Quantifying soil hydraulic properties and their uncertainties by modified GLUE method

Yifan Yan, Jianli Liu*, Jiabao Zhang, Xiaopeng Li, and Yongchao Zhao

Institute of Soil Science Chinese Academy of Sciences, Nanjing, Jiangsu, China

A b s t r a c t. Nonlinear least squares algorithm is commonly used to fit the evaporation experiment data and to obtain the 'optimal' soil hydraulic model parameters. But the major defects of nonlinear least squares algorithm include non-uniqueness of the solution to inverse problems and its inability to quantify uncertainties associated with the simulation model. In this study, it is clarified by applying retention curve and a modified generalised likelihood uncertainty estimation method to model calibration. Results show that nonlinear least squares gives good fits to soil water retention curve and unsaturated water conductivity based on data observed by Wind method. And meanwhile, the application of generalised likelihood uncertainty estimation clearly demonstrates that a much wider range of parameters can fit the observations well. Using the 'optimal' solution to predict soil water content and conductivity is very risky. Whereas, 95% confidence interval generated by generalised likelihood uncertainty estimation quantifies well the uncertainty of the observed data. With a decrease of water content, the maximum of nash and sutcliffe value generated by generalised likelihood uncertainty estimation performs better and better than the counterpart of nonlinear least squares. 95% confidence interval quantifies well the uncertainties and provides preliminary sensitivities of parameters.

K e y w o r d s: soil hydraulic properties, uncertainty, generalized likelihood uncertainty estimation, evaporation experiment

INTRODUCTION

Nonlinear least squares (NLLS) algorithm is one of the most popular approaches to optimise soil hydraulic parameters from measurements. NLLS optimises the unknown hydraulic parameters by minimising the variance between the predicted and observed measurements of water content, soil suction and hydraulic conductivity. It often provides quite acceptable fits to limited observations, which impresses us that a proper model of a specific system has been acquired. However, there are a number of issues in this algorithm and this back-calculate process is often overlooked, where the underestimated potential uncertainties in the assessments are hidden. One of the issues is equifinality, which means that different hydraulic parameter combinations can result in a similar response. Hence, many sets of model variables may be considered equal or almost equal simulations of the system, which can lead to huge uncertainties in model predictions and explanations, particularly in the case of limited observations and complex models. Another issue is that NLLS attempts to only find the sole 'optimal' parameter set that best fits the observed data, not a set of acceptable parameters. As a result, if the inverse solution has many local minima, using NLLS will lead the simulation to a local optimum, which is exactly not the 'optimal' parameter set at all.

In this study, the results of determining hydraulic parameters through evaporation experiments by NLLS algorithm are compared with another method called generalised likelihood uncertainty estimation (GLUE) developed by Beven and Binley (1992). The GLUE methodology is different from the most popular calibration procedures, in which the global 'optimal' parameter set is sought and assessment of parameter uncertainty is made. The GLUE methodology demands the rejection of the concept of a unique optimal parameter combination in specified model structures, instead, aiming to recognise the acceptability of different parameter sets. A likelihood measure and a cut-off threshold are used to separate behavioural from non-behavioural models. Parameter sets that result in likelihood values below the threshold are 'non-behavioural'.

*Corresponding author e-mail: jlliu@issas.ac.cn

Equifinality can be directly presented by the evaluation of different sets of parameters within the GLUE method. Meanwhile, variants of GLUE have been developed. These include revising GLUE based on the Markov Chain Monte Carlo sampling (Blasone *et al*., 2008), utilising a new program to adjust the prediction limit of a hydrological model in GLUE (Xiong and O'Conner, 2008), and so on.

GLUE methodology has been so widely used because of its operability and its adaptability to nonlinear systems, including those for which a unique calibration is not apparent. Since 1992, it has been widely used for model calibration and uncertainty assessment in hydrologic modelling (McMichael *et al*., 2006; Blasone *et al*., 2008), rainfall-runoff modelling (Lamb *et al*., 1998), and many other fields (Dzotsi *et al*., 2013; Hansson and Lundin, 2006), while its applications in soil water research are relatively scarce.

On the other hand, in recent years some authors criticised GLUE for not being formal Bayesian, and requiring subjective decisions on the likelihood measure and threshold (Christensen, 2004; Mantovan and Todini, 2006; Montanari, 2005). Blasone *et al*. (2008) deemed that GLUE was inapplicable for high-dimensional and complex estimation, because it cannot achieve robust and consistent estimates of behavioural models. Mantovan and Todini (2006) drew a conclusion that the GLUE methodology failed to guarantee the requirements of Bayesian inference process. Shortly afterwards, Beven *et al*. (2007) disproved that GLUE was created to handle real calibration problems in which both inputs and model structural errors played an important role. And if a proper likelihood measure was used, the results would be the same as the formal Bayes. Stedinger *et al.* (2008) also pointed that GLUE could be a useful methodology for uncertainty analyses and model calibration only if a proper likelihood measure and a behavioural threshold could be found. But other studies showed differently Yang *et al*. (2008) and Jin *et al*. (2010) pointed out that initial range of parameters and threshold value influenced the results of the GLUE method the most.

The issues mentioned above may cause a question on the applicability and reliability of GLUE. So, the following issues need to be addressed:

– for a model that is not so complex and not so high-dimensional as hydrology model, *i.e.* van Genuchten-Mualem model (Mualem, 1976), there is a need to know whether the modified GLUE we are using can achieve good performance within lower number of model simulations or not;
– there is a need to systematically evaluate and quantify the effects of threshold values and the sample sizes in GLUE on the parameter uncertainties for van Genuchten-Mualem model;
– there is a need to provide some quantitative criteria in choosing behavioural threshold and determining the sample sizes for this model.

The first purpose of this paper is, therefore, to compare systematically the performance of GLUE with NLLS algorithm in back-calculation of soil hydraulic parameters (soil water retention curve, θ (h) and unsaturated hydraulic conductivity, K (h)) in an evaporation experiment. The second purpose is to assess parameter uncertainties of van Genuchten-Mualem model, that is not as complex and high-dimensional as the hydrology model, by using GLUE method.

MATERIALS AND METHODS

In this section some details of the evaporation method which is used as the direct method in laboratory to simultaneously obtain θ (h) and K (h) are shown. Then the model use for soil hydraulic properties and the NLLS technique for searching the 'optimal' solution was introduced. Finally, some quantitative criteria in choosing behavioural threshold and determining the number of sample simulations for the GLUE methodology was provided.

All of the samples used in this study were collected from agricultural lands in Fengqiu County, Henan Province, in Northeast China (114.51°~ 114.60°E, 34.98°~ 35. 06°N), with an elevation of 65 to 72.5 m. Being situated close to the Yellow River, the soil of Fengqiu County is mainly formed by river alluviation. Fluvo-aquic soil has the widest distribution and accounts for 98.3% of the topsoil in this area. Three kinds of soil textures (loamy sand, sandy loam and sandy clay loam) are adopted in this study, the physical properties are summarised in Table 1. The soil sample of

T a b l e 1. Soil physical characteristics and corresponding range of feasible parameters used in the GLUE

Texture	Clay <0.002 mm	Silt 0.05~ 0.002 mm	Sand 2~0.05 mm	Bulk density	θ_r	θ_s	α (cm)	n	$\log K_0$ (log(cm h^{-1}))
	(%)			g cm^{-3}	cm^3 cm^{-3}				
Loamy sand (a)	6.5	8.3	85.2	1.439	0-0.1	0.41-0.46	0.001-0.13	1.2-3.0	-2-1
Sandy loam (b)	15.6	31.2	63.2	1.384	0-0.14	0.35-0.42	0.001-0.15	1.1-3.0	-2-1
Sandy clay loam (c)	24.2	26.5	49.3	1.518	0-0.15	0.36-0.4	0.001-0.12	1.0-2.8	-1.6-1

Loamy sand is undisturbed and it is collected from top soil; the other two are sieved before being packed in columns and from B horizon (60 cm depth).

Evaporation experiments were conducted on soil columns with a radius of 4.5 cm and length of 15 cm. During the evaporation experiments, the pressure head at specified times at three different depths (2.5, 7.5, and 12.5 cm) within the soil columns was monitored (Fig. 1), and also the soil samples were repeatedly weighed on a balance with 0.1 g accuracy to measure the evaporative water loss. It would take about 25 days to finish the experiment when the upper tensiometer reading value reached 760 cm. Then, the final water content was measured to calculate the total amount of water at each time.

An iterative process, the so-called 'wind method' (Arya, 2002), was involved in the process of analysing the data. Because at each time, only the average water content of the core was obtained, and a potential gradient exists along the core, an iterative process was needed to deduce the $\theta(h)$. Measured water storage acquired from weighing the soil sample at particular times was compared with the total predicted water storage. The water content values were updated based on the differences between the predicted and measured water storage. From these $\theta - h$ pairs, a new $\theta(h)$ was produced, and successive curve-fitting to each updated set of water contents could be achieved. This procedure was repeated until the change of the estimated water content values between iterations was <0.0001 $m^3 m^{-3}$.
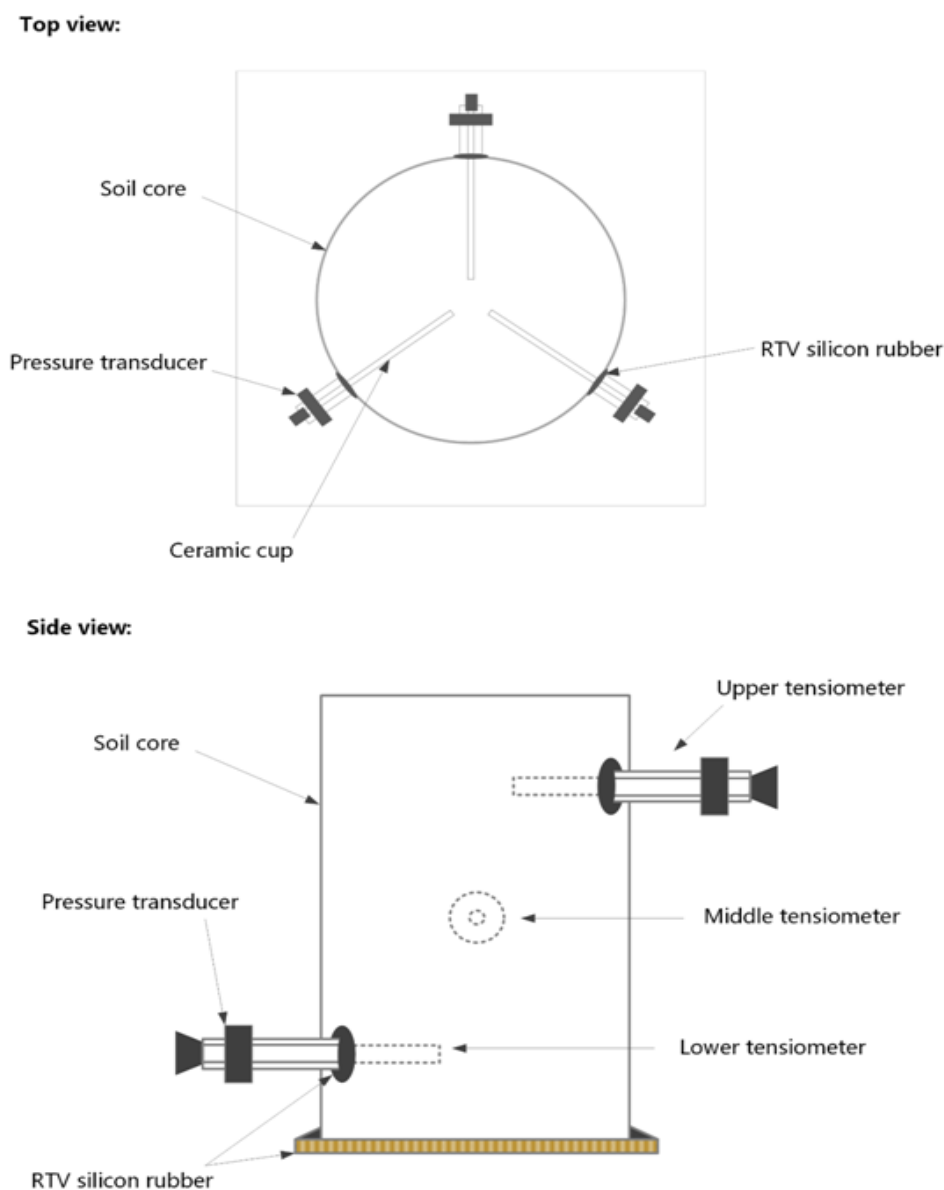

Fig. 1. Experimental setup for the evaporation method.

More information about this experiment and the 'Wind method' can be found in Wendroth (1993) and Šimůnek *et al.* (1998).

The interdependencies of water content, pressure head and conductivity were characterised by the θ (h) and K (h); in this study, van Genuchten-Mualem model was used:

$$S_e(h) = \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} = \begin{cases} \left(1 + |\alpha h|^n\right)^{-m} & h < 0 \\ 1 & h \geq 0 \end{cases}, \quad (1)$$

$$K(S_e) = K^* S_e^{1/2} [1 - (1 - S_e^{1/m})^m]^2, \quad (2)$$

where: $m = 1 - 1/n$, S_e (dimensionless) is the effective saturation, θ_s ($l^3\ l^{-3}$) is the saturated water content, θ_r ($l^3\ l^{-3}$) is the residual water content, n (dimensionless) and α (l^{-1}) are adjustable parameters. K^* ($l\ T^{-1}$) is a matching point hydraulic conductivity for $S_e = 1$. Although K^* is usually deemed as the saturated hydraulic conductivity Ks, a fitted matching point conductivity for slightly unsaturated conditions, $K^* = K_0 < Ks$, may be preferable for modelling unsaturated soils (Weynants *et al.*, 2009).

Parameter estimations were performed by using the RETC procedure (van Genuchten *et al.*, 1991). In order to estimate the unknown model parameters from observed θ (h) and K (h) at the same time, RETC used a weighted least-squares approach based on Marquardt maximum neighbourhood method that belongs to NLLS (Marquardt, 1963). A helpful text with background information on fitting equations to experimental data by using this method was given by Daniel and Wood (1971). The aim of the curve fitting process is to find an equation that maximises the sum of squares associated with the model, minimising the residual sum of squares (SSQ) at the same time. SSQ would be the objective function O (b) to be minimised in RETC. The general formula is:

$$O(b) = \sum_{i=1}^{N} \left\{ w_i [\theta_i - \hat{\theta}_i(b)] \right\}^2 +$$
$$\sum_{i=N+1}^{M} \left\{ w_i W_1 W_2 [Y_i - \hat{Y}_i(b)] \right\}^2, \quad (3)$$

where: b represents the unknown parameter vector, Y_i and $\hat{Y}_i$ are the observed and fitted conductivity data, w_i, W_1 and W_2 are weighting factors, and N is the number of retention data points, M is the total number of observed retention and conductivity data points. The weighting factor W_1 is to add extra flexibility to the parameter optimisation process. W_1 allows one to place more or less weight on the hydraulic conductivity data in their entirety, relative to the soil water retention data. Here we use 0.8, because conductivity data generally are also less precise than water content data (van Genuchten *et al.*, 1991).

More information about RETC and the optimisation approach could be found in its user manual (van Genuchten, 1991). To ensure the global minimum was found, RETC had been run several times with different initial values.

GLUE is based on a different philosophy of model calibration from NLLS. For the sake of understanding, the steps of using the GLUE methodology are described as follows.

First, each specified parameter is assigned with a sampling range. Information from expert knowledge, references and evaporation experiments are taken into account to determine the feasible ranges. The specific range of parameters is demonstrated in Table 1.

Parameter sets used in the GLUE method could be sampled from any probability distribution, in which uniform distributions are the most commonly reported. In this paper, uniform distributions and Monte Carlo sampling method were applied through the specific ranges to avoid strong prior assumption about the covariance of the parameters.

Every parameter set produces one model output, and that is compared with experimental data by using likelihood measure. Various likelihood measures have been put forward and evaluated in the literature (Christensen, 2004; Stedinger, 2008). In this paper the likelihood measure is the one most often applied, introduced by Nash and Sutcliffe (1970):

$$L(\theta_i | Y) = 1 - \frac{\sum_j^n (\theta_{ij} - \theta_{oj})^2}{\sum_j^n (\theta_{oj} - \theta_o)^2}, \quad (4)$$

$L(\theta_i | Y)$ is the likelihood measure of parameter i; θ_{ij} is the simulated response value at time j with parameter i; θ_{oj} is the observed response value at time j; θ_o is the mean value of observed data; n is the total number of observations. In this study, the weighting factor W_1 in Eq. (3) used by NLLS is 0.8, in order to maintain consistency; the final likelihood value L is defined as:

$$L = \frac{L(\theta_i | Y) + 0.8 L(K_i | Y)}{1.8}. \quad (5)$$

The smaller the difference between prediction and observation, the higher the L is ($L \leq 1$). Parameter sets that resulted in likelihood values lower than a certain threshold are termed as 'non-behavioural'. The 'behavioural' parameter sets are assigned and rescaled likelihood weights that sum to 1 and formulate a cumulative distribution.

Three indices are adopted as metrics for the 95% confidence interval (95CI) of soil water content and conductivity. The first is the percentage of observations covered by the 95CI (P_{95CI}), the second is the maximum of Nash and Sutcliffe value (*MNS*) and the third is the average relative interval length (*ARIL*) (Jin *et al.*, 2010). These metrics are expressed below:

$$P_{95CI} = \frac{NQ_{in}}{N_{obs}} 100\%, \quad (6)$$

NQ_{in} is the number of observations located within 95% CI; N_{obs} is the number of total observations.

$$MNS = \max_{i=1}^{N} \{NS_i\} = \max_{i=1}^{N} L(\theta_i | Y), \quad (7)$$

i is the acceptable sample index, N is the number of 'behavioural' runs. MNS is the maximum of Nash and Sutcliffe value:

$$ARIL = \frac{1}{n} \sum \frac{Limit_{upper,t} - Limit_{lower,t}}{R_{obs,t}}. \tag{8}$$

$Limit_{upper,t}$ and $Limit_{lower,t}$ are the upper and lower boundary values of 95% CI; n is the number of total observation; and $R_{obs,t}$ is the value of the observed data. The perfect simulation is the one that got $ARIL = 0$, the $P_{95CI} = 100\%$ and $MNS = 1$.

RESULTS AND DISCUSSION

Through laboratory experiments, 381, 177 and 252 data pairs of θ (h), and 249, 112 and 162 data pairs of K (h) for loamy sand, sandy loam and sandy clay loam were obtained, respectively. The 'optimal' hydraulic parameters for all kinds of soils were obtained by simultaneous fitting θ (h) and K (h) data by using NLLS algorithm. The results of parameter estimations and goodness of fit of all soil samples are illustrated in Table 2. The goodness of fit of the evaporation experiments by using NLLS is good, and R^2 is 0.7926, 0.9791 and 0.9631, respectively. The reason why R^2 for loamy sand is relatively low will be illustrated in the fourth part of this section.

The acceptable sampling rate was fixed to be 0.5% and the sample size was changed to: 5 000, 10 000, 15 000, 20 000, 25 000, 30 000, 35 000 and 40 000. Relations between MNS or threshold value and the number of sample simulations are shown in Fig. 2. It shows that the MNS values of b (sandy loam) and c (sandy clay loam) both increased with the increase of the number of sample simulations from 5 000 to 250 000. When the number of sample simulations was larger than 25 000, the MNS values of b (sandy loam) and c (sandy clay loam) decreased slightly, then almost kept stable. The corresponding threshold values had a similar tendency. For soil a (loamy sand), the MNS and the threshold value both reach the highest likelihood values when the sample size is 15 000. It shows that the number of sample simulations directly impact the per-

formance of GLUE when the sample size is not big enough. So, in order to reach a balance between the sample size and computational efficiency, in the following study of sensitivity of GLUE simulations we adopt the sample size equal to 15 000 for a (loamy sand) and 25 000 for b (sandy loam) and c (sandy clay loam). The number of sample simulations for van Genuchten-Mualem model in our research is quite a low as compared to McMichael et al. (2006), Li et al. (2010) and Jia and Culver (2012). They all use Monte Carlo assessment or GLUE for much more complicated models, such as conceptual catchment models and watershed models. Kuczera and Parent (1998) even explained that using a simple uniform prior probability distribution of model parameters over a relatively large region could lead to a consequence that even one good solution could not be sought after billions of model evaluations. Nevertheless, we obtained acceptable solutions (those simulations that have higher likelihood values) with relatively small sample size for all three kinds of soil textures. It meant that, judging from the respect of sample sizes, using GLUE methodology to estimate K (h) and θ (h) and analyse the parameter uncertainty from the evaporation experiment is a proper choice.

Previous researches have indicated the importance of the choice of threshold values for the likelihood measure. Two main methods have been recommended for defining the threshold values. One is to quantify the threshold value through adding a certain deviation to MNS, the other is to fix the percentage of the total number of simulations. The latter one is applied in this study and named as the acceptable sample rate (ASR) of 2, 1.5,1.2, 1, 0.8, 0.6, 0.5, 0.4, 0.25, and 0.1%. It is clearly seen in Fig. 3 that with an increase of ASR, the cut-off value decreases when the number of sample simulations is fixed.

The relationship among ASR, ARIL and P_{95CI} is shown in Fig. 4. Overall, the ARIL and the P_{95CI} both are increased with the growth of ASR. It means that the higher the ASR is, the better the P_{95CI}, but the worse the ARIL is. The explanation is that with the increase of ASR, the threshold

T a b l e 2. Results of parameter estimation and goodness of fit of evaporation experiment using NLLS

Texture	θ_r^a (cm³ cm⁻³)	θ_s^a (cm³ cm⁻³)	α^a (cm)	n^a	$\log K_0^a$ (log (cm h⁻¹))	R^2
Loamy sand	0	0.454 (0.383, 0.525)	8e-3 (6.6e-3, 8.4e-3)	1.598 (1.467, 1.729)	-0.953 (-1.039, -0.881)	0.793
Sandy loam	0.04 (0.002, 0.08) b	0.386 (0.375, 0.396)	3.2e-3 (3e-3, 3.4e-3)	1.553 (1.507, 1.599)	-1.484 (-1.51, -0.461)	0.979
Sandy clay loam	0	0.383 (0.374, 0.392)	4.3e-3 (4.2e-3, 4.5e-3)	1.596 (1.561, 1.632)	-0.748 -0.773, -0.725)	0.963

*When value <0.0001, it shows 0 in this table, a – optimised parameter, b – 95% CI determined by NLLS.
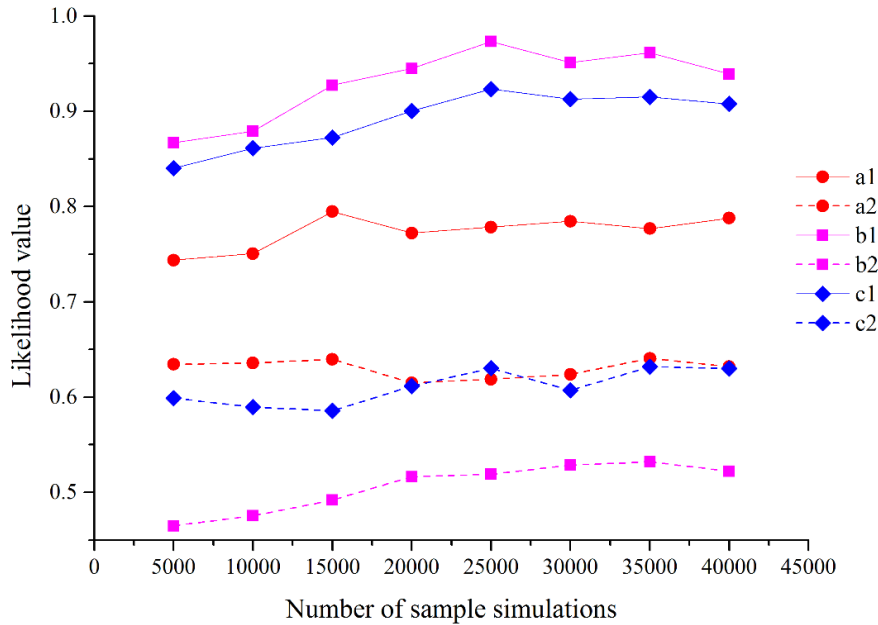
Fig. 2. Relationship between *MNS*, threshold value and number of sample simulations when the acceptable sampling rate is *ASR*=0.5%. Loamy sand (a), sandy loam (b), sandy clay loam (c); 1 and 2 means *MNS* and threshold value, respectively.
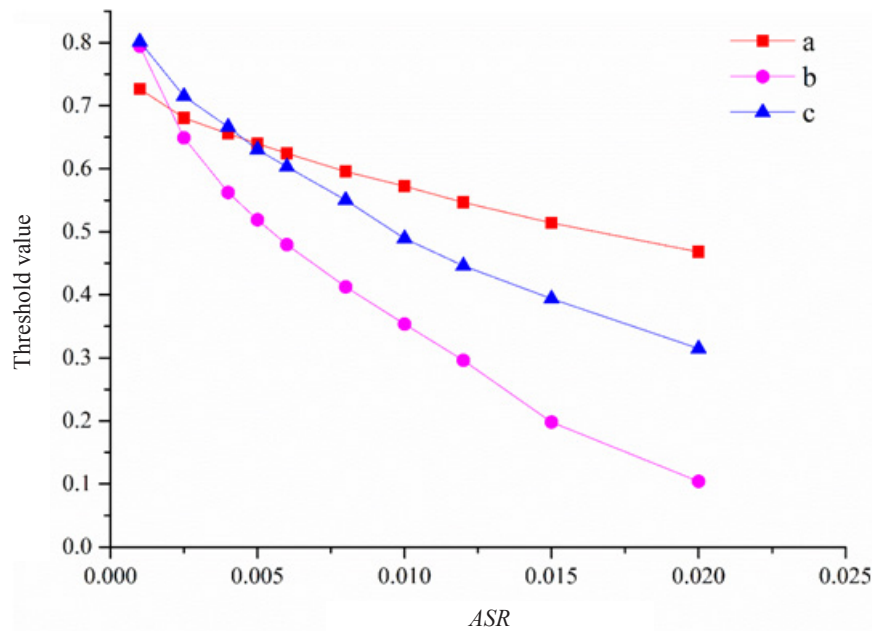

Fig. 3. Relationship between *ASR* and threshold value for three kinds of soil when the number of sample simulations is: 15 000 for soil loamy sand (a); and 25 000 for both soil: sandy loam (b), and sandy clay loam (c).

value decreases. So, the number of acceptable simulations increases, and more simulations that have lower likelihood values are involved in calculating 95% CI. The *ARIL* value inevitably increases with the broadening of 95% CI, in general. Meanwhile, more observed points fall into the 95% confidence interval.

The *ARIL* of *K* (*h*) and *θ* (*h*) increase fast when *ASR* increases from 0.1 to 0.4% (Fig. 4a). When *ASR* increases from 0.4 to 0.6%, the *ARIL* becomes insensitive to the changes in *ASR*. It means the goodness of the simulations

does not change a lot when *ASR* increases from 0.4 to 0.6%. When *ASR* is larger than 0.6%, the *ARIL* of *K* (*h*) shall increases rapidly. This demonstrates that when *ASR* is larger than 0.6%, a lot of simulations that perform poor in fitting *K* (*h*) will be involved in behavioural models. Based on this, *ASR* that is larger than 0.6% is completely out of the scope of consideration. The P_{95CI} of *K* (*h*) and *θ* (h) also increase fast when *ASR* increases from 0.1% to 0.4%, but then P_{95CI} of *θ* (h) still increases even though quite a little. So, 0.6% is chosen as the *ASR* of loamy sand. For sandy
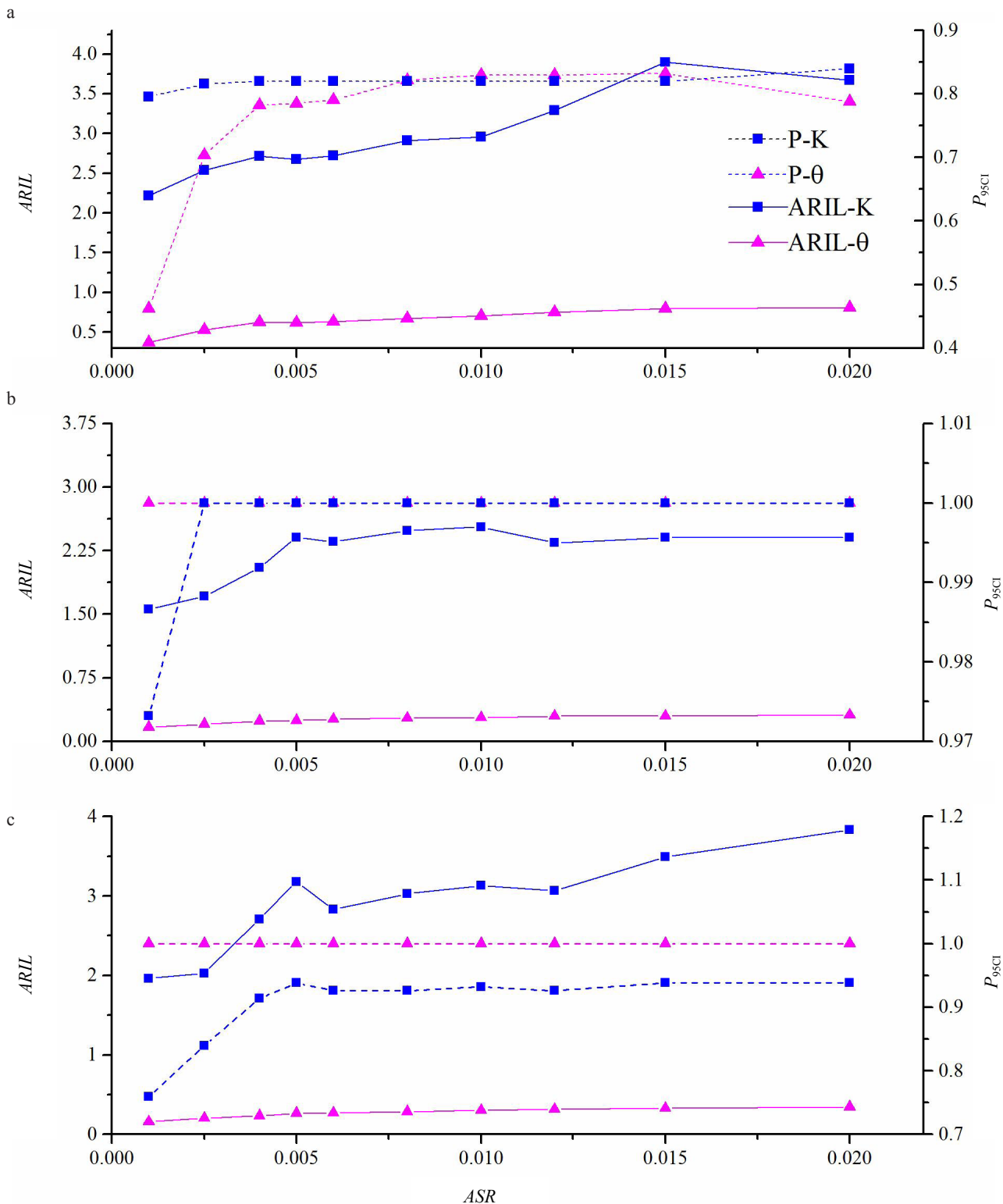
Fig. 4. Relationship among *ASR*, average relative interval length (*ARIL*) and P_{95CI} for three kinds of soils when the number of sample simulations is 15000 for soil: a – loamy sand; and 25000 for: b – sandy loam and c – sandy clay loam.
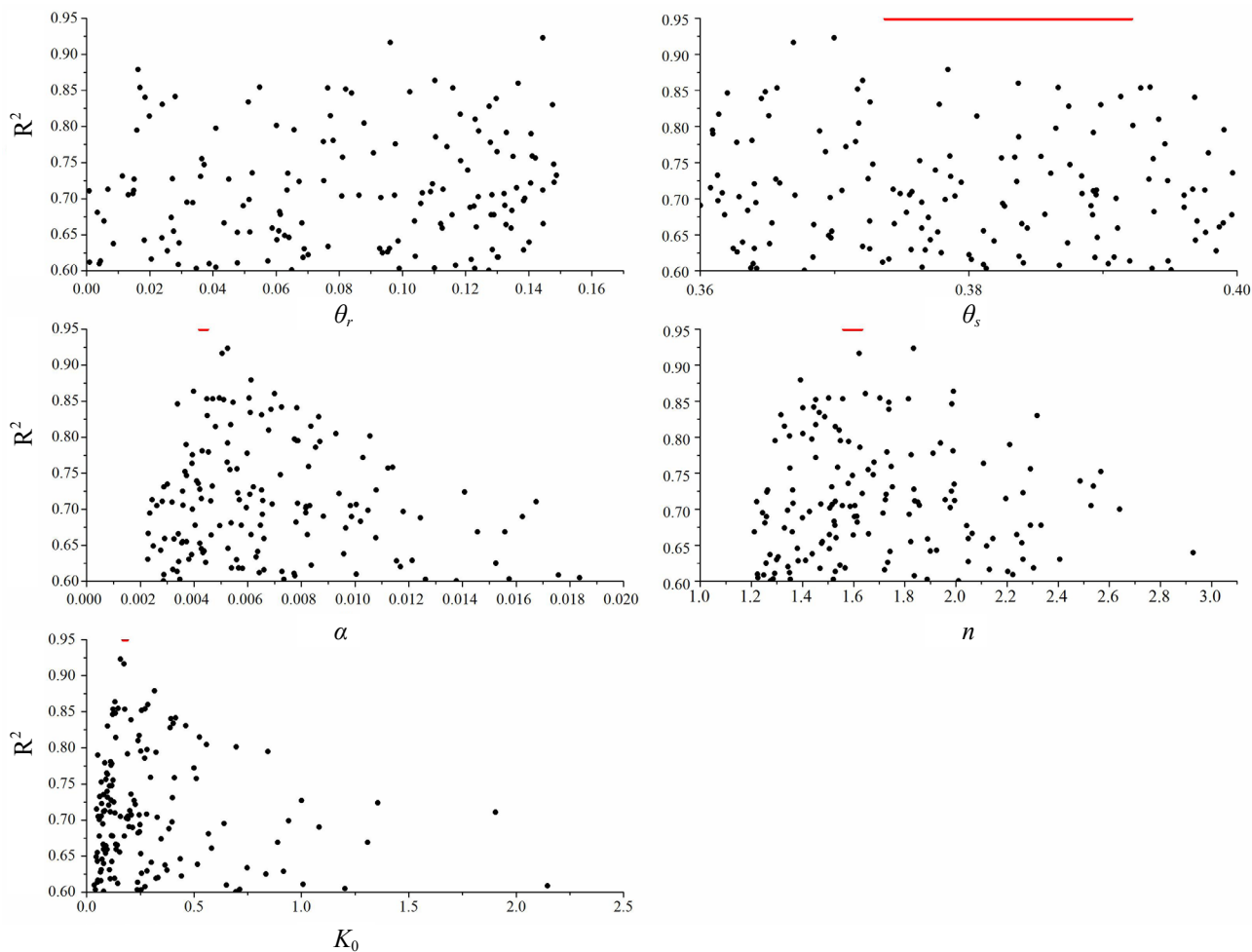
Fig. 5. Scatter plots of sandy clay loam (c) when N=25 000, R^2 represent likelihood value, the x axis is the parameter values. Each dot represents a single run of the simulation and its corresponding modelling effiency. The line on the top of the plots of θ_s, α, n, K_0 represents 95% confidence intervals determined using NLLS.

loam (Fig. 4b), the percentage of observations located within the 95% confidence interval for $K(h)$ and $\theta(h)$ is very high, larger than 0.97, so *ARIL* is the only consideration when choosing proper threshold value. Figure 4b clearly shows that the *ARIL* of $K(h)$ and $\theta(h)$ both gain the lowest value at *ASR* equal to 0.1%. So, 0.1% is chosen as the *ASR* of sandy loam. Analogously, 0.4% is a reasonable choice for sandy clay loam. Moreover, *ARIL* of $K(h)$ is much higher than that of $\theta(h)$, and this demonstrates that the variation and uncertainty of unsaturated conductivity is much larger than that of soil water content. Mohrath *et al.* (1997) investigated the errors due to the position of the tensiometers and the calibration of the transducer used in an evaporation experiment. The conclusion is that the water retention curve, compared to hydraulic conductivity, is less impacted by these errors. Little deviations in the position of the tensiometers (1-2 mm) and in transducers calibration (1-5%) can cause large uncertainties in the estimate of $K(h)$. The

results of Tamari *et al.* (1993) also reflect that estimation of the $\theta(h)$s by using the Wind method is not quite sensitive to experiment errors, but small uncertainties in tensiometric data extremely influence the $K(h)$s determined in wet conditions. Our findings are quite consistent with their results, showing a much higher uncertainty in the estimation of $K(h)$ than that of water retention.

The results of the Monte Carlo simulation for sandy clay loam (c) are shown in Fig. 5 as an example illustrating parameter uncertainty and sensitivity. It was called the scatter plots, showing the likelihood values for the range of defined parameters, θ_r, θ_s, α, n, K_0. The plots represent the goodness of fit for the parameters, and the closer the value of R^2 to 1, the better this parameter set performs. The threshold value in this figure is set at likelihood value equal to 0.6. The dots under 0.6 are regarded as 'non-behavioural' ones. From this figure we ascertain the acceptable range of defined parameter values. For instance, the initial range
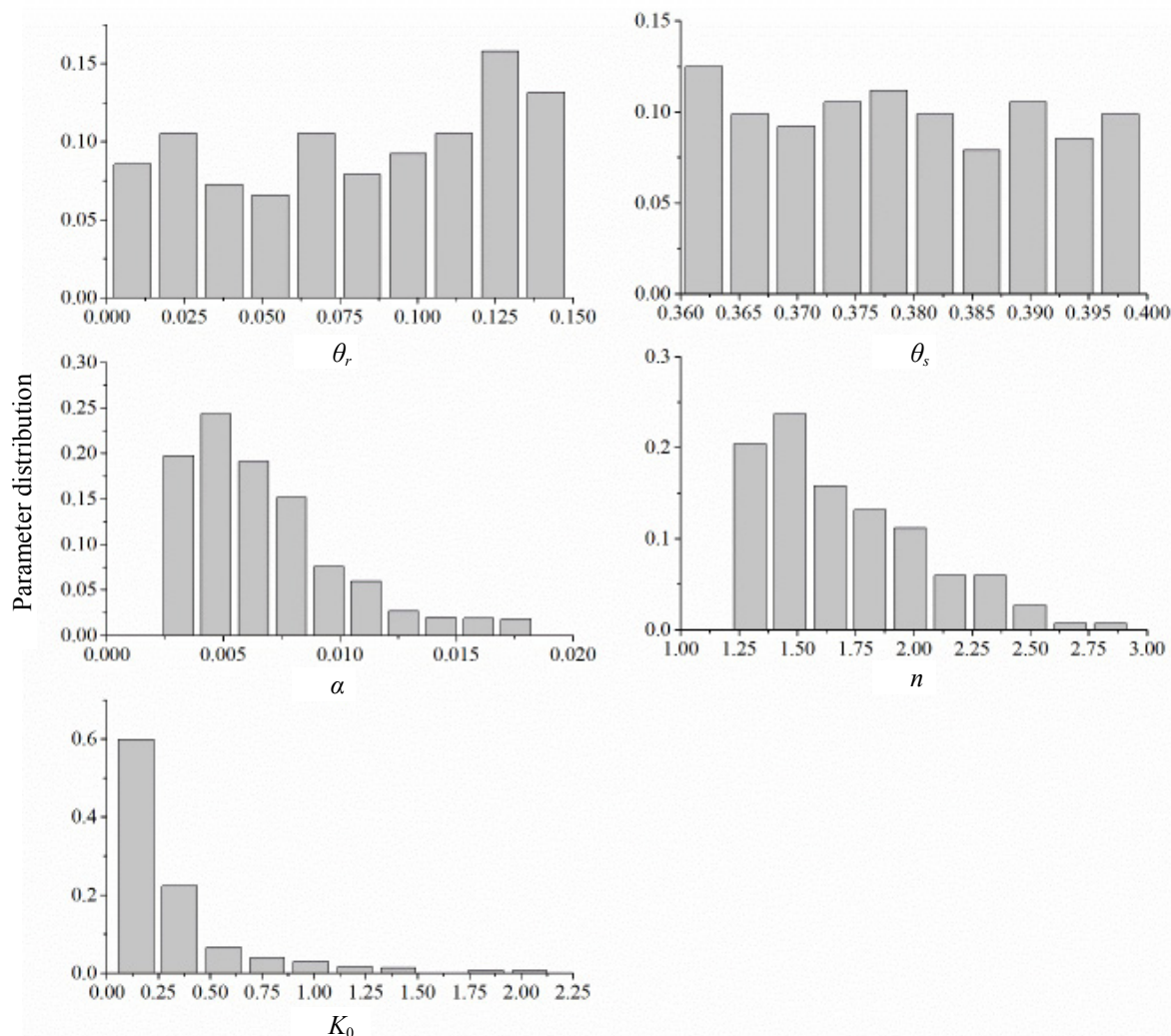
Fig. 6. Posterior distribution of parameters of sandy clay loam (c) when N = 25 000 and threshold value is 0.6. The x axis is parameter range, the y axis is parameter distribution.

of α used in GLUE is between 0.001 and 0.12 (Table 1), and in Fig. 5 the values of α are between 0.002 and 0.02 when threshold value is 0.6, and the range of parameters will keep narrow with the increase of threshold value. For example, if cut-off value is 0.8, α will be between 0.0034 and 0.0104. With the change of value range, the posterior probability distribution will change accordingly. These scatter plots clearly show that there is no sole 'optimal' solution; many parameter sets can fit the observations quite well (which have higher likelihood value), but the majority of these parameter sets fall outside of the 95% confidence limits (CL) determined by NLLS. Therefore, the results of the GLUE method manifest that the parameters estimated by NLLS are not as robust as suggested. So, using the sole 'optimal' solution obtained by NLLS to predict θ (h) and K (h) is very risky. For direct inspection, histograms of the marginal posterior parameter distributions derived by the GLUE methodology are plotted in Fig. 6. Ideally, the pre-

dicted uncertainty should be spread as small as possible, but consistent with observations, so that the parameter probability distribution function (PDF) is as sharp as possible (Gneiting *et al.*, 2003). The flatter the posterior parameter probability distribution of the parameter is, the lower the sensitivity of van Genuchten model to this parameter. The lower the sensitivity of the model to this parameter, the higher the uncertainty of this parameter to the model. The posterior PDFs of parameters θ_r, θ_s are similar to uniform distributions, but those for parameters α, n and K_0 are not – each of them shows obviously a well-defined peak and relatively centralised, especially for K_0. It means that when θ (h) and K (h) both are used in the parameter estimation procedure, relatively accurate estimations of α, n and K_0 are obtained, indicating these parameters are better identified and less uncertain. Whereas, the PDFs of parameters θ_r and θ_s are flat, indicating more uncertainty in the parameters. The 'dotty plot' and marginal posterior parameter distribu-
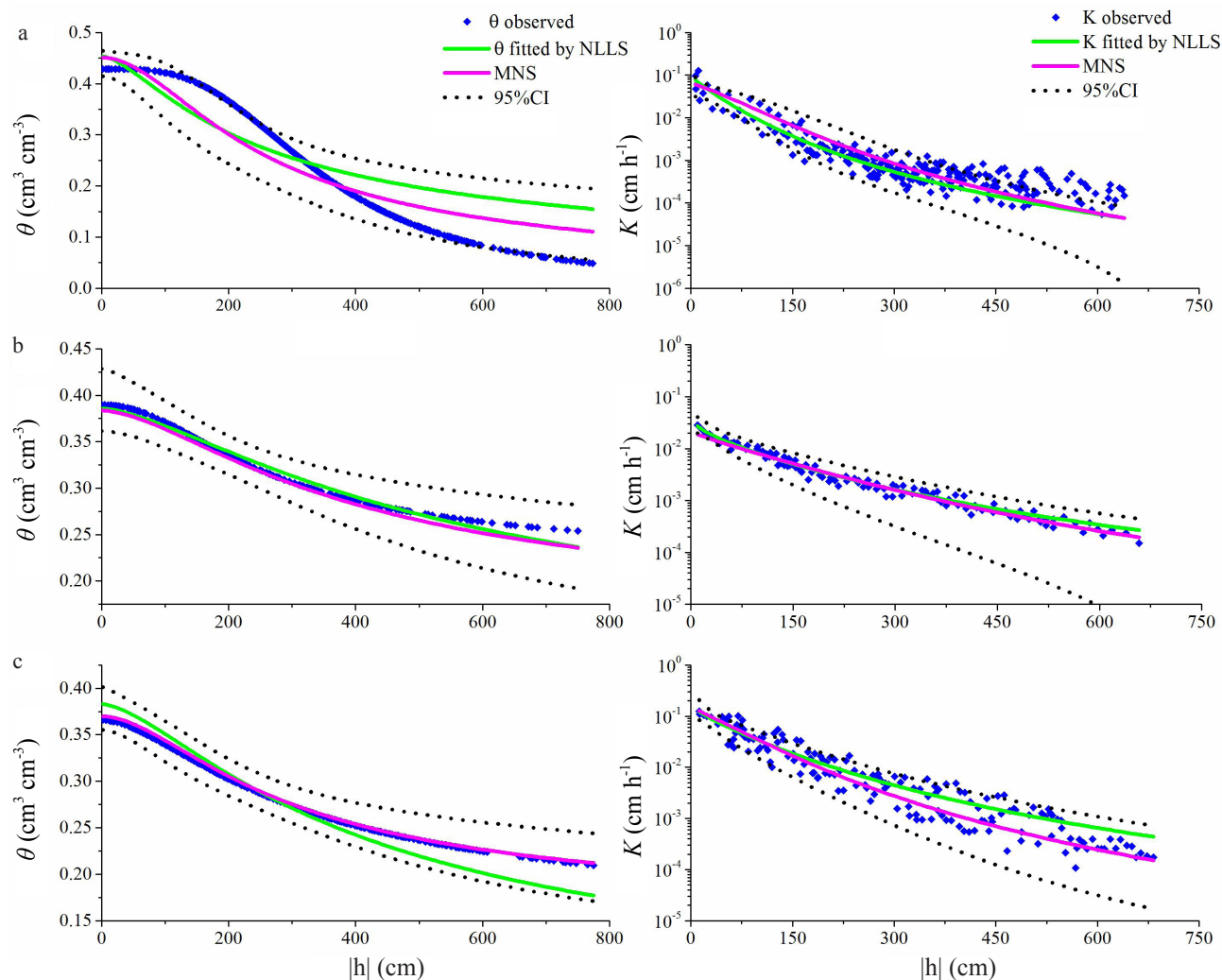
Fig. 7. Predicted water retention curve (left column) and hydraulic conductivity (right column) for: a – loamy sand, b – sandy loam, c – sandy clay loam. Blue dots represent values calculated by using wind method. The green curves are the best fitted ones with NLLS method (parameters in Table 2). The magenta curves are the fitted values by using *MNS* value of GLUE method (parameters in Table 3) and the black short dash lines represent the 95% confidence intervals.

tions of sandy loam and sandy clay loam show the same tendency. It demonstrates that the goodness of fitting of observed data is sensitive to α, n and K_0, but insensitive to the parameters θ_r and θ_s. And the parameters θ_r and θ_s have higher uncertainties than those of α, n and K_0, which is consistent with Shi *et al.* (2015) and Younes *et al.* (2013).

Measurements by using Wind method and predicted $\theta(h)$ and K (h) by deterministic inverse modelling by using NLLS solution and GLUE for all kinds of soils are shown in Fig. 7. The curve of *MNS* generated by GLUE is the simulation that has the maximum Nash-Sutcliffe value. The model prediction intervals shown in the figure are the 2.5th and 97.5th percentiles of the model output distributions computed by the GLUE method. It is calculated from the mean value of the prediction±1.96 times its standard deviation.

For loamy sand, the performance of the 'optimal' parameter combination of NLLS and *MNS* of GLUE are quite comparable. But both the predictions of water retention and conductivity are not so good as for sandy loam and sandy clay loam. It is because Wind method cannot calculate conductivity near saturation correctly because of the small range of pore-size distribution in sand. It can only develop below potentials of -50 to -100 cm in sand. Budiman and Damien (2005) conducted an evaporation experiment in seven soil textures (from heavy clay to sand) and they found the same condition for sand too. They explained that the gradient between tensiometers was very small near saturation, because at this stage water was held at its most common pore size. When the soil tension was lower than -55 cm, the rate of water loss increased quickly and the surface potential decreased fast. This leads to obstacles

T a b l e 3. Summary of posterior distribution for each parameter by GLUE method when N =15 000, *ASR* = 0.6% for loamy sand; N = 25 000, *ASR* = 0.1% for sandy loam; N = 25 000, *ASR* = 0.4% for sandy clay loam

Statistics		*MNS*	Minimum	Maximum	Mean	Variance	$P_{0.975}$	$P_{0.025}$
Loamy sand	θ_r (cm^3 cm^{-3})	0.009	0.001	0.09	0.041	0.025	0.088	0.003
	θ_s (cm^3 cm^{-3})	0.451	0.410	0.460	0.438	0.015	0.460	0.412
	α (cm)	0.006	0.003	0.017	0.007	0.003	0.016	0.004
	n	1.965	1.382	2.974	2.036	0.429	2.920	1.406
	$\log K_0$ (log (cm h^{-1}))	-1.182	-1.711	-0.545	-1.178	0.264	-0.607	-1.628
Sandy loam	θ_r (cm^3 cm^{-3})	0.124	0.007	0.138	0.079	0.040	0.137	0.009
	θ_s (cm^3 cm^{-3})	0.384	0.350	0.419	0.390	0.018	0.418	0.352
	α (cm)	0.004	0.001	0.025	0.007	0.005	0.023	0.002
	n	1.705	1.109	2.735	1.535	0.331	2.428	1.141
	$\log K_0$ (log (cm h^{-1}))	-1.492	-1.945	0.206	-1.216	0.498	-0.135	-1.925
Sandy clay loam	θ_r (cm^3 cm^{-3})	0.145	0.001	0.149	0.085	0.045	0.148	0.006
	θ_s (cm^3 cm^{-3})	0.367	0.360	0.400	0.379	0.012	0.429	0.361
	α (cm)	0.005	0.002	0.017	0.007	0.003	0.016	0.002
	n	1.835	1.212	2.640	1.699	0.340	2.539	1.245
	$\log K_0$ (log (cm h^{-1}))	-0.803	-1.332	0.280	-0.693	0.355	0.117	-1.271

*Percentiles $P_{0.975}$, $P_{0.025}$ are for the tail areas.

in the numerical simulation, where the collapse of water potential cannot be handled by the numerical solution. It has been proved by Wendroth *et al*. (1993) as being exactly the limitation of the evaporation method. Furthermore, the assumption of the Wind method that the potential decreases linearly with depth might not be set up. So, the measured water content and potential, and hence calculated conductivities, are not accurate when using the Wind method in loamy sand. Therefore, both NLLS and GLUE could not achieve satisfactory parameter estimations for van Genuchten model in loamy sand. Although Wind method is not the most exact way to obtain soil hydraulic properties for coarse texture soil, we still use this method because of its relatively low cost and easy operation, especially when the number of soil samples is large or for a field scale.

For sandy loam, the measured θ (*h*) at the dry end is higher than the sole 'optimal' solution of NLLS and the *MNS* of GLUE, but all of the water retention values still fall into the 95% CI. The one predicted by the NLLS and the *MNS* both perfectly fit the measured soil conductivity by Wind method, and 97.3% of the measured conductivity values locate in the 95% confidence interval. For sandy clay loam, *MNS* generated by GLUE approach performs

quite well for both θ (*h*) and *K* (*h*), but the one predicted by the inverse solution NLLS at the dry end is much lower than the measured water retention. It can be expected that, with the decrease of water content, using the sole 'optimal' parameter sets to forecast water retention curves will lead to larger uncertainty. It is very speculative to extrapolate beyond the range of measurement. This is consistent with the results of Budiman and Damien (2005) research. Those authors also found that with the decrease of water potential, the uncertainty in the wet end was low, which resulted in a good confidence in predicting θ_s. However, at the dry end, the uncertainty starts to widen. Šimůnek *et al*. (1998) also proved that extrapolation beyond the experimental value was associated with a high level of uncertainty.

Comparing the 95% confidence intervals of the predictions with the measured data, most of the measured data are located in the 95% confidence intervals. Except for loamy sand, only 5% at most of Wind values fell out of the 95% confidence interval, which means that parameter variability alone for sandy loam and sandy clay loam can compensate for other sources of errors, such as measurement and model structure errors, and thus, it can account for the total output uncertainty. For loamy sand, the evaporation experiment

cannot calculate conductivity near saturation correctly, so we cannot come to a conclusion about the main source of uncertainty. The exact results for all soils are presented in Table 3. The 'optimal' parameter set of all three kinds of soils (Table 2) is included in the 95% CI generated by GLUE, except for θ_r of loamy sand and sandy clay loam (Table 3). The 'optimal' parameter set and the ranges of parameters θ_s, α, n (Tables 2 and 3) are quite acceptable compared to the lower and upper limits from the literature (Mertens *et al.*, 2004). The 'optimal' θ_r of loamy sand and sandy clay loam in Table 2 is 0, which means these values are lower than 0.0001. The same results were obtained for silty loam and clay in Wendroth *et al.* (1993) study, while the value of K_0 in this study is much larger and varies drastically for three kinds of soil textures (sandy loam, silty loam and clay). The 'optimal' K_0 determined by NLLS for loamy sand in this study falls into the range of Mertens *et al.* (2004), but it is much lower than the average. It may be related to (a) the combination of the van Genuchten-Mualem model that does not allow for large changes of soil hydraulic conductivity near saturation, and (b) the use of the Wind's method that does not allow to monitor the conditions near saturation closely.

Overall, the performance of NLLS and GLUE is comparable in terms of predicting $\theta(h)$ and $K(h)$ simultaneously, but at the dry end, GLUE performs better. What is more important is that GLUE is able to estimate the uncertainty of the predicted hydraulic properties and can be used to conduct sensitivity analysis at the same time, while the NLLS cannot. The GLUE approach demonstrates intuitively that there is no single reasonable set of parameters to predict soil water flow. Many different combinations of parameters can achieve the same good performance. *MNS* is an exact example that directly shows the same results as the scatter plots do.

CONCLUSIONS

1. The 'optimal' solution of nonlinear least squares by optimisation within retention curve fits the observed data quite well. However, this alone is not sufficient to imply that the 'optimal' solution is the real best fit to the data and the assumed model is definitely a correct representation of the evaporation process. Good fitting results (small residual variance) of the nonlinear least squares optimisation theory do not mean small parameter uncertainties. The generalised likelihood uncertainty estimation methodology, on the other hand, clearly demonstrates that a wider range of parameter sets can give acceptable simulations of water retention and hydraulic conductivity curve.

2. In the generalised likelihood uncertainty estimation method, simulation of $\theta(h)$ and $K(h)$ is very sensitive to the choice of sample sizes and the threshold value. Insufficiency of simulation samples influences the maximum of Nash and Sutcliffe value when the acceptable sample rate is fixed. Compared to high-dimensional and complex estimation problems in hydrology modelling, the number of sample simulations for $\theta(h)$ and $K(h)$ is quite acceptable.

3. The uncertainty of water retention curve is much lower than that of hydraulic conductivity curve. Extrapolating the water retention and hydraulic conductivity beyond measurement has a high uncertainty. Using the sole 'optimal' parameter set of nonlinear least squares is quite risky to predict $\theta(h)$ and $K(h)$, especially at the dry end.

4. Although the modified generalised likelihood uncertainty estimation we used in this study does not overcome all the lively debated points of generalised likelihood uncertainty estimation, it seems adequate to fit $\theta(h)$ and $K(h)$ simultaneously and to quantify the parameter uncertainty for the van Genuchten-Mualem model.

5. The uncertainty of water capacity and conductivity propagating from measurement to prediction is of significant importance in soil-water management, while generalised likelihood uncertainty estimation provides a way to quantify the uncertainty in the process of prediction, so it helps us to lower the risks of further prediction when high uncertainty exists.

Conflict of interest: The Authors declare that do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

REFERENCES

Arya L.M., 2002. Wind and hot-air methods. In: Methods of Soil Analysis. Part 4. Physical Methods (Eds J.H. Dane, G.C. Topp). SSSA Book, SSSA, Madison, WI, USA.

Beven K., 1993. Prophecy, reality and uncertainty in distributed hydrological modeling. advances in water resources. Advances in water resources, 16: 41-51. doi:10.1016/0309-1708(93)90028-e.

Beven K. and Binley A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrological Processes, 6: 279-298. doi:10.1002/hyp.3360060305.

Beven K., Smith P., and Freer J., 2007. Comment on 'Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology' by P. Mantovan and E. Todini. J. Hydrology, 338: 315-318.

Blasone R.-S., Vrugt J.A., Madsen H., Rosbjerg D., Robinson B.A., and Zyvoloski G.A., 2008. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov chain Monte Carlo sampling. Advances Water Resources, 31: 630-648. doi:10.1016/j.advwatres.2007.12.003.

Brazier R.E., Beven K.J., Anthony S.G. and Rowan J.S., 2001. Implications of model uncertainty for the mapping of hillslope-scale soil erosion predictions. Earth Surface Processes and Landforms, 26: 1333-1352. doi:10.1002/esp.266.

Budiman M. and Damien J.F., 2005. Estimating soil hydraulic properties and their uncertainty: the use of stochastic simulation in the inverse modelling of the evaporation method. Geoderma, 126: 277-290. doi: 10.1016/j.geoderma.2004.09.015.

Christensen S., 2004. A synthetic groundwater modelling study of the accuracy of GLUE uncertainty intervals. Nordic Hydrology, 35: 45-59.

Daniel C. and Wood F. S., 1971. Fitting Equations to Data. Wiley-Interscience, New York, USA.

Dzotsi K.A., Basso B., and Jones J.W., 2013. Development, uncertainty and sensitivity analysis of the simple SALUS crop model in DSSAT. Ecological Modelling, 260, 62-76. doi:10.1016/j.ecolmodel.2013.03.017.

Gneiting T., Raftery A.E., Balabdaoui F., and Westveld A., 2003. Verifying probabilistic forecasts: Calibration and sharpness. Proc. Workshop on Ensemble Forecasting, 69: 243-268.

Hansson K. and Lundin L.C., 2006. Equifinality and sensitivity in freezing and thawing simulations of laboratory and in situ data. Cold Regions Sci. Technol., 44: 20-37. doi: 10.1016/j.coldregions.2005.06.004.

Jia Y. and Culver T.B., 2012. Robust optimization for total maximum daily load allocations. J. Logic Language Information, 21: 117-139.

Jin X., Xu C.-Y., Zhang Q., and Singh V.P., 2010. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J. Hydrology, 383: 147-155. doi:10.1016/j.jhydrol.2009. 12.028.

Kuczera G. and Parent E., 1998. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. J. Hydrology, 211: 69-85.

Lamb R., Beven K., and Myrabo S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. Advances in Water Res., 22: 305-317. doi:10.1016/s0309-1708(98)00020-7.

Li L., Xia J., Xu C.-Y., and Singh V.P., 2010. Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. J. Hydrology, 390: 210-221. doi: 10.1016/j.jhydrol.2010.06.044.

Mantovan P. and Todini E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. J. Hydrology, 330: 368-381. doi:http://dx.doi. org/10.1016/j.jhydrol.2006.04.046.

Marquardt D.W., 1963. An Algorithm for Least Square Estimation of Non-Linear Parameters. J. Society Industrial Applied Mathematics, 11: 431-441.

McMichael C.E., Hope A.S., and Loaiciga H.A., 2006. Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. J. Hydrology, 317: 307-324. doi:10.1016/j. jhydrol.2005.05.023.

Mertens J., Madsen H., Feyen L., Jacques D., and Feyen J., 2004. Including prior information in the estimation of effective soil parameters in unsaturated zone modelling. J. Hydrology, 294: 251-269. doi: 10.1016/j.jhydrol.2004.02.011.

Minasny B. and Field D.J., 2005. Estimating soil hydraulic properties and their uncertainty: the use of stochastic simulation in the inverse modelling of the evaporation method. Geoderma, 126: 277-290. doi:http://dx.doi.org/10.1016/j. geoderma.2004.09.015.

Mohrath D., Bruckler L., Bertuzzi P., Gaudu J.C., and Bourlet M., 1997. Error analysis of an evaporation method for determining hydrodynamic properties in unsaturated soil. Soil Sci. Soc. America J., 61: 725-735.

Montanari A., 2005. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. Water Resources Res., 41, 224-236, doi:10.1029/2004wr003826.

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

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

Schaap M.G. and Leij F.J., 2000. Improved prediction of unsaturated hydraulic conductivity with the Mualem-van Genuchten model. Soil Sci. Soc. America J., 64: 843-851.

Shi L., Song X., Tong J., Zhu Y., and Zhang Q., 2015. Impacts of different types of measurements on estimating unsaturated flow parameters. J.f Hydrology, 524: 549-561.

Stedinger J.R., Vogel R.M., Lee S.U., and Batchelder R., 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Res. Res., 44. doi:10.1029/ 2008wr006822.

Šimůnek J., van Genuchten M.T.V., and Wendroth O., 1998. Parameter estimation analysis of the evaporation method for determining soil hydraulic properties. Soil Sci. Soc. America J., 62: 894-905.

Tamari S., Bruckler L., Halbertsma J., and Chadoeuf J., 1993. A simple method for determining soil hydraulic properties in the laboratory. Soil Sci. Soc. America J., 57: 642-651.

van Genuchten M.T.V., Leij F.J., Yates S.R., and Williams J.R., 1991. The RETC Code for Quantifying Hydraulic Functions of Unsaturated Soils.

Wendroth O., Ehlers W., Hopmans J.W., Kage H., Halbertsma J., and Wosten J.H.M., 1993. Reevaluation of the evaporation method for determining hydraulic functions in unsaturated soils. Soil Sci. Soc. America J., 57: 1436-1443.

Weynants M., Vereecken H., and Javaux M., 2009. Revisiting vereecken pedotransfer functions: introducing a closed-form hydraulic model. Vadose Zone J., 8: 86-95. doi: 10.2136/vzj2008.0062.

Xiong L., O'Connor K.M., and Kieran M., 2008. An empirical method to improve the prediction limits of the GLUE methodology in rainfall-runoff modeling. J. Hydrology, 349: 115-124.

Yang J., Reichert P., Abbaspour K.C., Xia J., and Yang H., 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. J. Hydrology, 358: 1-23. doi:10.1016/j.jhydrol.2008.05.012.

Younes A., Mara T.A., Fajraoui N., Lehmann F., Belfort B., and BeydounH., 2013. Use of global sensitivity analysis to help assess unsaturated soil hydraulic parameters. Vadose Zone J., 12, 27-38, doi:10.2136/vzj2011.0150.

Zhang D., Beven K., Mermoud A., 2006. A comparison of nonlinear least square and GLUE for model calibration and uncertainty estimation for pesticide transport in soils. Advances Water Res., 29: 1924-1933.