Sensitivity Analysis and Identification of Parameters to the Van Genuchten Equation

Guangzhou Chen,1,2 Lifeng Jiao,2 and Xinhai Li2

1National Experimental Teaching Center of Water Pollution Control and Wastewater Resource, Anhui Jianzhu University, Hefei 230601, China
2School of Environment and Energy Engineering, Anhui Jianzhu University, Hefei 230601, China

Correspondence should be addressed to Guangzhou Chen; chgzhs@163.com

Received 18 July 2016; Revised 20 October 2016; Accepted 23 October 2016

Van Genuchten equation is the soil water characteristic curve equation used commonly, and identifying (estimating) accurately its parameters plays an important role in the study on the movement of soil water. Selecting the desorption and absorption experimental data of silt loam from a northwest region in China as an instance, Monte-Carlo method was firstly applied to analyze sensitivity of the parameters and uncertainty of model so as to get the key parameters and posteriori parameter distribution to guide subsequent parameter identification. Then, the optimization model of the parameters was set up, and a new type of intelligent algorithm—difference search algorithm was employed to identify them. In order to overcome the fault that the base differencesearch algorithm needed more iterations and to further enhance the optimization performance, a hybrid algorithm, which coupled the difference search algorithm with simplex method, was employed to identification of the parameters. By comparison with other optimization algorithms, the results show that the difference search algorithm has the following characteristics: good optimization performance, the simple principle, easy implement, short program code, and less control parameters required to run the algorithm. In addition, the proposed hybrid algorithm outperforms the basic difference search algorithm on the comprehensive performance of algorithm.

1. Introduction

The suction of unsaturated soil water is a function of soil moisture ratio and the curve that represents its relevance is described as the soil water characteristic curve or the soil water holding curve. The curve reflects the physical properties of soil such as the soil moisture retention property and the soil formation, so it is an important indicator of the soil basic hydraulic characteristic, and it is also the foundation of obtaining other soil moisture kinetic parameters and the soil moisture content. It is of great importance to the research of soil moisture retention; therefore, the curve has invariably been the hot issue of research in the fields such as soil physics, geohydrology, conservancy engineering, irrigation, and water conservancy [1, 2].

There are many complex factors that influence the soil water characteristic curve, so it is difficult to deduce the exact relation theoretically. Therefore, the researchers put forward many empirical functions to describe the relevance based on the empirical data by great amount of experimental researches. It included the common models: Brooks-Corey model [3], Gardner model [4, 5], Van Genuchten model [6], and Gardner-Russo model [7]. In numerous mathematical models established, Van Genuchten model was widely used because of the model’s good fitting degree to the measured data and the definite parametric meanings [8, 9].

In practice, the key to apply the model is to obtain four parameters of the model accurately. To estimate parameters of the model, the scholars have presented many methods to solve the issue above. In terms of traditional optimization algorithms, Wang et al. [10], Li et al. [11], and Xiao et al. [1] used the traditional optimization—simplex method to carry out the model parameters’ optimizing solving; Xu et al. [12] fitted the sandy soil moisture curve equation parameters with the Picard iteration. Ma et al. [13] used the nonlinear damping least square method to fit the equation parameters;
Wang et al. [14] used the pattern search algorithm which has stronger local search ability to identify the equation parameters. In terms of software applications, Peng and Shao [15] and Yang et al. [16] used the nonlinear fitting function in the MATLAB software to determine the model parameters; Liu et al. [17] obtained the parameters to be estimated by relying on the software of DPS (data processing system) based on the Marquardt method which combined the straight forward solution with the steepest descent method.

Because the equation was complex and the above parameters’ estimation belonged to the nonlinear problems, the conventional traditional optimization methods would tend to be trapped into local extremum; moreover, the algorithm was somewhat sensitive to the selection of initial iteration points, which often leads to the less algorithmic adaptability. In recent years, with the development of intelligent computing, many scholars have adopted the different heuristic intelligent algorithms to do it, among which the emerging swarm intelligence optimization algorithms have been used widely for their strong global optimization performance. Chen and Ma [18] and Du and Zhang [19] used particle swarm optimization (PSO) and stochastic particle swarm optimization to solve the equation parameters optimization, respectively. Liao et al. [20] used genetic algorithm and simulated annealing algorithm to estimate parameters. Xu et al. [21, 22] used real-coded multisubpopulation genetic algorithm and differential evolution algorithm to optimize the equation parameters, respectively. In order to further improve the optimal performance of swarm intelligence algorithm, some ameliorative and hybrid intelligent algorithms were put forward to solve the question. Guo et al. [23] structured a genetic algorithm combined with Levenberg-Marquardt algorithm to solve the equation parameters, which turned out that the hybrid genetic algorithm were improved in the calculated performance. Liu et al. [24] put forward an improved particle swarm optimization (based on dynamic neighbors and local search strategy), which showed that the method had better solving accuracy. Li et al. [25] identified the parameters based on hybrid immune genetic algorithm which was the combination of nonlinear fitting function and immune genetic algorithm (LIGA). Xing et al. [26] introduced harmony search (HS) algorithm and put forward a kind of harmony search optimization calculation method (IHGS) based on global information to solve the above problem. Mo et al. [27] used improved artificial glowworm swarm optimization algorithm to solve the parameters.

The above achievements evaluated the results of the model according to the fitting effect given by optimization algorithms, so they lacked uncertainty analysis for the model, which was the essential link in the process of modeling. To evaluate the uncertainty of model parameters and every parameter’s quantitative impact on the model results, the parameter sensitivity analysis was very necessary. According to the results of the sensitivity analysis, it could be determined what were the main factors that affected the accuracy of the model, and it could be conducive to cognition the characteristics of the model [28]. In addition, the above intelligent optimization algorithms needed to set many running control parameters of algorithm at running. If the control parameters of algorithm were set unreasonable, it will not to get better optimization results. Moreover, some intelligent algorithms’ calculation principles were relatively complex, they were difficult for the realization of programming, and the program codes were long, and since the program was long, it resulted the inconvenience of the application. Therefore, it is still valuable research task to continue to seek a new optimization algorithm, which has the simple calculation principle, few control parameters of algorithm, and good optimization performance.

In view of the above two reasons, this paper firstly conducted sensitivity analysis of parameters of VG model and uncertainty analysis of the model. Then, one used a new type of swarm intelligence optimization-difference search algorithm for the identification of model parameters. To further enhance the optimal performance, the basic difference search algorithm was improved to compose a hybrid algorithm by mixing with simplex method. The optimization results were compared with the ones of other swarm intelligent algorithms.

2. Van Genuchten Equation

VG model was put forward by the scholar Van Genuchten in 1980 [6]; it describes the relation between energy and quantity in the soil water, and it can be expressed as the following specifically:

\[ \theta = \theta_r + \frac{\theta_s - \theta_r}{1 + (\alpha h)^{m}}, \]

where \( \theta \) is for the volume moisture content (cm\(^3\)·cm\(^{-3}\)), \( h \) is soil matrix potential, \( \theta_r \) is residual water content (cm\(^3\)·cm\(^{-3}\)), \( \theta_s \) is the saturated water content of the soil (cm\(^3\)·cm\(^{-3}\)), \( \alpha \) and \( n \) are the fitting shape parameters of soil water holding curve, and \( m = 1 - 1/n, n > 1 \).

3. Principle of Hybrid Algorithm

3.1. Basic Difference Search Algorithm (DSA). Difference search algorithm (DSA) was a kind of new algorithm developed to apply for numerical optimization, and it was put forward by scholar Civicioglu [29] in 2012. The idea comes from the idea of the organisms’ migration, which uses a moving behavior similar to Brownian motion. Subjected to the constraints of regional survival resources, many organisms show the seasonal migration behavior in a year. In the process of migration movement, a lot of migration organisms make up a superorganism. By the migration behavior, they move from one habitat to a richer place, which is more helpful to survive. The evolution of group behavior is accompanied with the process of artificial super organisms’ continuous migrating to the global optimal solution. The detailed calculation steps of basic difference search algorithm are referred to the literatures [29, 30]. About the optimization performance for continuous unconstrained optimization problems, comparison studies were done in [29] between DS algorithm and eight widely used algorithms, including PSO, ABC, DE, and Gravitational Search Algorithm (GSA), and...
the results showed that DS algorithm was more powerful [29, 31]. For the constrained global optimization problems, Liu et al. [31] further developed it to solve generalized constrained optimization problems.

3.2. Modified Difference Search Algorithm. The key of a meta-heuristic algorithm is the ability of the following two aspects: exploitation and exploration. The basic difference search algorithm has the strong ability of global optimization; however, its convergent speed is slow, which means it needs more iteration to get better optimization results. Simplex is a direct search optimization method, which has fast local searching optimization ability. Therefore, in order to make full use of the global search and local search ability, the hybrid algorithm composed of difference search algorithm and simplex method was proposed. Chen et al. employed this improved algorithm to identify parameters of the 2-chlorophenol oxidation model, and the result showed it was powerful [32].

This paper tested two mixed ways: (1) in series combination, we first used difference search algorithm to obtain the approximate optimal solution and then used simplex method around the optimal solution to search again the local in order to get a better solution; (2) the embedded mode, after every iteration of the difference search algorithm, did the local search by simplex method to the optimal solution of this iteration calculation, and if it is better than the original solution, replace the old one. Afterwards, run the difference search and keep the circulation until termination condition was satisfied.

4. Case Study

4.1. Experimental Data. In the soil water characteristic curve, the four independent undetermined parameters \( \theta_1, \theta_2, \alpha, \) and \( n \) needed to be identified. The experimental data in the literature [10] were selected to do the experiment, which was the desorption and absorption experimental data of silt loam from a northwest region; details were shown in Table 1.

4.2. Objective Function. The form of objective function is as follows:

$$\min \quad f = \sum_{i=1}^{k} \left[ \hat{\theta}_i (h) - \theta_i (h) \right]^2$$

s.t. \( m = 1 - \frac{1}{n}, \quad n > 1, \)

where \( \hat{\theta}_i (i = 1, 2, \ldots, k) \) is the actual measuring soil moisture content, \( \theta_i (i = 1, 2, \ldots, k) \) is the moisture content of VG equation calculation, \( k \) is for the number of experimental data, and \( n \) is the fitting shape parameters of soil water holding curve.

4.3. Sensitivity Analysis of Model Parameters. The Monte-Carlo Analysis Toolbox is developed by Wagener et al., and it is a comprehensive analysis and visualization tool, which adopts MATLAB GUI technology [33]. The tool can analyze the structure of model, the sensitivity, the parameters, and the uncertainty of output. The core content of toolbox adopts the concept of regional sensitivity analysis (RSA) and extends the generalized likelihood uncertainty analysis technique (GLUE) [34]. The paper used the MCAT to carry out sensitivity analysis and uncertainty analysis. The sampling number was set 10000, the square error’s judging threshold number was set 10000, the square error’s judging threshold was set 1, and the following results were obtained.

4.3.1. Dotty Plots of Objective Function. Horizontal axis of Figure 1 and other figures showed the values of four variables (\( x_{1mc-\theta_1}, x_{2mc-\theta_2}, x_{3mc-\alpha}, x_{4mc-n} \)), and vertical axis of Figure 1 showedsse (sum of squares for error), which was the objective function. Figure 1 presented the relationship that the value of objective function changed along with the change of the independent variable. Seen from Figure 1, the surface shape of dotty plots for \( \theta_1 \) (\( x_{1mc} \)) had relatively minimum area, and the surface shape of dotty plots for the other three parameters had not obvious corresponding area. The above situation showed that parameter \( \theta_1 \) was easier to be identified, and the other three parameters were difficult to be identified.

4.3.2. Posteriori Parameter Distribution. The horizontal axis of Figure 2 showed the values of four parameters (\( x_{1mc-\theta_1}, x_{2mc-\theta_2}, x_{3mc-\alpha}, x_{4mc-n} \)), and the vertical axis of Figure 2 showed D(sse), which was the likelihood (transformed from sse). Firstly, the initial parameter distributions of the simulating group were sampled according to the uniform distribution, the range of every parameter was divided into 20 containers of equal width, and then the likelihoods in each container were added up and divided by the sum of parameter values within the container. In contrast, the posterior probability distribution was shown in Figure 2. Therefore, the regions of higher bars in Figure 2 showed that their corresponding parameter ranges had better model performance, which could be used to verify whether the result of optimization algorithm was scientific and reasonable.

4.3.3. Regional Sensitivity Analysis Plot. The vertical axis of Figure 3 was the value for cumulated normalized sse. The curves in Figure 3 were cumulative frequency distribution curves; the larger shape differences of the curves showed

<table>
<thead>
<tr>
<th>Type</th>
<th>0.5</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>50</th>
<th>100</th>
<th>130</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silty loam</td>
<td>Desorption</td>
<td>0.363</td>
<td>0.347</td>
<td>0.279</td>
<td>0.165</td>
<td>0.095</td>
<td>0.081</td>
<td>0.069</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>Absorption</td>
<td>0.362</td>
<td>0.340</td>
<td>0.269</td>
<td>0.141</td>
<td>0.082</td>
<td>0.069</td>
<td>0.061</td>
<td>0.056</td>
</tr>
</tbody>
</table>
**Figure 1:** Dotty plots of objective function ($x_1mc-\theta_s$, $x_2mc-\theta_s$, $x_3mc-\alpha$, $x_4mc-n$).

**Figure 2:** Posteriori parameter distribution.
that the parameter was more sensitive. Seen from Figure 3, parameter $\theta_r$ was the most sensitive parameter; the sensitivity of $\alpha$, $n$, and $\theta_s$ was decreased successively.

4.3.4. 3D Surface Plot of Every Two Parameters. The 3D plot in Figure 4 reflected the interaction relationship of any two parameters relative to objective function values. Seen from Figure 4, the low sse region of the two parametric comprehensive effect could be found, which revealed the better areas of parametric composite, and this could contribute to judging the rationality of the optimization results.

4.3.5. GLUE Variable Uncertainty Plot. Figure 5 reflected the cumulative probability distribution and probability density function of the output variables calculated using a selected objective (transformed to likelihood). The results showed peak output range under 0.21418 and above 1.1627 had 5% probability; that was to say, probability of laying in range [0.21418, 1.1627] was 95%.

Through the above sensitivity analysis, it contributed to identify the important model input control parameters and analyze the main uncertainty range of model output.

4.4. Identification of Model Parameters. The desorption and absorption experimental data of silt loam were used to identify the model parameters. In the experiment, the number of populations was set to 30, the total number of iterations was set to 1000, size factor was valued by the Gamma distribution (1, 0.5), and the rest were default parameters; therefore, the running control parameters of difference search algorithm were less. Taking into account the actual physical meaning of the parameters, the estimate range of the four parameters was $[0, 1]$, $[0, 1]$, $[0, 1]$, and $[1, 10]$. To prevent the effect of the random factors, 20 times were run, and the statistical results were shown in Table 2. The simplex method used the Nelder-Mead simplex method; the control parameters of simplex were as follows: total iteration was 3000, the termination tolerance was $1e^{-20}$, termination tolerance for the objective function value was $1e^{-20}$, and maximum number of function evaluations was 3000.

Through the experiments, it was found that the simplex method was very sensitive to the initial value. For DSA + simplex (series) method, parameter settings were the same as that of the basic algorithm; total iteration was 1000. For DSA + simplex method (embedded), the total number of iterations only was set 50 times by trying. Statistical results of the two improved algorithms were also shown in Table 2.
Figure 4: 3D surface plot of any two parameters.
Likelihood

Cumulative density plot (5% prediction quantile = 0.21418, 95% prediction quantile = 1.1627)

Density plot

Peak output

Figure 5: GLUE variable uncertainty plot.

Table 3: Objective function values obtained by different algorithms only for the desorption data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Objective function</th>
<th>Algorithm</th>
<th>Objective function</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLPSO [24]</td>
<td>2.3219e-03</td>
<td>Basic DSA</td>
<td>7.5404e-04</td>
</tr>
<tr>
<td>IAGSOA [27]</td>
<td>7.5404e-04</td>
<td>Two improved DSA</td>
<td>7.5404e-04</td>
</tr>
</tbody>
</table>

Seen from Table 2, basic difference search algorithm could search to the minimum and showed better optimization performance. Compared with the basic DSA, the performance of DSA + simplex (series) and DSA + simplex (embedded) had the better performance. To compare the optimization performance of algorithms, other algorithmic calculation results only for the desorption data were shown in Table 3.

Seen from Table 3, it can be concluded that basic DSA and improved DSA have greater optimization performance, better than those of most algorithms for the test of desorption data. In this paper, only for the desorption data, the identification results of four parameters successively were 0.0659, 0.3588, 0.0127, and 4.0443, respectively, and the objective function value was 7.5404e-04. Considering desorption and absorption processing, the objective function value for simplex method was 3.2721e-03 [10] and that of DNLPSO was 2.1317e-03 [24]; the objective function value of DSA in this paper was 1.9537e-03, and the identification results of four parameters successively were 0.0649, 0.3552, 0.0127, and 4.7480.

5. Conclusion

This paper selected the soil water characteristic curve equation to analyze the sensitivity of the model parameters and identify parameters. The main conclusions were as follows.

(1) Through the sensitivity analysis, the results showed that parameter $\theta_1$ was easier to be identified. Parameter $\theta_1$ was the most sensitive, followed by $\alpha$, $n$, and $\theta_2$. In addition, the posterior probability distribution plot of the parameters and the relationship plot of the parameters reflected the optimal area and position of the parameter distributions based on the objective function; it could be helpful to judge the identification results of intelligent optimization algorithms.

(2) The basic difference search algorithm had the advantages such as simple principle, easy realization, and few running parameters set, and it showed the better optimization performance through the practice application. The results of the improved algorithms which combined simplex in this paper showed that the proposed hybrid algorithm outperformed the basic difference search algorithm on the algorithm robustness and convergence.

Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

Acknowledgments

This work is supported by the following foundations: National Natural Science Foundation of China (Grant no. 41471422); Anhui Provincial Natural Science Foundation (Grant no. 1508085MD67); Key Project of Outstanding Young Talents of Universities in Anhui Province (gxyqZD2016148).
References


