depth (for 7 locations). The residual decreases initially and then becomes asymptotic to exhibit some level of convergence.

[20] Few examples of accepted K_{sat} fields for WC11 are shown in Figures 6 and 7 for C1 and C2, respectively. All these different K_{sat} fields demonstrated potential to describe same soil moisture dynamics with the applied initial and boundary conditions for C1 and C2. The plots in Figures 6 and 7 also illustrate the uncertainty associated with K_{sat} in an agricultural setup. Note that given the conditions of the WC11, meteorological forcings, boundary conditions, and soil types the numerical results show effective K_{sat} fields that are capable to produce similar surface soil moisture dynamic conditions. The KLE-MCMC algorithm represents this in the

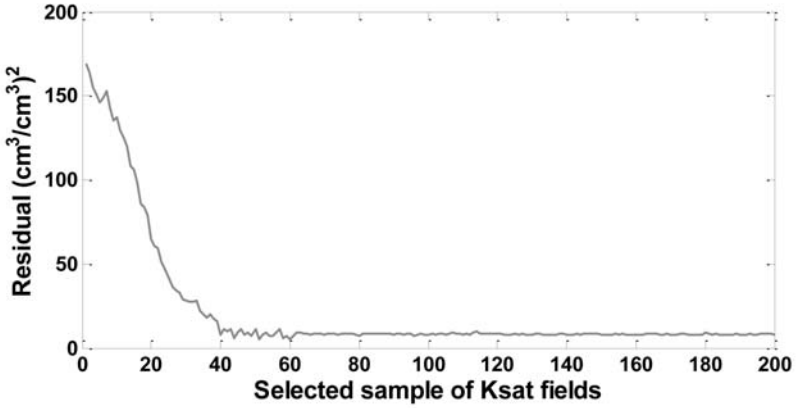

Figure 5. Residuals of accepted K_{sat} (cm/d) fields during SMEX05 from the MCMC for case 1 (C1) having no correlation among K_{sat}, α, and n.

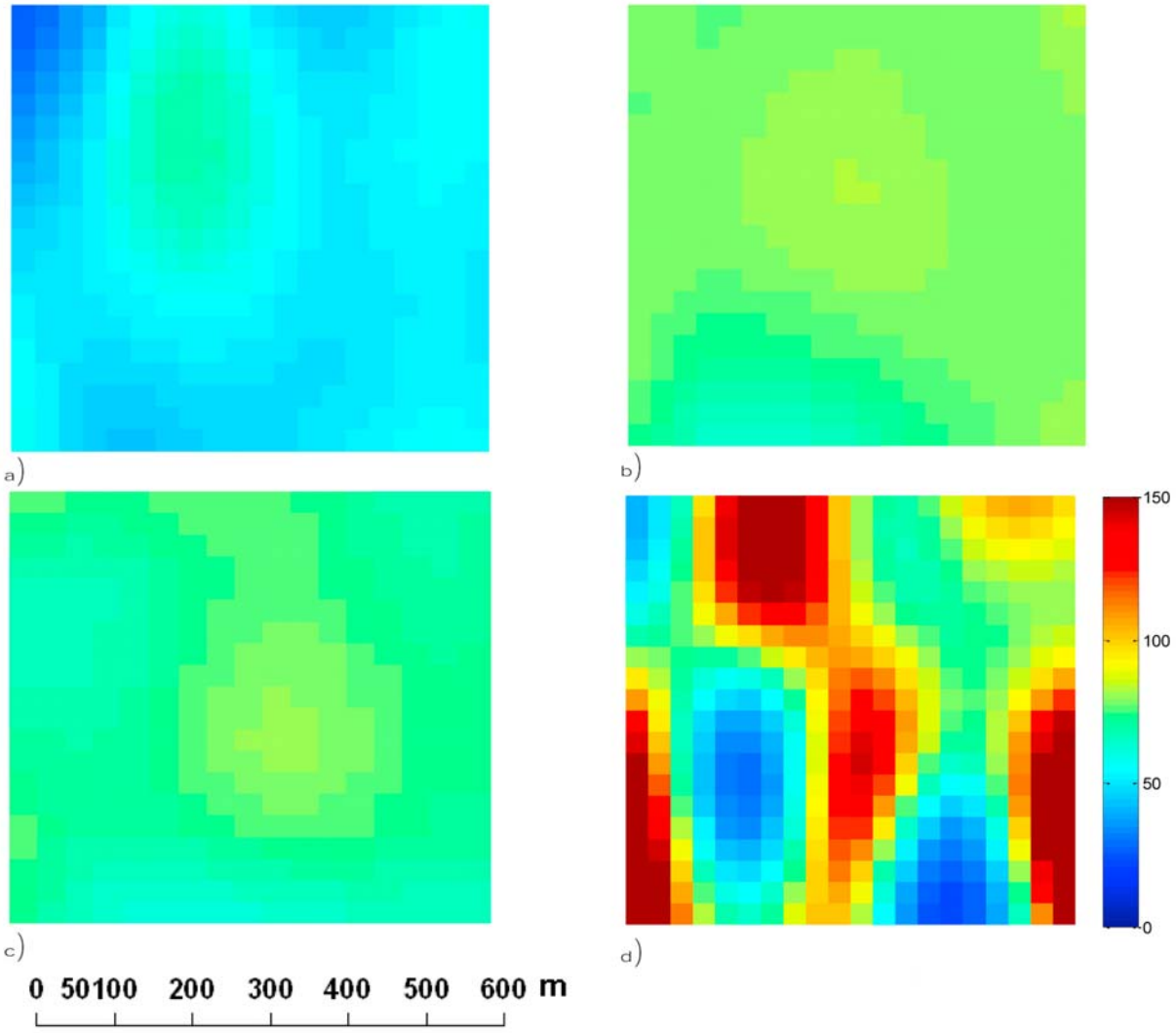

Figure 6. (a–d) Four randomly selected highly probable K_{sat} (cm/d) fields during SMEX05 for C1 having no correlation among K_{sat}, α, and n.

form of K_{sat} field spatial stochasticity, and hence, resulting in some cases of nonuniqueness. Uncertainties in hydraulic conductivity in agricultural fields were also reported in other studies. For example, in a silty loam soil, *Coutadeur et al.* [2002] and *Logsdon and Jaynes* [1996] found that tillage increased the spatial variation of K_{sat} measured with a disc infiltrometer. In WC11, uncertainty in K_{sat} fields is influenced by clustering of glacial till soil. This clustering phenomenon is due to differential deposition of soil layers by glacial drifts during its formation or subsequent incremental deposition due to wind drifts [*Mohanty et al.*, 1991]. The impact of soil type (hydraulic properties), and influence of lateral flow due to high water table/water content during wet conditions also dictate considerably the realization of stochastic K_{sat} fields from the employed algorithm. This highlights the need for researchers to use caution when using deterministic K_{sat} data as model input because of stochastic nature of K_{sat}. In addition, model users need to consider both the variability/uncertainty of K_{sat} data associated with specific soil and field location and understand how the method of determination may influence its value.

[21] The log saturated hydraulic conductivity $Y(x) = \log(K_{sat}(x))$ in the model domain is second-order stationary with a separable exponential covariance function (equation (8)). Various combinations of correlation length L_1 and L_2 in the horizontal directions (x and y axes) from equation (8) were obtained for highly probable realizations of K_{sat} fields. The highly probable K_{sat} fields have similar soil moisture dynamics and that lead to high acceptance probability in MCMC. This exemplifies the stochastic nature of K_{sat} that gives similar soil moisture evolution in WC11 for the study period. It is noteworthy that the mean horizontal correlation length of the K_{sat} is nearly 0.5 times of domain length (~300 m) in either direction (L_1 and L_2) for C1. However, in case of C2 a mean correlation length is nearly 0.65 times of domain length (i.e., 400 m). The increase in correlation length for C2 is attributed to the characteristics of van Genuchten α fields corresponding to K_{sat} fields that introduces more flexibility i.e., degree of freedom in describing the three-dimensional flow of water in the soil profile. In other words, in addition to saturated hydraulic conductivity, soil water retention functions (and unsaturated hydraulic conductivity functions [e.g., *Shouse and Mohanty*, 1998] are capable to explain the nature

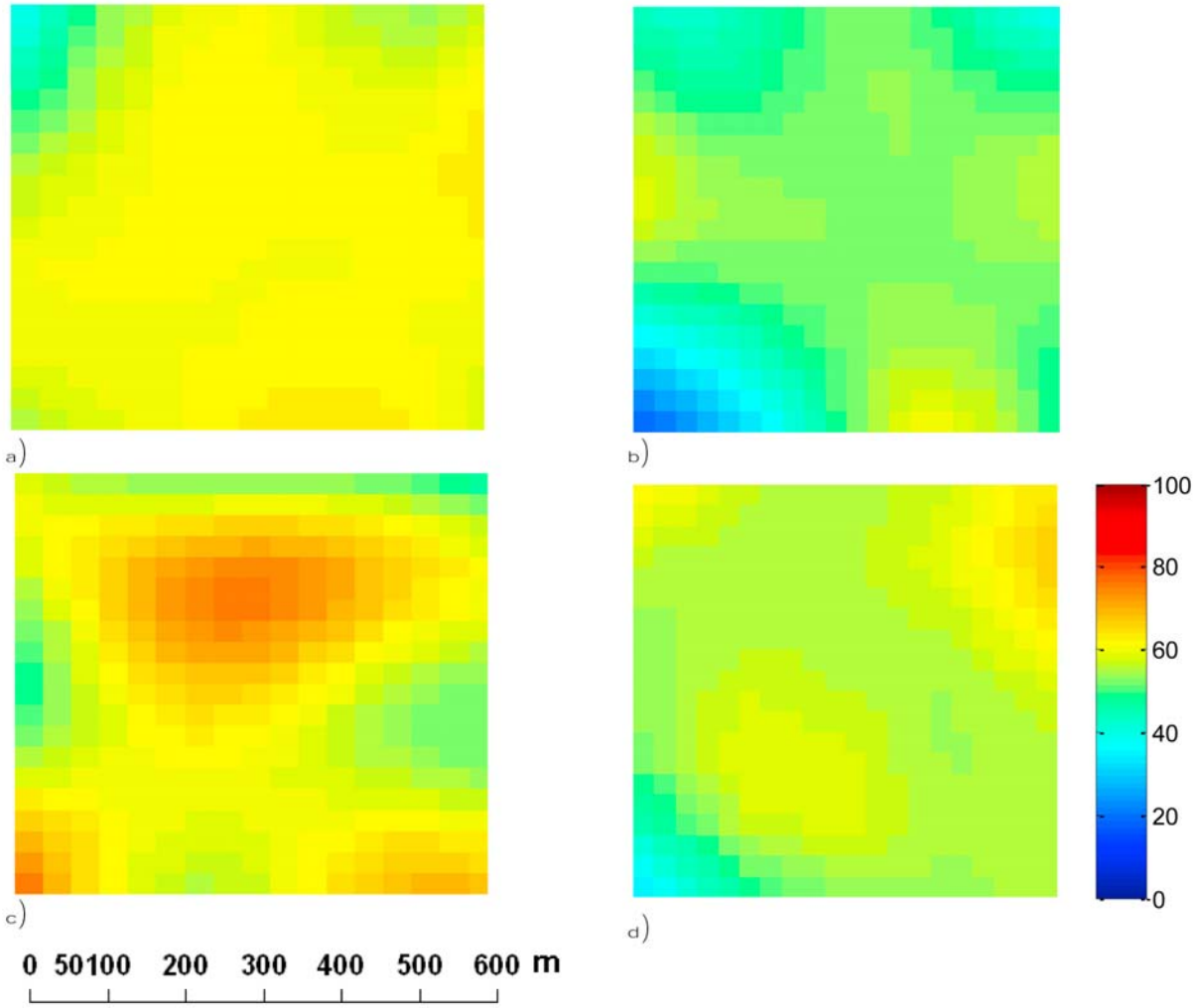

Figure 7. (a–d) Four randomly selected highly probable K_{sat} (cm/d) fields during SMEX05 for C2 having correlation of $K_{sat} \propto \alpha^2$ and n as constant.

of water movement in the soil profile within the study domain. Close examination of the domain (Figure 1d) also reveals the spatial continuity of soil type in the range of 0.2–0.7 times of the domain length. Under similar field conditions, *Mohanty et al.* [1991] reported a nested correlation structure with correlation lengths ranging from 23 m to 60 m and attributed the effect to the agricultural practices and soil of the same origin and/or their depositional patterns. One major difference between this study and *Mohanty et al.* [1991] is the treatment of K_{sat} field. This study evaluates effective saturated hydraulic conductivity for the whole soil profile or vadose zone (up to 4.5 m) in a grid block, whereas, the results of *Mohanty et al.* [1991] are particularly based on the measurements at 15 cm and 30 cm depths. The study highlights the underlying horizontal correlation length of the whole soil profile (0–4.5 m) instead of the spatial correlation on the soil surface.

[22] A mean of 20,000 for constant of proportionality P observed for C2, also concurs with the characteristics of the predominant soil types (clay loam and loam) within the domain. For C1, the 100 accepted K_{sat} fields show a range of 45 cm/d to 400 cm/d. In case of C2, the 100 K_{sat} fields have a much smaller range of 45 cm/d to 80 cm/d. This corroborates to the fact that the inclusion of parameter correlation (i.e., $K_{sat} \propto \alpha^2$) reduces the uncertainties (nonuniqueness) attached to K_{sat} within the domain. The high correlation length (L_2) for C2 also indicates this feature. Note that the influence of topography on correlation structure of K_{sat} is not apparent in this study because the relief of the study domain is low and the hydrologic modeling was conducted for a flat surface. The selection of study domain was made to avoid the influence of topography on K_{sat} characterization.

[23] Undisturbed soil cores for K_{sat} measurement collected at different locations along the soil moisture sampling transects in WC11 are shown in Figure 1b. The soil cores were taken at the land surface and 30 cm depth. The purpose of measuring K_{sat} of these soil cores was to evaluate the representativeness of K_{sat} characterization and the performance of the KLE-MCMC algorithm. This is because in the MCMC algorithm K_{sat} is treated as a stochastic parameter based on dominant soil texture from the SSUROGO 30 m resolution data for a numerical grid block of 30 m × 30 m × 4.5 m, hence, represented as a distribution without any formal aggregation or upscaling. In other words, we assumed that the aggregation protocol used for developing the SSURGO database at 30 m resolution by USDA-NRCS is good enough

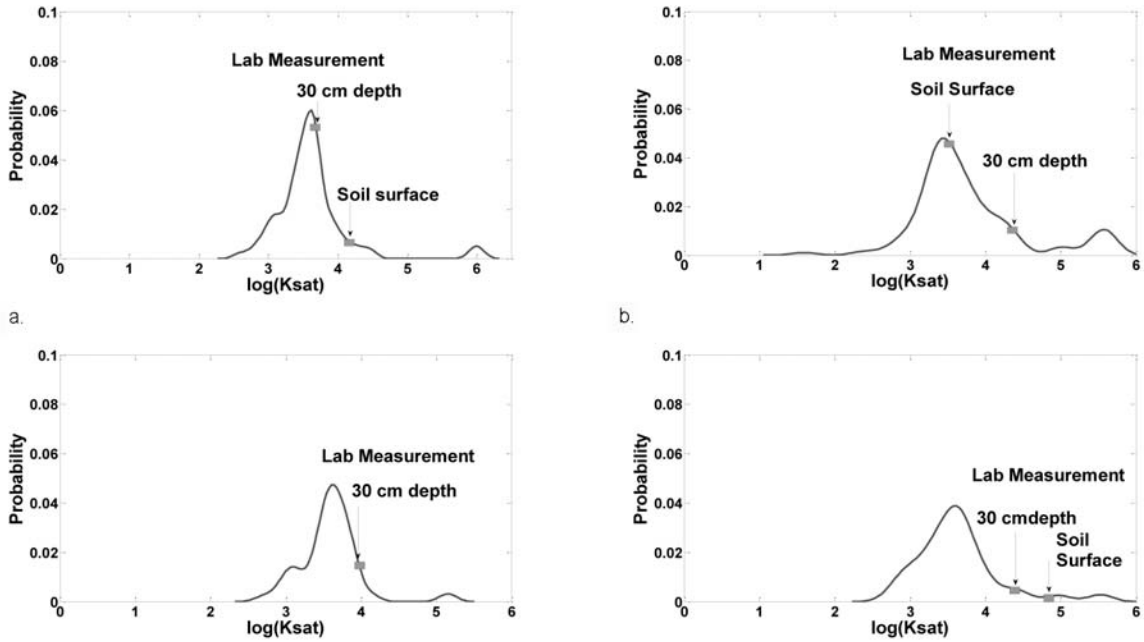

Figure 8. Probability distribution functions of K_{sat} (cm/d) at specific locations during SMEX05 for comparison with field measurements for C1 having no correlation among K_{sat}, α, and n.

for this purpose. For comparative evaluation between the laboratory measurements and simulated values, the distribution of K_{sat} realizations at seven particular sampling locations (Figure 1b) were extracted from the accepted stochastic K_{sat} fields. The distributions of $\log(K_{sat})$ from all the accepted 100 realizations of overlaying grid cell for the four distinct locations are shown in Figures 8 and 9 for C1 and C2, respectively, along with the lab measured K_{sat} at the same locations. The PDFs also exhibit the amount of uncertainty involved in K_{sat} within the particular grid cells of the domain. Few interesting features of K_{sat} are apparent in Figures 8 and 9. Noteworthy among them is very similar distributions with nearly same mean, standard deviation, and range for all the four locations. Similar statistics were also obtained for the remaining three locations not shown here. A distinct difference is visible between C1 (Figure 8) and C2 (Figure 9); that is, the overall variability is greatly reduced in case of C2. This demonstrates the decrease in uncertainty is due to parameter

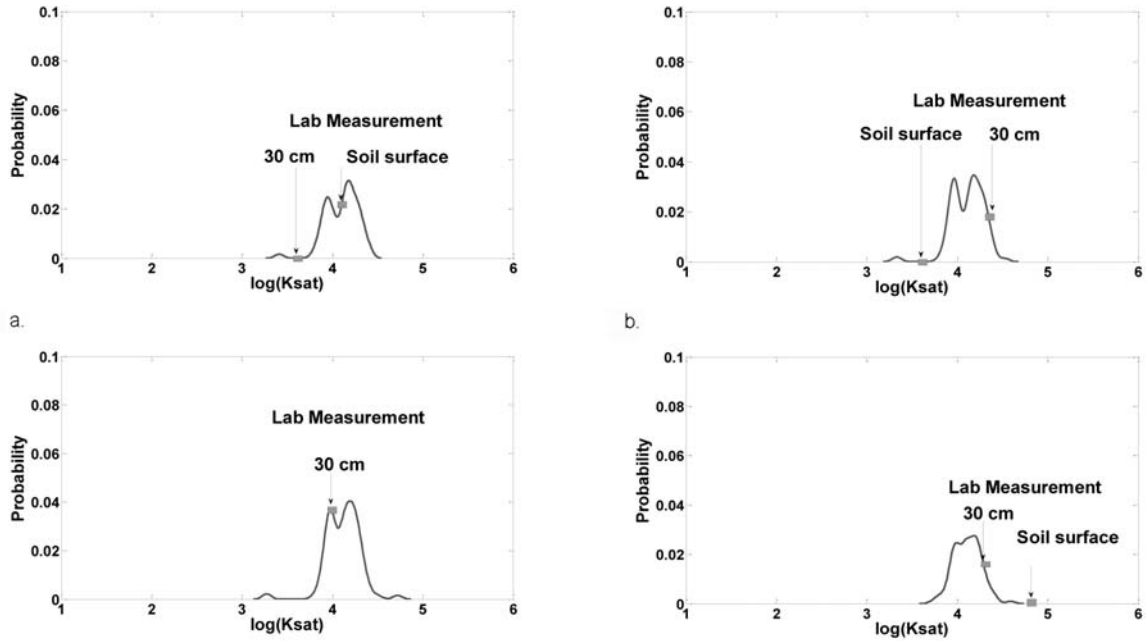

Figure 9. Probability distribution functions of K_{sat} (cm/d) at specific locations during SMEX05 for comparison with field measurements for C2 having correlation of $K_{sat} \propto \alpha^2$ and n as constant.

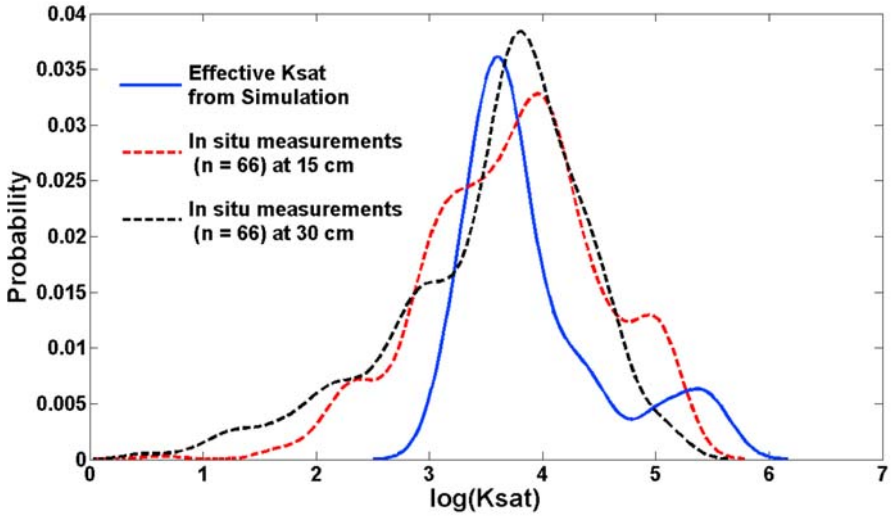

Figure 10. Probability distribution function from all 100 realizations of the study domain for C2 plotted against the probability distribution functions of in situ measurements of K_{sat} [*Mohanty et al.*, 1991] from nearby fields using the Guelph permeameter technique.

correlation as $K_{sat} \propto \alpha^2$. The multimodality feature in these probability density functions PDFs (Figures 8 and 9) should not be confused with the lognormal distribution of K_{sat} for a horizontal domain size of 600 m × 600 m with a spatial resolution of 30 m × 30 m that is obtained from the KLE part of the algorithm. The bimodal distributions (Figures 8 and 9) come from the K_{sat} values of a particular grid block. The K_{sat} values of the grid block are from different realizations of the accepted K_{sat} fields. The accepted K_{sat} fields suffer from nonuniqueness that leads to such multimodal distribution for a grid block of dimension 30 m × 30 m × 4.5 m. Laboratory measurements of K_{sat} for these locations also demonstrate distinct characteristics. Except in one location (Figures 8b and 9b), the lab measured K_{sat} for soil surface is greater than that of the K_{sat} at 30 cm depth. This phenomenon is usually observed as the K_{sat} is expected to decrease with depth because of the changes in soil density, texture, and structure. One obvious feature in Figures 8 and 9 (except in Figure 8a) is the high probability of measured K_{sat} at 30 cm depth than at the soil surface, making it a more reliable estimate. In most of the sampling locations, the measured value of K_{sat} at 30 cm depth is closer to the mean of the predicted K_{sat}. No soil core samples were collected in WC11 below the depth of 30 cm due to time and cost limitations associated with the sampling protocols during SMEX05. From the above discussion, it is apparent that samples from deeper depths would further enhance the testing of the numerical results. The low probability associated with the soil surface K_{sat} is due to higher measured values as compared to the predictions. Higher measured values could be attributed to the presence of biological macropores and structural fractures in the core samples collected near the soil surface. Under field conditions, such macropores would be expected to terminate in the subsoil because of the decrease in porosity with depth, from soil swelling, and clay hydration. At a sampling site (Figures 8b and 9b) if the lab measured K_{sat} on soil surface is lesser than the K_{sat} at 30 cm depth, it highlights another dimension of variability. Anthropogenic manipulations (e.g., tractor tracks, trails) that lead to soil compaction may explain such occurrences. The presence of biomass and high concentration of organic matter could also impede water conduction at the soil surface. *Bouma* [1980] also suggested a representative elementary volume of 10,000 cm^3 for K_{sat} measurement in clay soil for minimizing experimental variability in the laboratory. However, the soil cores used for this study were much smaller than the representative elementary volume. The PDFs of the KLE-MCMC predicted vadose zone effective K_{sat} (Figures 8 and 9) in the selected grid cells (of 30 m × 30 m × 4.5 m resolution) within the model domain that encompassed the laboratory measured values reflect the existence of variability. For comparison of vadose zone effective K_{sat} of a grid block with the lab measured values, more intensive sampling for K_{sat} is required on the soil surface and at various soil profile depths to test the results completely.

[24] Additional validation was performed using in situ K_{sat} data of *Mohanty et al.* [1991]. Probability distribution functions for K_{sat} measurements at 15 and 30 cm depths were made to compare with the probability distribution of effective K_{sat} (Figure 10) of the whole study domain (600 m × 600 m) for case 2 (C2). A close agreement in terms of mean and spread is clearly visible in Figure 10. As expected the spread of effective K_{sat} is the least as compared to the spread of in situ K_{sat} measurements at 15 cm and 30 cm because the effective K_{sat} represents an integrated value for the whole soil profile to a depth of 4.5 m. This exercise further imparts validity to the newly developed KLE-MCMC algorithm and its capability to capture the stochasticity found in K_{sat}.

[25] To further assess the performance of the KLE-MCMC algorithm, conditions of the WC11 field during the SMEX02 (2002) field campaign were also applied. The importance of this assessment is to evaluate the effects of change in vegetation cover (i.e., between corn and soybean), anthropogenic manipulation (i.e., tillage) and variation in the local forcings (i.e., atmospheric conditions) on the evolution of K_{sat} fields [e.g., *Das Gupta et al.,* 2006]. The change in vegetation cover influences the status of soil moisture and ultimately modeled K_{sat} by having different evapotranspiration rates and related infiltration characteristics. Tillage and addition of biomass to

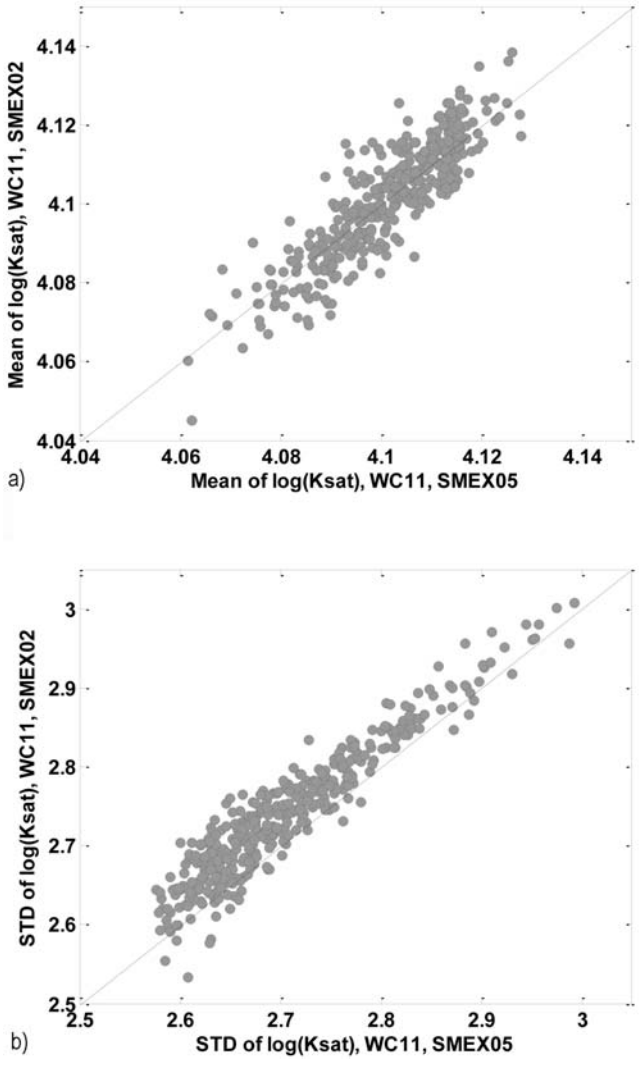

Figure 11. Scatterplots of (a) mean of log(K_{sat}) and (b) standard deviation of log(K_{sat}) at 30 m × 30 m × 4.5 m (grid cell) of the study domain for the SMEX05 and SMEX02 periods.

the topsoil also greatly manipulates the K_{sat} field. During the latter part of the SMEX02 campaign greater amount of precipitation was also observed that led to wetter conditions with very high soil moisture contents in clay textured soil. With these varied hydrologic and environmental conditions during SMEX02 and SMEX05, the vadose zone effective K_{sat} fields modeled by using the KLE-MCMC algorithm were compared. The comparison of the mean and variance of the PDFs shows statistically insignificant difference between the vadose zone effective K_{sat} fields estimated during the SMEX02 and SMEX05 periods. Even though the surface conditions were entirely different between the two years, not much influence is observed on the outcome of vadose zone effective K_{sat}. We suggest that the surface condition may affect K_{sat} near the soil surface and not the hydraulic properties of deeper soil leading to a minimal influence on the effective K_{sat} of the entire profile or vadose zone. It is also noteworthy to observe similarity in the correlation length L of the K_{sat} fields during the SMEX02 and SMEX05 periods. To quantitatively compare the accepted 100 K_{sat} realizations for the SMEX02 and SMEX05 periods, the mean and standard deviations of log(K_{sat}) were calculated for every grid block (30 m × 30 m × 4.5 m) of the study domain. Figure 11 illustrates the comparison based on every grid block within the study domain. High correlation of 0.9 (Figure 11a) and 0.95 (Figure 11b) is observed for the mean and standard deviation of log(K_{sat}), respectively, between the SMEX02 and SMEX05 periods. This suggests that the vadose zone effective K_{sat} fields have experienced statistically insignificant change over time irrespective of the change in vegetation cover and anthropogenic manipulation (tillage) of the soil surface in the WC11 field.

6. Conclusion

[26] In this paper, we characterized saturated hydraulic conductivity (K_{sat}) for an agricultural field (WC11) in the Walnut Creek watershed, Iowa. Soil moisture measurements taken during a field campaign were used to assess the vadose zone effective K_{sat}. The Karhunen-Loève expansion (KLE) with Markov chain Monte Carlo (MCMC) technique in conjunction with a physically based vadose zone model (SWMS_3D) were used for the study under two different scenarios: case 1, van Genuchten soil hydraulic parameters except K_{sat} are constant and are based on the soil type of the grid block within the modeling domain; and case 2, K_{sat} is correlated with the van Genuchten parameter α as $K_{sat} \propto \alpha^2$. The MCMC algorithm selected K_{sat} fields with a low acceptance rate, but displayed a reasonable level of convergence. The selected K_{sat} fields revealed an average spatial correlation length of nearly 400 m within a study domain of 600 m × 600 m × 4.5 m, which is quite reasonable considering the influence of clay soil texture. Laboratory measurements of K_{sat} using soil cores collected from the WC11 field at soil surface and 30 cm depth are used to test the predicted effective K_{sat} fields. The probability distribution functions (PDFs) of the predicted effective K_{sat} for both C1 and C2 scenarios encompassed the laboratory measurements and also illustrated the variability associated within the domain. In general, the uncertainty of K_{sat} reduces when the parameter correlation, $K_{sat} \propto \alpha^2$, is used. Furthermore, the laboratory measured K_{sat} values from the soil cores collected at 30 cm depth were found more representative of the field conditions in comparison with the predicted effective K_{sat} for the entire profile. Comparison of effective K_{sat} with in situ measurements at 15 and 30 cm depths from a nearby field with similar hydrologic conditions (i.e., terrain, land use, and glacial till soil) also displays good agreement and highlights the capability of the algorithm to capture stochasticity present in K_{sat} values.

[27] Application of the same algorithm in separate field conditions (i.e., different years) of the WC11 domain was used to assess the impact on effective K_{sat} fields due to different vegetation cover and local forcings. Results reveal statistically insignificant difference in the effective K_{sat} for the study domain between the two years 2002 (SMEX02) and 2005 (SMEX05). Although the study shows that the proposed algorithm performed reasonably well to characterize and parameterize the evolution of vadose zone effective K_{sat} for an agricultural field, further studies are required to evaluate the performance in larger spatial domains, for different hydroclimatic regions, and diverse topographic conditions to extend the application. We also plan to use more

efficient sampling techniques such as two-stage MCMC to sample K_{sat} in our future efforts. Other aspect of this KLE-MCMC algorithm that needs attention in future work is the quantification of computational cost with increasing model spatial domain.

[28] **Acknowledgments.** The research was funded by NSF (CMG/DMS) grant 0621113. We would like to acknowledge the partial support of LANL-SAHRA, NASA-JPL, and NASA-THP grants for this work.

References

Azevedo, A. S., R. S. Kanwar, and R. Horton (1998), Effect of cultivation on hydraulic properties of an Iowa soil using tension infiltrometers, *Soil Sci.*, *163*, 22–29, doi:10.1097/00010694-199801000-00004.

Bellin, A., and Y. Rubin (1996), HYDRO_GEN: A new random field generator for correlated properties, *Stochastic Hydrol. Hydraul.*, *10*(4), 253–278, doi:10.1007/BF01581869.

Bouma, J. (1980), Field measurement of soil hydraulic properties characterizing water movement through swelling clay soils, *J. Hydrol.*, *45*, 149–158, doi:10.1016/0022-1694(80)90011-6.

Carsel, F. C., and R. S. Parrish (1988), Developing joint probability distributions of soil water retention characteristics, *Water Resour. Res.*, *24*, 755–769, doi:10.1029/WR024i005p00755.

Celia, M. A., and E. T. Bouloutas (1990), A general mass-conservative numerical solution for the unsaturated flow equation, *Water Resour. Res.*, *26*, 1483–1496.

Coutadeur, C., Y. Coquet, and J. Roger-Estrade (2002), Variation of hydraulic conductivity in a tilled soil, *Eur. J. Soil Sci.*, *53*, 619–628, doi:10.1046/j.1365-2389.2002.00473.x.

Das Gupta, S., B. P. Mohanty, and J. M. Kohne (2006), Soil hydraulic conductivities and their spatial and temporal variations in a vertisol, *Soil Sci. Soc. Am. J.*, *70*, 1872–1881, doi:10.2136/sssaj2006.0201.

Feddes, R. A., P. J. Kowalik, and H. Zaradny (1978), *Simulation of Field Water Use and Crop Yield*, *Simul. Monogr.*, vol. 188, Pudoc, Wageningen, Netherlands.

Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, in *Bayesian Statistics 4*, edited by J. M. Bernardo et al., pp. 169–193, Oxford Univ. Press, Oxford, U. K.

Gupta, R. K., R. P. Rudra, W. T. Dickinson, and G. J. Wall (1992), Stochastic analysis of groundwater levels in a temperate climate, *Trans. Am. Soc. Agric. Eng.*, *35*, 1167–1172.

Gupta, R. K., R. P. Rudra, W. T. Dickinson, N. K. Patni, and G. J. Wall (1993), Comparison of saturated hydraulic conductivity measured by various field methods, *Trans. ASAE*, *36*, 51–55.

Hillel, D. (1980), Measurement of unsaturated hydraulic conductivity of soil profile in situ: Internal drainage methods, in *Fundamentals of Physics*, pp. 213–220, Academic, San Diego, Calif.

Hills, R. G., D. B. Hudson, and P. J. Wierenga (1992), Spatial variability at the Las Cruces trench site, in *Proceedings of the International Workshop on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, edited by M. T. van Genuchten, F. J. Leij, and L. J. Lund, pp. 529–538, Univ. of Calif., Riverside.

Javaux, M., and M. Vanclooster (2006), Three-dimensional structure characterisation and modelling of transient flow in an undisturbed variably saturated heterogeneous monolith, *J. Hydrol.*, *327*, 516–524, doi:10.1016/j.jhydrol.2005.11.064.

Javaux, M., J. Vanderborght, R. Kasteel, and M. Vanclooster (2006), Three-dimensional modeling of the scale- and flow rate-dependency of the dispersion in a heterogeneous unsaturated sandy monolith, *Vadose Zone J.*, *5*, 515–528, doi:10.2136/vzj2005.0056.

Klute, A., and C. Dirksen (1986), Hydraulic conductivity and diffusivity: Laboratory methods, in *Methods of Soil Analysis, Part 1: Physical and Mineralogial Methods*, 2nd ed., edited by A. Klute, pp. 687–734, Am. Soc. of Agron., Madison, Wis.

Lewandowska, J., A. Szymkiewicz, K. Burzynski, and M. Vauclin (2004), Modeling of unsaturated water flow in double-porosity soils by the homogenization approach, *Adv. Water Resour.*, *27*, 283–296.

Loague, K., and G. A. Gander (1990), R-5 revisited: 1. Spatial variability of infiltration on a small rangeland catchment, *Water Resour. Res.*, *26*, 957–971.

Loève, M. (1977), *Probability Theory*, 4th ed., Springer, Berlin.

Logsdon, S. D., and D. B. Jaynes (1996), Spatial variability of hydraulic conductivity in a cultivated field at different times, *Soil Sci. Soc. Am. J.*, *60*, 703–709.

Lu, Z., and D. Zhang (2004), Conditional simulations of flow in randomly heterogeneous porous media using a KL-based moment-equation approach, *Adv. Water Resour.*, *27*, 859–874, doi:10.1016/j.advwatres.2004.08.001.

McBratney, A. B. (1998), Some considerations on methods for spatially aggregating and disaggregating soil information, *Nutr. Cycling Agroecosyst.*, *50*, 51–62, doi:10.1023/A:1009778500412.

McLaughlin, D., and L. R. Townley (1996), A reassessment of groundwater inverse problem, *Water Resour. Res.*, *32*, 1131–1161, doi:10.1029/96WR00160.

Meek, B. D., W. R. Detar, D. Rolph, E. R. Rechel, and L. M. Carter (1990), Infiltration rate as affected by an alfalfa and no-till cotton cropping system, *Soil Sci. Soc. Am. J.*, *54*, 505–508.

Metropolis, N., and S. Ulam (1949), The Monte Carlo method, *J. Am. Stat. Assoc.*, *44*, 335–341, doi:10.2307/2280232.

Mohanty, B. P. (1999), Scaling of hydraulic properties of a macroporous soil, *Water Resour. Res.*, *35*, 1927–1931, doi:10.1029/1999WR900050.

Mohanty, B. P., and R. S. Kanwar (1994), Spatial variability of residual nitrate-nitrogen under two tillage systems in central Iowa: A composite three-dimensional resistant and exploratory approach, *Water Resour. Res.*, *30*, 237–251, doi:10.1029/93WR02922.

Mohanty, B. P., and Z. Mousli (2000), Saturated hydraulic conductivity and soil water retention properties across a soil-slope transition, *Water Resour. Res.*, *36*, 3311–3324, doi:10.1029/2000WR900216.

Mohanty, B. P., R. S. Kanwar, and R. Horton (1991), A robust-resistant approach to interpret spatial behavior of saturated hydraulic conductivity of a glacial till soil under no-tillage system, *Water Resour. Res.*, *27*, 2979–2992, doi:10.1029/91WR01720.

Mohanty, B. P., M. D. Ankeny, R. Horton, and R. S. Kanwar (1994a), Spatial analysis of hydraulic conductivity measured using disc infiltrometers, *Water Resour. Res.*, *30*, 2489–2498, doi:10.1029/94WR01052.

Mohanty, B. P., R. S. Kanwar, and C. J. Everts (1994b), Comparison of saturated hydraulic conductivity measurement methods for a glacial-till soil, *Soil Sci. Soc. Am. J.*, *58*, 672–677.

Mohanty, B. P., R. Horton, and M. D. Ankeny (1996), Infiltration and macroporosity under row crop agricultural field in a glacial till soil, *Soil Sci.*, *161*(4), 205–213, doi:10.1097/00010694-199604000-00001.

Nemes, A., M. G. Schaap, F. J. Leij, and J. H. M. Wosten (2001), Description of the unsaturated soil hydraulic database UNSODA version 2.0, *J. Hydrol.*, *251*, 151–162, doi:10.1016/S0022-1694(01)00465-6.

Nielsen, D. R., J. W. Biggar, and K. T. Erh (1973), Spatial variability of field measured soil-water properties, *Hilgardia*, *42*, 215–259.

Paige, G. B., and D. Hillel (1993), Comparison of three methods for assessing soil hydraulic properties, *Soil Sci. Soc. Am. J.*, *155*, 175–189.

Roy, R. V., and S. T. Grilli (1997), Probabilistic analysis of flow in random porous media by stochastic boundary element, *Eng. Anal. Bound. Elem.*, *19*, 239–255, doi:10.1016/S0955-7997(97)00009-X.

Sharma, M. L., R. J. W. Barron, and M. S. Fernie (1987), Areal distribution of infiltration parameters and some soil physical properties in lateritic catchments, *J. Hydrol.*, *94*, 109–127, doi:10.1016/0022-1694(87)90035-7.

Shouse, P. J., and B. P. Mohanty (1998), Scaling of near-saturated hydraulic conductivity measured using disc infiltrometers, *Water Resour. Res.*, *34*, 1195–1205, doi:10.1029/98WR00318.

Simunek, J., K. Huang, and M. T. van Genuchten (1995), The SWMS_3D code for simulating water flow and solute transport in three-dimensional variably saturated media, version 1.0, *Res. Rep. 139*, U.S. Salinity Lab., Agric. Res. Serv., U.S. Dep. of Agric., Riverside, Calif.

Smajstrla, A. G. (1990), *Agricultural field scale irrigation requirements simulation (AFSIRS) model*, technical manual, Univ. of Fla., Gainesville.

Smith, R. E., and B. Diekkruger (1996), Effective soil water characteristics and ensemble soil water profiles in heterogeneous soils, *Water Resour. Res.*, *32*, 1993–2002, doi:10.1029/96WR01048.

van Genuchten, M. T. (1980), A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*, 892–898.

Wager, M. G., and H. P. Denton (1989), Influence of crop and wheel traffic on soil physical properties in continuous no till corn, *Soil Sci. Soc. Am. J.*, *53*, 1206–1210.

Wang, J. S. Y., and T. N. Narasimhan (1992), Distribution and correlations of hydrologic parameters of rocks and soils, in *Proceedings of the International Workshop on Indirect Methods for Estimating the Hydraulic*

Properties of Unsaturated Soils, edited by M. T. van Genuchten, F. J. Leij, and L. J. Lund, pp. 169–176, Univ. of Calif., Riverside.

Warrick, A. W., and D. R. Nielsen (1980), Spatial variability of soil physical properties in the field, in *Application of Soil Physics*, edited by D. Hillel, pp. 319–344, Academic, San Diego, Calif.

Wierenga, P. J. (1985), Spatial variability of soil water properties in irrigated soils, paper presented at Workshop of the International Society of Soil Science and the Soil Sciences Society of America, Int. Soc. of Soil Sci., Las Vegas, Nev.

Yang, J., D. D. Zhang, and Z. Lu (2004), Stochastic analysis of saturated-unsaturated flow in heterogeneous media by combining Karhunen-Loeve expansion and perturbation method, *J. Hydrol.*, *294*, 18–38, doi:10.1016/j.jhydrol.2003.10.023.

N. N. Das and B. P. Mohanty, Department of Biological and Agricultural Engineering, Texas A&M University, College Station, TX 77843-2117, USA. (bmohanty@tamu.edu)

Y. Efendiev, Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA.