Hydrology and Earth System Sciences Discussions

Geostatistical modeling of spatial variability of water retention curves

H. Saito[1], K. Seki[2], and J. Šimůnek[3]

[1]Department of Ecoregion Science, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan
[2]Faculty of Business Administration, Toyo University, Tokyo 112-8606, Japan
[3]Department of Environmental Sciences, University of California, Riverside, CA 92521, USA

Correspondence to: H. Saito (hiros@cc.tuat.ac.jp)

HESSD

5, 2491–2522, 2008

Geostatistical modeling of water retention curves

H. Saito et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

|◄|►|
|◄|►|

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Geostatistical modeling of water retention curves

H. Saito et al.

Abstract

This study compares the performance of two geostatistical approaches, parametric and non-parametric, to evaluate the spatial distribution of water retention curves. Data used in this study were obtained from the Las Cruces trench site database that contains water retention data for 448 soil samples. In a commonly used parametric approach, three standard water retention models, i.e. Brooks and Corey (BC), van Genuchten (VG), and log-normal (LN), were first fitted to each data set. For each model, a cross validation procedure was used to estimate parameters at each sampling location, allowing computation of prediction errors. In a rarely used non-parametric approach, a cross validation procedure was first used to directly estimate water content values for eleven pressure heads at each sampling location and then the three water retention models were fitted using the same automated procedure to compute prediction errors. The results show that the non-parametric approach significantly lowered prediction errors for the VG model, while moderately reducing them also for the LN and BC models.

1 Introduction

It is well known that water flow in the vadose zone is strongly influenced by spatial variability of soil hydraulic properties and is thus subject to large uncertainties. Consequently, for example, predictions of soil moisture distributions in the vadose zone or estimates of contaminant arrival time to groundwater strongly rely on robust estimates of the spatial distribution of soil hydraulic parameters. However, since exhaustive sampling is practically impossible in most cases, we always only have an incomplete knowledge of the spatial distribution of these parameters. Geostatistical interpolation techniques are therefore used to estimate unknown values of these parameters at unsampled locations from available observations.

Unlike for water flow in saturated systems, predictions of variably-saturated water flow in soils depend not only on knowledge of saturated hydraulic conductivities, but

H. Saito et al.

also on knowledge of water retention and hydraulic conductivity characteristics that are usually described by various functional forms. The water retention characteristic for a given soil is usually obtained by measuring a series of pressure head and water content data pairs on a core sample and then by fitting the constructed discrete curve
5 using a simple common model with a closed-form function, such as the well-known van Genuchten model (van Genuchten, 1980). Predictions of water flow in soils using the numerical model are thus greatly improved when the estimated hydraulic parameters adequately represent the highly nonlinear relationships between the water content, pressure head and hydraulic conductivity. Although accurate estimation of water re-
10 tention curves (WRC) or their model parameters is rarely a goal of a study, it is one of the most crucial steps in modeling water flow and solute transport in the vadose zone. It has also been known that the soil hydraulic parameters are not only soil-type dependent, but also spatial-location dependent.

  Although some soil hydraulic model parameters have been derived in a purely em-
15 pirical way, many studies are based solely on these parameters. For example, Zhu and Mohanty (2002) averaged the widely used van Genuchten "parameters" to simulate large-scale infiltration and evaporation processes. In Oliveira et al. (2006), stochastic fields of "parameters" were generated using conditioning parameters to simulate variably-saturated water flow in soils. A common theme in these studies is that soil
20 hydraulic parameters, regardless of their derivation, can be treated as meaningful parameters. The question then arises whether or not we can consider these derived (i.e. averaged or randomly generated) parameters to be equivalent to those representing soil physical properties that are measured experimentally. To our knowledge, there have been no studies that would rigorously investigate this question.

25 This study therefore aims to quantify how well (or bad) the spatial distribution of soil hydraulic model parameters represents the spatial distribution of actual water retention curves. Two approaches are compared to analyze the spatial distribution of retention data and retention parameters. In the first approach retention curve model parameters are first evaluated at particular locations from measured retention data and these

parameters are then estimated directly for locations without measurements. In the second approach retention data (i.e. water contents for particular pressure heads) are first estimated geostatistically for locations without measurements and retention curve model parameters are then fitted to these estimated retention curves (Fig. 1). In this study, the former approach is referred to as the "parametric" (or P) approach, while the latter is referred to as the "non-parametric" (or NP) approach. The performance of P and NP approaches to estimate the spatial distribution of WRC model parameters is then evaluated.

2 Materials and methods

2.1 Water retention data and models

The soil hydraulic properties (e.g. water retention curves) used in this study were obtained from the Las Cruces Trench Site database (Wierenga et al., 1989). The database was generated as part of a comprehensive field study conducted in southern New Mexico near Las Cruces for validating and testing numerical models of water flow and solute transport in the unsaturated zone (Wierenga et al., 1991). A 24.6-m long by 6.0-m deep trench wall was excavated and a total of 450 samples were taken from nine layers. From each layer, 50 samples were taken for every 0.5 m. Soil samples (undisturbed and disturbed) were then analyzed in the laboratory for soil properties, such as the bulk density, the saturated hydraulic conductivity and the soil water retention curve. Soil water retention curves were determined at 448 sampling locations, $\boldsymbol{u}_i$, $i=1, 2\ldots448$ (Fig. 2). At each location, the water contents, $\theta(\boldsymbol{u}_i; h_j)$, were measured at eleven pressure heads, h_j, $j=1, 2\ldots11$, of 0, −10, −20, −40, −80, −120, −200, −300, −1000, −5000, and −15 000 cm H_2O. While undisturbed soil cores were used for the wet range (−300 cm to 0 cm), disturbed soil samples were used with a standard pressure plate apparatus for the dry range (−15 000 cm to −1000 cm). More details on the experimental procedures can be found in Wierenga et al. (1989). The soil charac-

Geostatistical modeling of water retention curves

H. Saito et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

⏮   ⏭

◀   ▶

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

terization and infiltration experiments at this site have been analyzed in a number of studies, including recent articles by Rockhold et al. (1996), Oliveira et al. (2006) and Twarakavi et al. (2008).

Figure 2 shows the spatial distribution of observed saturated water contents (i.e. when the pressure head is equal to 0 cm or h_1) at the site. The saturated water content data are, as expected, spatially heterogeneous with mean, maximum, and minimum values of 0.322, 0.529, and 0.218, respectively. Although not shown here, water contents measured for other pressure heads are also spatially heterogeneous. Figure 3 depicts experimental water retention curves obtained for seven different depths (z) at x=10.75 m. There are some discontinuities between water contents observed for pressure heads of −300 cm and −1000 cm due to the difference in measurement procedures (Hills et al., 1993). Not surprisingly, none of seven curves are identical. This confirms the importance of analyzing the spatial distribution of water retention curves at this site.

Water retention curves are usually approximated using one of the common analytical models. These models are then used to estimate the relationship between the (unsaturated) hydraulic conductivity and the water content using the pore-size distribution models (e.g. Mualem, 1976) as the soil-water retention curve is considerably easier and more accurate to measure than unsaturated hydraulic conductivities. Therefore, in this study, only water retention curve data are analyzed.

Among many available retention curve models, the Brooks and Corey (BC) model (Brooks and Corey, 1964), the van Genuchten (VG) model (van Genuchten, 1980), and lognormal pore size distribution (LN) model (Kosugi, 1996) are the most widely used analytical expressions to represent the dependence of the water content on the capillary pressure head for unimodal pore systems. Detailed discussions on the commonly used soil hydraulic models can be found in Leij et al. (1997). Analytical expressions of the aforementioned three closed-form models used in this study are listed in Table 1.

The parameter S_e is an effective saturation defined as follows:

$$S_e(h) = \frac{\theta - \theta_r}{\theta_s - \theta_r} \qquad (1)$$

where θ_s and θ_r are the saturated and residual water contents ($m^3\,m^{-3}$), respectively. As can be seen from Table 1, in addition to θ_s and θ_r, all three models require two other shape (or empirical) parameters, leading to a total of four parameters representing the retention curve models.

An application of these water retention functions requires reliable estimates of the parameters for a soil of interest. While θ_s is relatively easy to obtain experimentally, the other retention parameters usually have to be estimated indirectly by fitting the analytical function to the experimental water retention data using an optimization approach, such as a non-linear least-squares minimization approach (e.g. implemented in the RETC code). In most optimization procedures, the performance is improved if the initial estimates are close to the "true" solution of the inverse problem. In other words, if initial estimates are too far from the "true" solutions, the solution may not converge to the global minimum but to the local minimum of the objective function. Seki (2007) recently developed a program code that uses a full-automatic procedure to estimate soil water retention model parameters. The program automatically selects initial estimates based on observations so that the user does not have to make his/her selections. Details of the approach will not be discussed in this paper, but, it should be noted that the full-automatic approach worked very well in most cases when applied to a number of water retention data from the UNSODA database (Leij et al., 1996; Seki, 2007). Such an approach is especially attractive when a great number of retention curves needs to be fitted simultaneously.

2.2 Geostatistical theory

To estimate an unknown value of a given soil property at unsampled locations, a geostatistical estimation procedure, also known as kriging, was used in this study. Con-

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

sider the problem of estimating the value of a soil attribute z (e.g. water content or soil hydraulic parameters) at an unsampled location $\boldsymbol{u}$, where $\boldsymbol{u}$ is a vector of spatial co-ordinates. The available information consists of values of the variable z at N locations $\boldsymbol{u}_i$, $i=1, 2 \ldots N$. All univariate kriging estimates are variants of the general regression estimate $z^*(\boldsymbol{u})$ defined as:

$$z^*(\boldsymbol{u}) - m(\boldsymbol{u}) = \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha(\boldsymbol{u})\, [z(\boldsymbol{u}_\alpha) - m(\boldsymbol{u}_\alpha)] \tag{2}$$

where $\lambda_\alpha(\boldsymbol{u})$ is the weight assigned to datum $z(\boldsymbol{u}_\alpha)$ and $m(\boldsymbol{u})$ is the trend component of the spatially varying attribute (Goovaerts, 1997). In practice, only observations closest to $\boldsymbol{u}$ are retained, that is the $n(\boldsymbol{u})$ data within a given neighborhood or window $W(\boldsymbol{u})$ centered on $\boldsymbol{u}$, while the influence of those farther away are discarded (Saito and Goovaerts, 2000). One of the most common kriging estimators is ordinary kriging (or OK), which estimates an unknown value as a linear combination of neighboring observations:

$$z^*(\boldsymbol{u}) = \sum_{\alpha=1}^{n(\boldsymbol{u})} \lambda_\alpha^{OK}(\boldsymbol{u}) z(\boldsymbol{u}_\alpha) \tag{3}$$

In OK, unlike a constant value used in simple kriging, the mean (or trend) at each estimation location (i.e. local mean, $m(\boldsymbol{u})$) is implicitly re-estimated. OK weights λ_α^{OK} are determined so as to minimize the error or estimation variance $\sigma^2(\boldsymbol{u}) = Var Z^*(\boldsymbol{u}) - Z(\boldsymbol{u})$ under the constraint of an unbiased estimate (Goovaerts, 1997). These weights are obtained by solving the system of linear equations known as the ordinary kriging system:

$$\begin{cases} \displaystyle\sum_{\beta=1}^{n(\boldsymbol{u})} \lambda_\beta^{OK}(\boldsymbol{u}) = \gamma(\boldsymbol{u}_\alpha - \boldsymbol{u}_\beta) + \mu_{OK}(\boldsymbol{u}) = \gamma(\boldsymbol{u}_\alpha - \boldsymbol{u})\,\alpha = 1, \ldots, n(\boldsymbol{u}) \\ \displaystyle\sum_{\beta=1}^{n(\boldsymbol{u})} \lambda_\beta^{OK}(\boldsymbol{u}) = 1 \end{cases} \tag{4}$$

Geostatistical modeling of water retention curves

H. Saito et al.

where $\mu_{OK}(\boldsymbol{u})$ is the Lagrange parameter, $\gamma(\boldsymbol{u}_\alpha - \boldsymbol{u}_\beta)$ is the semivariogram between observations at $\boldsymbol{u}_\alpha$ and $\boldsymbol{u}_\beta$, and $\gamma(\boldsymbol{u}_\alpha - \boldsymbol{u})$ is the semivariogram between the datum location $\boldsymbol{u}_\alpha$ and the location being estimated $\boldsymbol{u}$. The semivariogram $\gamma(\boldsymbol{h})$ models the variability between observations separated by a vector $\boldsymbol{h}$. The only information required by the system (Eq. 4) is the semivariogram values, γ, for different separation distances. These values are readily derived from the semivariogram model fit to experimental values (i.e. linear model of regionalization):

$$\hat{\gamma}(\boldsymbol{u}) = \frac{1}{2N(\boldsymbol{h})} \sum_{\alpha=1}^{n(\boldsymbol{u})} [z(\boldsymbol{u}_\alpha) - z(\boldsymbol{u}_\alpha + \boldsymbol{h})]^2 \tag{5}$$

where $N(\boldsymbol{h})$ is the number of data pairs for a given separation vector $\boldsymbol{h}$. The choice of the model is limited to functions that ensure a positive definite covariance function matrix of the left-hand-side of the kriging system (Eq. 4). Spatial correlations often vary with direction, and such a case requires one to compute semivariograms in different orientations and fit anisotropic (direction-dependent) semivariogram models. Details of model fitting can be found in Deutsch and Journel (1998), Goovaerts (1997), and Kitanidis (1997).

2.3 Prediction performances: P vs. NP

To compare the P and NP approaches in terms of reproduction of WRC, $S_e(h; \boldsymbol{u}_i)$, two standard validation techniques were used: a cross-validation, in which one observation at a time is temporarily removed from the dataset and re-estimated from remaining data, and a jack-knife, in which the dataset is divided into two non-overlapping prediction and validation sets (Fig. 4). The number of observation locations held in each set is 199 (N_p) and 249 (N_v), respectively. In the jackknife technique, the prediction set is used to estimate values at all locations in the validation set so that estimation errors can be explicitly calculated. Figure 4 depicts the location map of both prediction and validation sets used in the jackknife technique. For both approaches, three

Geostatistical modeling of water retention curves

H. Saito et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

WRC models, i.e. BC, VG, and LN, were considered. For both approaches, unknown values were estimated using ordinary kriging (Eq. 3). Note that, regardless of the approach, semivariogram models used in kriging are those fitted to semivariograms for all 448 observations. Details about P and NP approaches used in this study are given below.

2.3.1 Parametric approach (P)

1. Parameters for three water retention functions (BC, VG, and LN) are first obtained for all 448 water retention curves $S_e(h; \boldsymbol{u}_i)$ using the automatic fitting procedure (Seki, 2007).

2. For each soil hydraulic parameter, unknown values are estimated at all 448 locations for the cross-validation and at 249 locations for the jackknife using ordinary kriging (OK).

3. For each model, a discrepancy between observed water retention curves $S_e(h; \boldsymbol{u}_i)$ and those calculated from estimated parameters $S_e(h; \boldsymbol{u}_i)$ is obtained at each location by computing differences in water contents at corresponding eleven pressure heads. Mean absolute error (MAE) and mean square error (MSE) are used in this study

$$\text{MAE} = \frac{1}{N} \sum_{i=i}^{N} \left[\frac{1}{11} \sum_{j=1}^{11} |\hat{\theta}(\boldsymbol{u}_i; h_j) - \theta(\boldsymbol{u}_i; h_j)| \right] \quad (6)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=i}^{N} \left[\frac{1}{11} \sum_{j=1}^{11} (\hat{\theta}(\boldsymbol{u}_i; h_j) - \theta(\boldsymbol{u}_i; h_j))^2 \right] \quad (7)$$

where N is 448 for the cross-validation and 249 for the jackknife, $\hat{\theta}(\boldsymbol{u}_i; h_j)$ is the water content at the location $\boldsymbol{u}_i$ calculated from estimated parameters for the pressure head

h_j, and $\theta(\mathbf{u}_i; h_j)$ is the observed water content at the location $\mathbf{u}_i$ for the pressure head h_j.

2.3.2 Non-parametric approach (NP)

1. Unknown water content values corresponding to eleven pressure heads are estimated using OK at 448 locations for the cross-validation and 249 locations for the jackknife.

2. Using the automatic fitting procedure (Seki, 2007), parameters for three WRC models (i.e. BC, VG, and LN) are obtained at each estimated location.

3. For each model, a discrepancy between observed water retention curves and those calculated from parameters obtained in the previous step is obtained by computing MAE and MSE as done for the P approach.

4. The impact of the number of data pairs to construct WRC when estimating parameters is investigated by repeating 1 through 3 using only the following six pressure heads: 0, -20, -80, -200, -1000, and $-15\,000$ cm. This approach is referred to as NP6 in the remainder of the paper, while the approach that considers all eleven pressure heads is referred to as NP11.

3 Results and discussions

The results are divided into two sections. The first section summarizes experimental semivariograms computed from water retention data and model parameters. In the second section, different approaches (P and NP) are compared in terms of describing spatial distribution of soil hydraulic parameters.

3.1 Semivariograms

Figure 5 shows the experimental semivariograms of WRC model parameters with fitted geometric anisotropy models. Semivariograms of water contents corresponding to eleven pressure heads, h_1–h_{11}, with fitted geometric anisotropy models are depicted in Fig. 6. All these semivariograms were calculated for 448 observations. Except for α_v there is clear anisotropy in all semivariograms with major spatial continuity observed in the horizontal direction. While horizontal semivariograms, in general, are all well structured, vertical semivariograms in most cases fluctuate a lot and are not smooth. There is not a sufficient number of pairs in the vertical direction to obtain well structured semivariograms. Existence of soil horizons would also not result in spatial correlation of soil variables in the vertical direction at the scale of observations. All experimental semivariograms were fitted using either exponential or spherical models and they all display a clear nugget effect. Ranges and sill values vary depending upon the variable. For example, as expected, semivariograms for the saturated (θ_s) and residual (θ_r) volumetric water contents are all similar for all three models (Fig. 5, top two rows). However, other parameters have semivariograms of different shapes, which was expected as the spatial variability of these parameters varies.

As for the semivariograms of water contents at given pressure heads (Fig. 6), the ranges of water contents near saturation in the horizontal direction (i.e. major range) are generally larger than those for drier conditions. This suggests that the spatial continuity of larger pores is more profound than that of smaller pores. When pressure heads are smaller than −40 cm, the shape of the semivariograms becomes almost identical. From the Laplace equation for the capillary rise, the radius of the pore corresponding to the pressure head of −40 cm can be calculated to be about 0.35 mm (assuming that the contact angle is equal to 0°). Therefore, it can be concluded that spatial distributions of pores smaller than 0.35 mm are similar in this soil.

3.2 Prediction performance

Using both parametric and non-parametric approaches, water retention curve models can be geostatistically estimated at any location. Figure 7 shows differences between P and NP approaches in terms of predicting a water retention curve at a given loca-
5 tion using the LN model. In this example, a location where observed retention data (squares) were available was selected to investigate whether or not one approach outperforms the other in terms of reproducing the WRC. Due to the exactitude property of kriging, data at this location were not included in kriging. While the blue line is obtained with parameters predicted by the P approach, the red line is calculated with parame-
10 ters obtained by the NP approach. In the P approach, model parameters were directly estimated using OK from surrounding conditioning parameters. In the NP approach, the water retention curve was geostatistically constructed by estimating water content values corresponding to eleven pressure heads, h_1–h_{11}. Geostatistically constructed water retention data are shown in Fig. 7 using triangles. The LN model (the solid red
15 line in Fig. 7) was then fitted to the predicted retention data to obtain model parameters. While the model predicted using the P approach could not reproduce the S-shape observed in retention data, the model obtained using the NP approach could capture the trend of WRC well. This confirms that depending upon the chosen approach, resulting WRC models can be significantly different. In the following section, both approaches
20 were compared in a more comprehensive manner using cross-validation.

Figures 8 and 9 depict mean absolute errors (MAE) and mean square errors (MSE) calculated for all three approaches, i.e. P, NP11, and NP6, using cross-validation and jack-knife procedures for all three WRC models. In general, the NP approach resulted in smaller prediction errors regardless of the used WRC model, except when the LN
25 model was used in the jack-knife procedure. Decreases in prediction errors were much larger when the VG model was used, regardless of the adopted validation procedure, because the prediction performance of the VG model for the P approach was much worse than that of the other two models. At one location, (x, z)=(12.75, 3.21), the fitted

parameter α_v (=28.16) was three orders of magnitude greater than the rest, where the mean of fitted α_v is 0.12 and the median is 0.04 (Fig. 10). Although the fitting itself was good at this particular location (Fig. 11), the large α_v value (outlier) affected the estimation of α_v values at surrounding locations leading to relatively poor performance of the P approach for the VG model. This kind of problem can be avoided if the NP approach is used. For the other two WRC models, average prediction errors decreased by about 10–15% when the NP approach was used compared to prediction errors for the P approach, in which there were no outliers in initially estimated parameters.

For the P approach, prediction errors vary depending upon the WRC model used. On the other hand, prediction errors do not depend upon the model used for the NP approach. This is not a surprising result since in the NP approach three models were fitted to the same reconstructed WRCs. Therefore, all prediction errors calculated for the NP approach are mainly the summary of model fitting results but do not directly reflect the accuracy of geostatistical estimation as those for the P approach.

There is no distinct trend in whether or not the NP11 approach is better than NP6 or vice versa. While NP6 outperforms NP11 when the VG model is used, NP11 has slightly smaller average errors or almost the same errors compared to NP6 for the BC and LN models. This indicates that reducing the number of $\theta-h$ pairs by about half does not necessarily worsen prediction performances as long as the non-parametric approach is adopted. This result is quite important since the main reason for the P approach to be preferred over the NP approach in most studies is that the number of variables one needs to analyze can be significantly reduced. The number of parameters required in the WRC models used in this study is four, which is significantly smaller than original eleven $\theta-h$ pairs used to construct water retention curves. To use all pairs in the NP approach, one needs to perform kriging eleven times, which requires significantly more effort than performing kriging only four times in the P approach. However, it was just shown here that when the number of data pairs is reduced to six, which is slightly more than four, the non-parametric approach is better than the parametric approach. With a little extra work required than for the P approach, we can improve the

Geostatistical modeling of water retention curves

H. Saito et al.

description of the spatial variability of water retention curves using the NP approach.

Figures 12 and 13 show spatial distributions of locations where NP11 outperformed P and vice versa in terms of MAE and MSE, respectively, in cross-validation. There is no specific trend observed, such as, for example, a cluster of white or black squares.

5   This means that the prediction of water retention curves was fairly unbiased. Proportions of white squares for both NP11 and NP6 (spatial distributions are not shown) are summarized in Table 2 for all cases. For all models, the non-parametric approaches outperformed the parametric approach at more than 60% of sampling locations. Especially for the VG model, the NP approaches resulted in smaller prediction errors at

10   more than 80% of sampling locations. Although the difference between NP11 and NP6 is small, NP11 was slightly better than NP6 for BC and LN,, while NP6 was better than NP11 for VG. Overall, these percentages confirm the superiority of the NP approach to predict WRCs.

4   Conclusions

15   This study compares the performance of two geostatistical approaches, commonly used parametric (P) and newly proposed non-parametric (NP), to estimate the spatial distribution of water retention curve parameters. In the former approach retention curve model parameters are first evaluated at sampled locations from observed retention data and these parameters are then estimated directly for locations without mea-

20   surements. In the latter approach retention data are first estimated using kriging for locations without measurements and retention curve model parameters are then fitted to these estimated retention curves. Standard validation techniques (cross-validation and jackknife) were used to compare two approaches in terms of estimation of the spatial distribution of water retention curves.

25   Both validation results showed that regardless of the retention model selected, the NP approach lowered prediction errors compared to the P approach. Improvements were especially large for the VG model because an outlier in initially estimated pa-

Geostatistical modeling of water retention curves

H. Saito et al.

rameters led to the poor estimate of that parameter at surrounding locations in the P approach. There is always a risk of including parameter outliers in the estimation process as long as the P approach is used. In the NP approach, such a problem hardly occurs. In addition, the NP approach performed better than the P approach even when the number of data pairs used to construct retention curves was reduced from eleven to six. This makes the NP approach comparable to the P approach in terms of workload because the number of variables used in the P approach is four, while that of the NP approach now becomes only 6. The P approach has been preferred mainly because one can reduce the computational time and workload. However, as this study shows, with a little bit of extra work, one can achieve much better estimation of the spatial distribution of water retention curves by using the NP approach.

Overall, this study shows the superiority of the NP approach over the P approach to evaluate the spatial distribution of water retention curves, which is the initial but critical step to model variably-saturated water flow and solute transport in field-scale heterogeneous soils. In future studies, the actual impact of the NP approach on water flow and solute transport needs to be still investigated.

References

Deutsch, C. V. and Journel, A.: GSLIB: Geostatistical Software Library and User's Guide, Oxford Univ. Press, New York, 384 pp., 1998.

Hills, R. G., Wierenga, P., Luis, S., McLaughlin, D., Rockhold, M., Xiang, J., Scanlon, B., and Wittmeyer, G.: INTRAVAL – Phase II model testing at the Las Cruces Trench Site, Tech. rep., US Nuclear Regulatory Commission – NUREG/CR-6063, Washington, D.C., 140 pp., 1993.

Goovaerts, P.: Geostatistics for Natural Resources Evaluation, Oxford Univ. Press, New-York, 483 pp., 1997.

Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, New York, 271 pp., 1997.

Leij, F. J., Alves, W. J., van Genuchten, M. T., and Williams, J. R.: The UNSODA-Unsaturated Soil Hydraulic Database. User's manual Version 1.0, Report EPA/600/R-96/095, National Risk Management Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Cincinnati, OH, 103 pp., 1996.

Leij, F. J., Russell, W. B., and Lesch, S. M.: Closed-Form Expressions for Water Retention and Conductivity Data, Ground Water, 35(5), 848–858, 1997.

Oliveira, L. I., Demond, A. H., Abriola, L. M., and Goovaerts, P.: Simulation of solute transport in a heterogeneous vadose zone describing the hydraulic properties using a multistep stochastic approach, Water Resour. Res., 42, W05420, doi:10.1029/2005WR004580, 2006.

Rockhold, M. L., Rossi, R. E., and Hills, R. G.: Application of similar media scaling and conditional simulation for modeling water flow and tritium transport at the Las Cruces Trench Site, Water Resour. Res., 32(3), 595–609, 1996.

Saito, H. and Goovaerts, P.: Geostatistical interpolation of positively skewed and censored data in a dioxin contaminated site, Environ. Sci. Technol., 34(19), 4228–4235, 2000.

Seki, K.: SWRC fit - a nonlinear fitting program with a water retention curve for soils having unimodal and bimodal pore structure, Hydrol. Earth Syst. Sci. Discuss., 4, 407–437, 2007, http://www.hydrol-earth-syst-sci-discuss.net/4/407/2007/.

Twarakavi, N. K. C., Saito, H., Šimůnek, J., and van Genuchten, M. T.: A New Approach to Estimate Soil Hydraulic Parameters using Only Soil Water Retention Data, Soil Sci. Soc. Am. J., 72(2), 471–479, 2008.

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

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

van Genuchten, M. T., Lesch, S. M., and Yates, S. R.: The RETC code for quantifying the hydraulic functions of unsaturated soils, Version 1.0, US Salinity Laboratory, USDA, ARS, Riverside, CA, 93 pp., 1991.

Wierenga, P. J., Toorman, A. F., Hudson, D. B., Vinson, J., Nash, M., and Hills, R. G.: Soil physical properties at the Las Cruces trench site, NUREG/CR-5441, Nucl. Regul. Comm., Washington, D. C., 92 pp., 1989.

Wierenga, P. J., Hills, R. G., and Hudson, D. B.: The Las Cruces trench site: characterization, experimental results, and one-dimensional flow predictions, Water Resour. Res., 27, 2695–2705, 1991.

Zhu, J. and Mohanty, B. P.: Spatial averaging of van Genuchten hydraulic parameters for steady-state flow in heterogeneous soils: a numerical study, Vadose Zone J., 1, 261–272, 2002.

Geostatistical modeling of water retention curves

H. Saito et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Water retention, S_e, models used in this study.

Models	Expression	Parameters
Brooks and Corey	$S_e(h) = \begin{cases} 1 & h \leq h_b \\ (\frac{h}{h_b})^{-\lambda} & h > h_b \end{cases}$	h_b [cm] λ [−]
van Genuchten	$S_e(h) = \frac{1}{[1+|\alpha h|^n]^m}$	n [−] α g cm^{-1}] m (=1−1/n) [−]
Kosugi	$S_e(h) = \frac{1}{2} erfc \left\{ \frac{\ln(h/h_l)}{\sqrt{2}\sigma} \right\}$	h_l [cm] σ [−]

Table 2. Percentages (%) of locations where predictions errors (MAE and MSE) for the NP approach were smaller than those for the P approach in cross-validation.

Model	BC		VG		LN	
	NP11	NP6	NP11	NP6	NP11	NP6
MAE	71.7	70.1	81.0	82.1	62.1	59.4
MSE	71.4	68.8	80.4	82.4	62.7	60.7
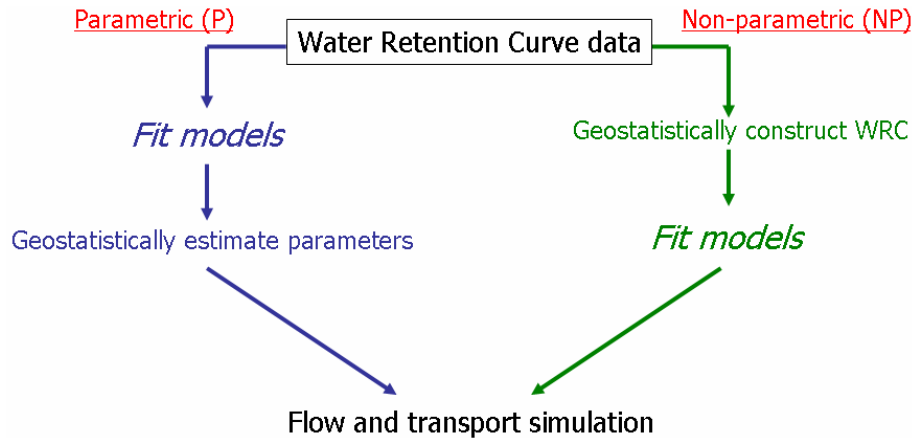
Fig. 1. A difference between the parametric and non-parametric approaches used in this study.
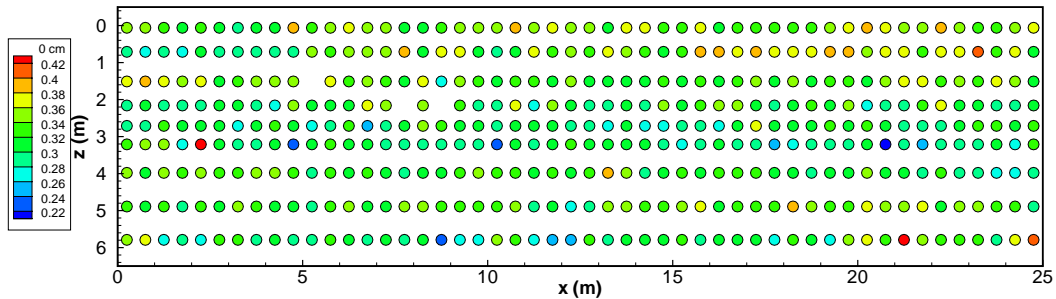
Fig. 2. Location map of saturated water content data (h=0 cm) at the Las Cruces Trench site. The vertical axis represents the depth from the surface.
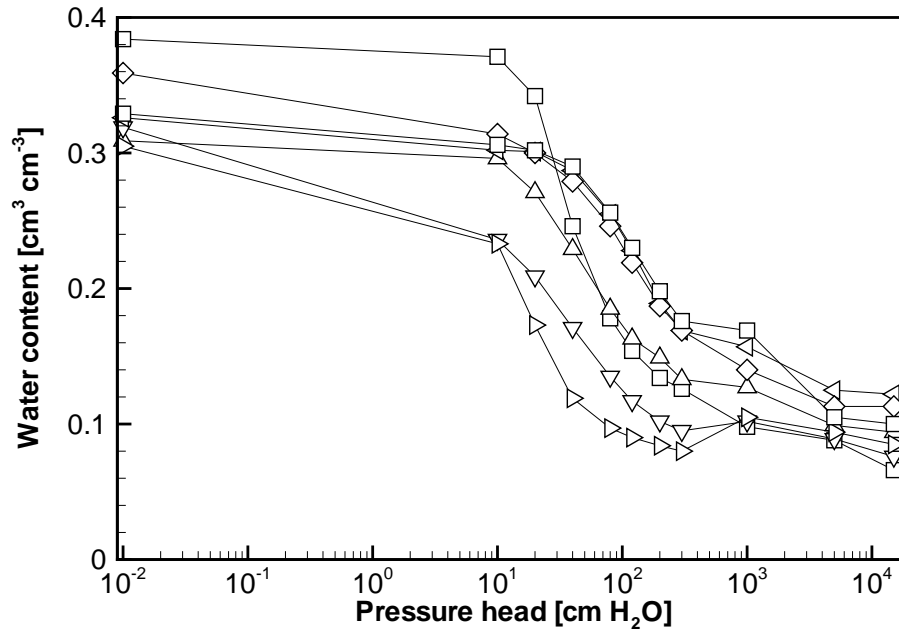
Fig. 3. Water retention curves obtained for seven different depths at x=10.75 m.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
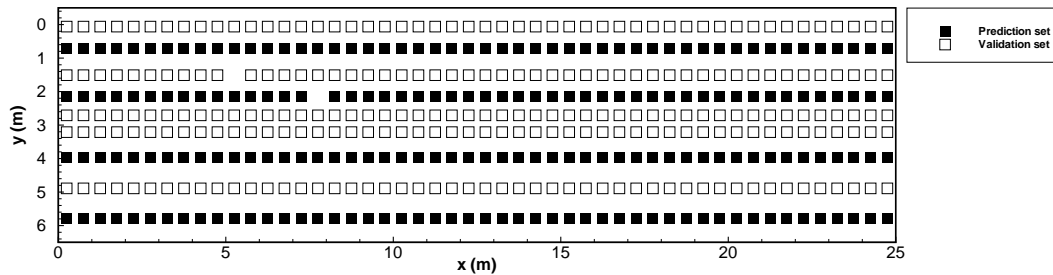
Fig. 4. Locations used for prediction (closed squares) and validation (open squares) purposes in the jackknife procedure.
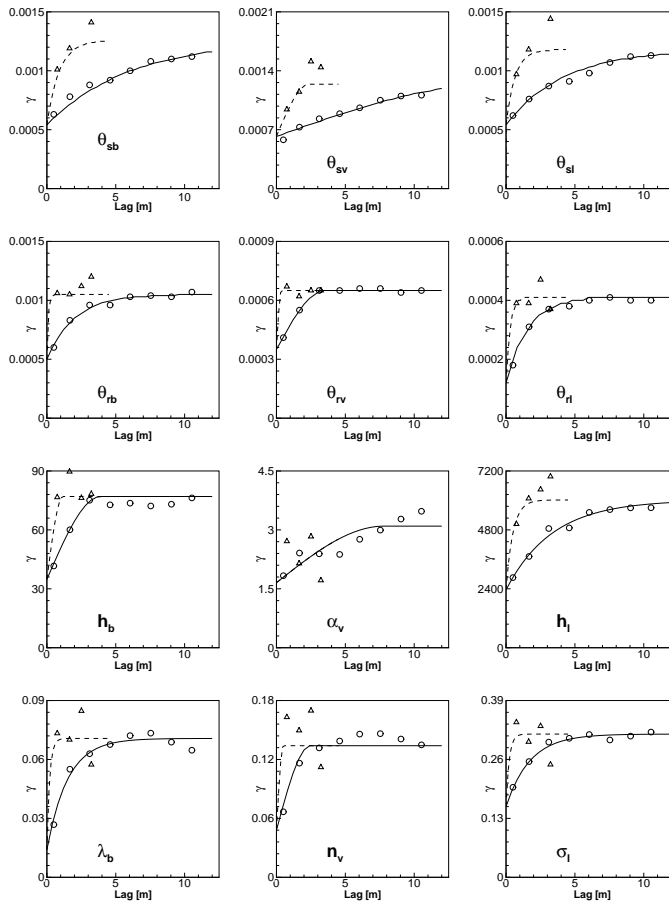
Fig. 5. Experimental semivariograms of water retention curve model parameters with fitted geometric anisotropy models to horizontal (circle, solid line) and vertical (triangle, dashed line) directions. Semivariograms were calculated from model parameters at 448 sampling locations. Subscripts b, v, and l correspond to BC, VG, and LN models, respectively.
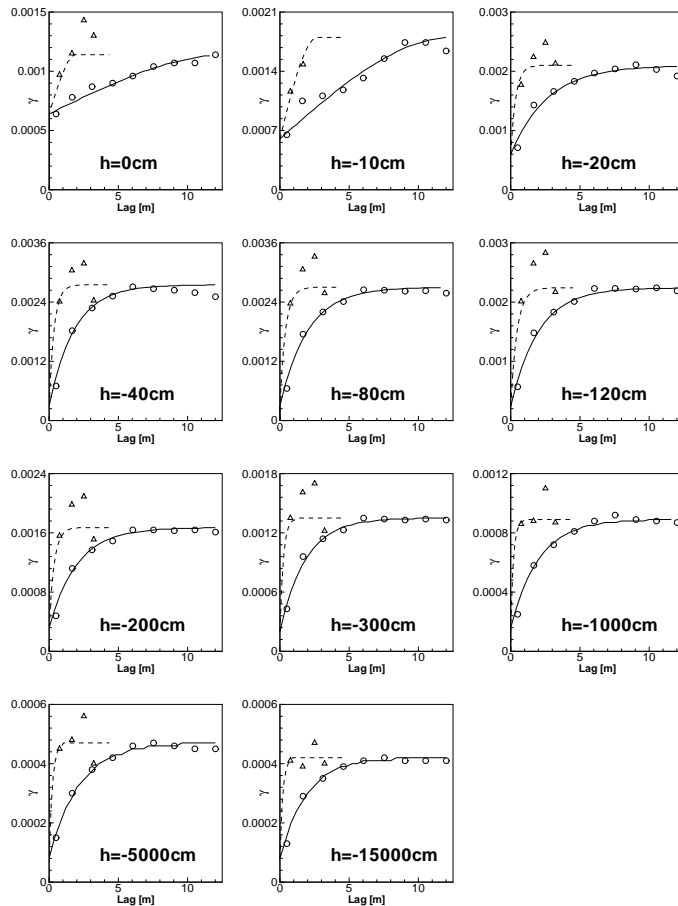
Fig. 6. Experimental semivariograms of water retention data (water contents for a given pressure head) with fitted geometric anisotropy models to horizontal (circle, solid line) and vertical (triangle, dashed line) directions. Semivariograms were calculated from retention data obtained at 448 sampling locations.
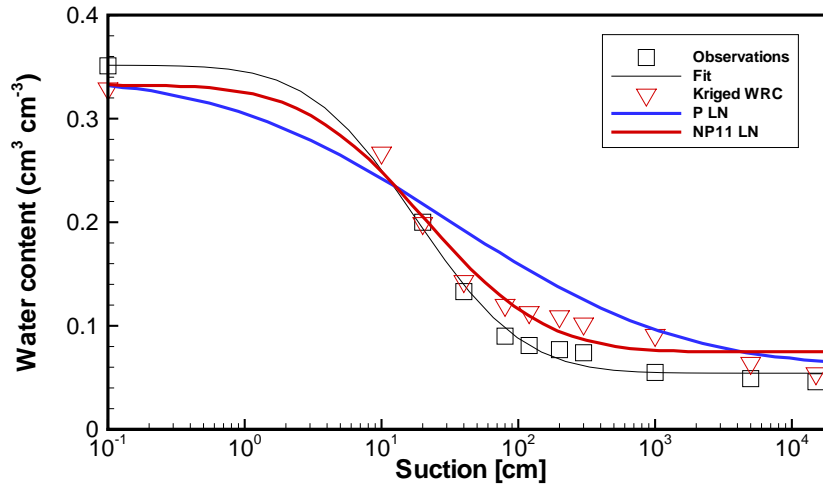
Fig. 7. Water retention data observed (square), fitted using the LN model (solid line), and kriged (triangle) at (x, y)=(7.25, 2.16). In the non-parametric approach, LN model was fitted to the kriged water retention curve (red). In the parametric approach, LN model parameters were kriged to obtain the P-LN curve.
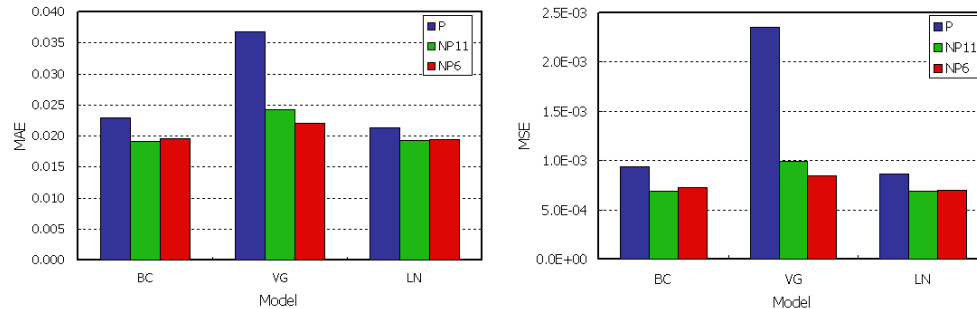
Fig. 8. The impact of the choice of the approach on mean absolute errors (MAE) and mean square errors (MSE) of retention curve predictions for the cross-validation procedure. Three WRC models (BC, VG, and LN) were compared.

Fig. 9. The impact of the choice of the approach on mean absolute errors (MAE) and mean square errors (MSE) of retention curve predictions for the jack-knife procedure. Three WRC models (BC, VG, and LN) were compared.
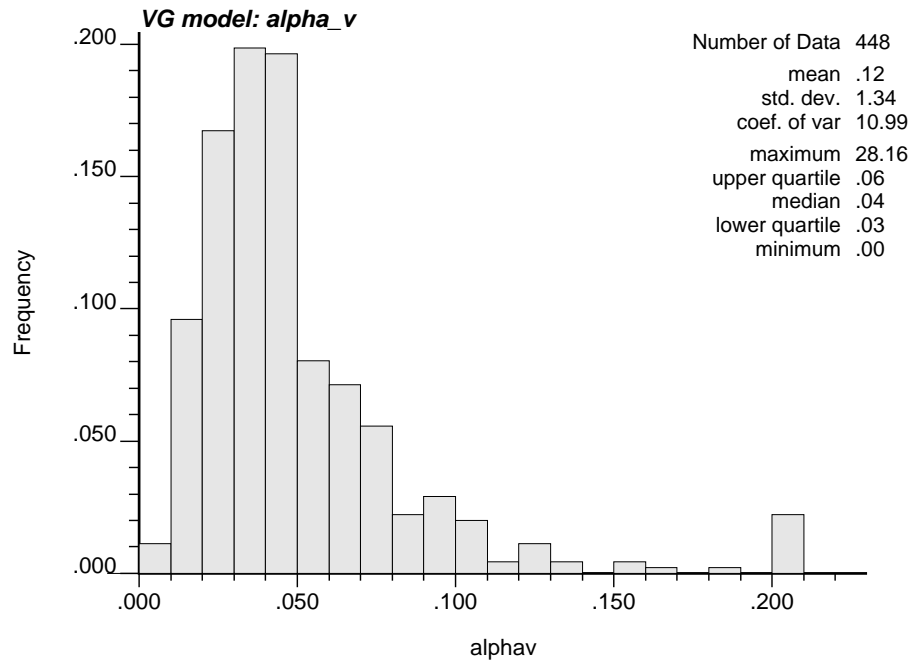
Fig. 10. Histogram of α_v parameter fitted for 448 retention curves.
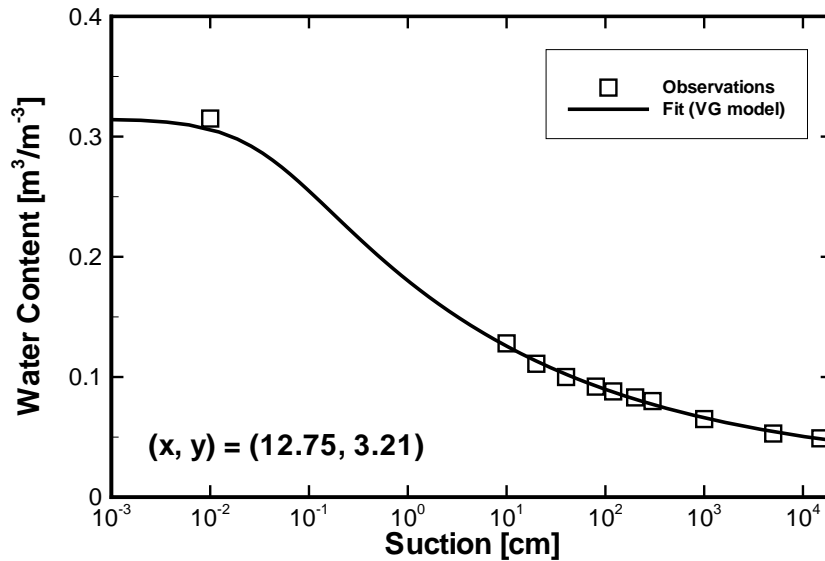
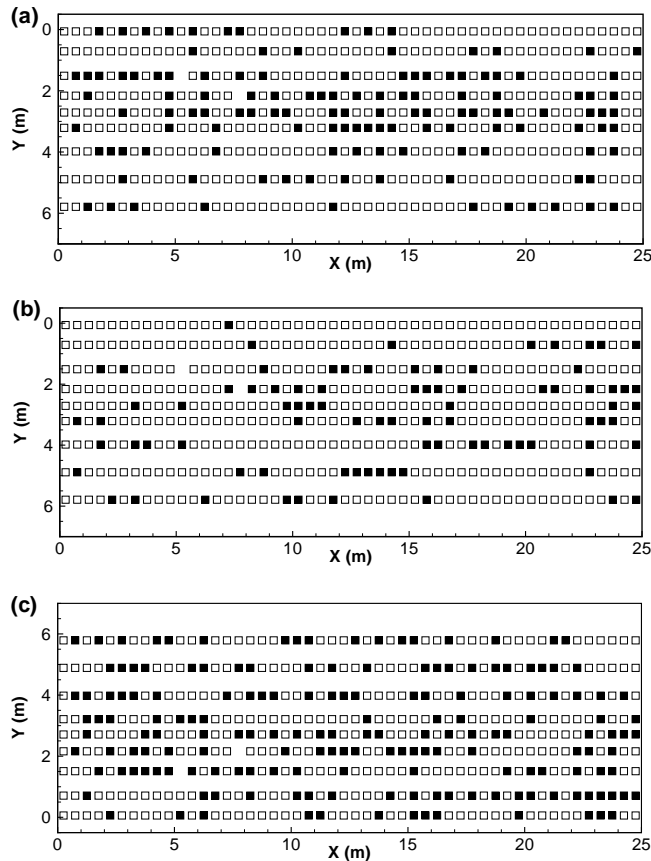**Fig. 11.** Observed water retention data (square) with the VG model fitted at (x, y)=(12.75, 3.21).

Fig. 12. Locations where NP11 outperformed P in terms of mean absolute errors (i.e. smaller MAE) are depicted by open squares, while closed squares indicate locations where P outperformed NP for **(a)** BC, **(b)** VG, and **(c)** LN models.

Geostatistical modeling of water retention curves

H. Saito et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version
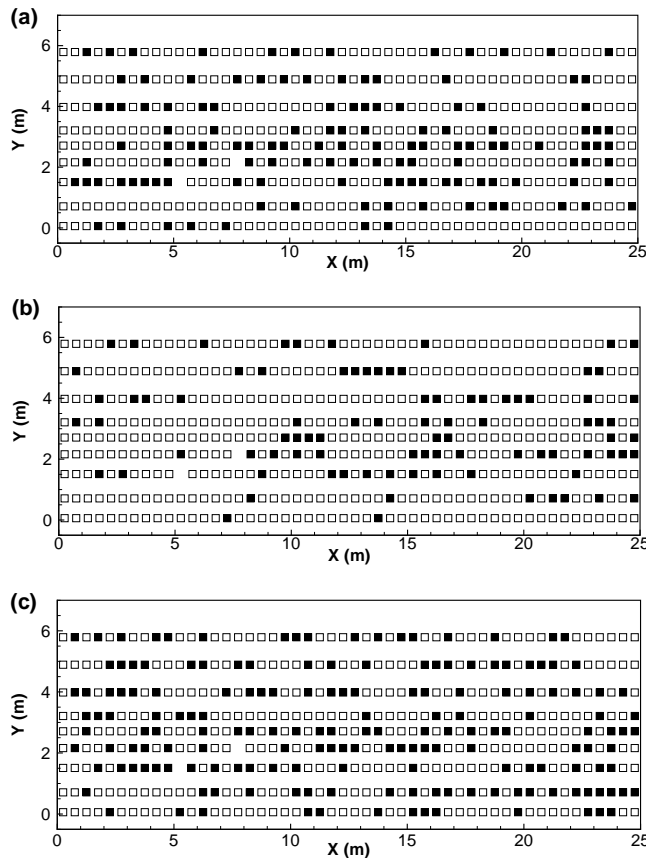
Interactive Discussion

Fig. 13. Locations where NP11 outperformed P in terms of mean square errors (i.e. smaller MSE) are depicted by open squares, while closed squares indicate locations where P outperformed NP for **(a)** BC, **(b)** VG, and **(c)** LN models.