AN APPLICATION OF CHAOS GRAY-ENCODED GENETIC ALGORITHM FOR PHILIP INFILTRATION MODEL

by

Kai-Wen WANG[a], Xiao-Hua YANG[b*], Yu-Qi LI[c], Chang-Ming LIU[a], and Xing-Jian GUO[a]

[a] Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Sciences, Beijing, China
[b] State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing, China
[c] College of Architecture and Landscape Architecture, Peking University, Beijing, China

Original scientific paper
https://doi.org/10.2298/TSCI1804581W

To improve the precision of parameters' estimation in Philip infiltration model, chaos gray-coded genetic algorithm was introduced. The optimization algorithm made it possible to change from the discrete form of time perturbation function to a more flexible continuous form. The software RETC and Hydrus-1D were applied to estimate the soil physical parameters and referenced cumulative infiltration for seven different soils in the USDA soil texture triangle. The comparisons among Philip infiltration model with different numerical calculation methods showed that using optimization technique can increase the Nash and Sutcliffe efficiency from 0.82 to 0.97, and decrease the percent bias from 14% to 2%. The results indicated that using the discrete relationship of time perturbation function in Philip infiltration model's numerical calculation underestimated model's parameters, but this problem can be corrected a lot by using optimization algorithm. We acknowledge that in this study the fitting of time perturbation function, Chebyshev polynomial with order 20, did not perform perfectly near saturated and residue water content. So exploring a more appropriate function for representing time perturbation function is valuable in the future.

Key words: *Philip infiltration model, gray-encoded genetic algorithm, USDA soil texture triangle, Chebyshev polynomial*

Introduction

The infiltration process happening on the land surface is an important part in hydrology, soil physics, environment, and agriculture science. Many empirical, semi-empirical, numerical, and approximate analytical models have been developed to stimulate the infiltration process [1-10]. Finding a simple, robust, efficient, and stable physical infiltration model and making it correctly stimulate the interaction between surface and subsurface hydrological cycle are the objectives of all the infiltration model studies.

Richards' equation (RE) [11] describes the moisture movement in unsaturated porous media, which is the well-known governing function of infiltration and unsaturated flow. Most of the physical infiltration models are derived from the RE, which can guarantee the physical meaning of different parameters and its measurability. The RE derived infiltration

* Corresponding author, e-mail: xiaohuayang@bnu.edu.cn

models can not only predict the infiltration process with measurable parameters, but also help understand the infiltration process such as the sorption's role in infiltration process. Philip [12-15] presented the first analytical solution from RE and gave a two-term infiltration model, *i. e.*, Philip infiltration model (PIM) with corresponding numerical methods. The PIM assumed a time perturbation series around sorption process and strictly defined the sorptivity to describe the sorption process. Based on the physical definition presented by Philip, infiltration process has become clearer [16]. The early stage of infiltration is controlled by capillary suction, sometimes it can be approximated by sorption. As time goes by, the gravity plays bigger role and the values of the second-term in PIM become larger.

The infiltration process, soil moisture movement, and soil physical properties are all described in complex non-linear relationship, and using optimization algorithm to estimate parameters is essential. Because genetic algorithm (GA) is a global optimization algorithm, different GA based on genetic evolution of species are universally applied in hydrology, environmental and soil science [17-19]. Sometimes, it is difficult for some GA to find the optimization solutions nearby the boundary of the search space quickly, which means GA would need a large amount of computation. To solve these problems, we use chaos gray-coded genetic algorithm (CGGA) proposed by Yang *et al.* [20] in this study, which combined chaos technology to reduce computations and improved the calculation precision in parameters' estimation.

Although Philip presented detailed numerical methods to calculate his model, PIM can be improved by using optimization algorithm. Using efficient computer optimization technique could improve the calculation accuracy and applicability of the model. The objective of this paper is to revisit PIM and introduce optimization algorithm, *i. e.* (CGGA), into it. Note, the application of optimization technique in this paper is different from calibration procedure in hydrological model which uses observation or other reference data to optimize parameters.

Governing equation in unsaturated soil

The water flow in unsaturated soil is traditionally described by RE, which combines the principle of water balance and the Buckingham-Darcy equation [21]. The modified form of RE is used widely to deduce the analytical solutions. The 1-D RE for vertical infiltration is given below, which describes soil moisture movement in a homogeneous, isotropic and rigid porous medium under isothermal conditions:

$$\frac{\partial z}{\partial t} = -\frac{\partial}{\partial \theta}\left[D(\theta)\frac{\partial \theta}{\partial z}\right] + \frac{\mathrm{d}K(\theta)}{\mathrm{d}\theta} \tag{1}$$

$$\theta = \theta_\mathrm{i}, \qquad t = 0, z \geq 0 \tag{2}$$

$$\theta = \theta_0, \qquad t > 0, z = 0 \tag{3}$$

where z [cm] is the vertical co-ordinate positive downward, z = 0 – the depth at soil surface, θ [cm^{3}cm^{-3}] – the volumetric water content, θ_i – the initial volumetric water content, θ_0 – the volumetric water content at soil surface, t – the time, $K(\theta)$ [cms^{-1}] – the unsaturated hydraulic conductivity, and $D(\theta)$ [cm^{2}s^{-1}] – the diffusivity.

The water retention characteristic and unsaturated hydraulic conductivity relationship are calculated by widely used van Genuchten [22] and Mualem [23] models. The van Genuchten and Mualem models are written in the following form:

$$\theta = \theta_\mathrm{r} + \frac{\theta_\mathrm{s} - \theta_\mathrm{r}}{\left[1 + |\alpha h|^n\right]^m} \tag{4}$$

$$D(\theta) = \frac{(1-m)K_\mathrm{s}}{\alpha m(\theta_\mathrm{s} - \theta_\mathrm{r})} \frac{\theta - \theta_\mathrm{r}}{\theta_\mathrm{s} - \theta_\mathrm{r}}^{\frac{1}{2} - \frac{1}{m}} \left[\left(1 - \frac{\theta - \theta_\mathrm{r}}{\theta_\mathrm{s} - \theta_\mathrm{r}}^{\frac{1}{m}}\right)^{-m} + \left(1 - \frac{\theta - \theta_\mathrm{r}}{\theta_\mathrm{s} - \theta_\mathrm{r}}^{\frac{1}{m}}\right)^{m} - 2 \right] \tag{5}$$

$$K(\theta) = K_\mathrm{s} \sqrt{\frac{\theta - \theta_\mathrm{r}}{\theta_\mathrm{s} - \theta_\mathrm{r}}} \left\{ 1 - \left[1 - \left(\frac{\theta - \theta_\mathrm{r}}{\theta_\mathrm{s} - \theta_\mathrm{r}} \right)^{\frac{1}{m}} \right]^m \right\}^2 \tag{6}$$

where h [cm] is the suction head, θ_r – residue water content, θ_s – saturated water content, K_s – the saturated hydraulic conductivity, α [cm^{-1}] – the scaling parameter that is inversely proportional to mean pore diameter, n – the soil water characteristic curve index, and $m = 1 - 1/n$.

Philip two-term infiltration model

Philip [5, 12] assumed that infiltration can be described as the time perturbation around sorption process, which caused by the effect of gravity. The analytical solutions of sorption is got by Boltzmann transformation, and adding the time perturbation series is:

$$z(\theta, t) = \eta_1(\theta)t^{1/2} + \eta_2(\theta)t^{2/2} + ...\eta_i(\theta)t^{i/2} + ... = \sum_{i=1}^{\infty} \eta_i(\theta)t^{i/2} \tag{7}$$

where $\eta_i(\theta)$ is the different function of θ, which can be numerically calculated by the iterative algorithm provided by Philip [15, 24]. In order to simplify the calculation process, only the first two terms in this series are used in infiltration calculation process.

Introducing eq. (7) into eq. (1) and combing the principle of continuity is the well-known Philip two-term infiltration model [5]:

$$I(t) = St^{0.5} + (K_\mathrm{i} + A_2)t = t^{0.5} \int_{\theta_\mathrm{i}}^{\theta_0} \eta_1(\theta)\mathrm{d}\theta + t \int_{\theta_\mathrm{i}}^{\theta_0} \eta_2(\theta)\mathrm{d}\theta + K(\theta_\mathrm{i})t \tag{8}$$

$$i(t) = 0.5St^{-0.5} + (K_\mathrm{i} + A_2) = \frac{1}{2} t^{-0.5} \int_{\theta_\mathrm{i}}^{\theta_0} \eta_1(\theta)\mathrm{d}\theta + \int_{\theta_\mathrm{i}}^{\theta_0} \eta_2(\theta)\mathrm{d}\theta + K(\theta_\mathrm{i}) \tag{9}$$

where S [cms$^{-0.5}$] is the sorptivity which is equal to $\int_{\theta_\mathrm{i}}^{\theta_0} \eta_1(\theta)\mathrm{d}\theta$, K_i is the hydraulic conductivity at initial water content and A_2 [cms^{-1}] is equal to $\int_{\theta_\mathrm{i}}^{\theta_0} \eta_2(\theta)\mathrm{d}\theta$.

Philip introduced the numerical methods [15, 24] to calculate S and A_2 with an iterative process. But the numerical methods can only give a discrete relationship of $\eta_1(\theta)$, $\eta_2(\theta)$. Using different numerical integration methods influence the computational accuracy. To get the relationship of $\eta_2(\theta)$ needs to estimate the derivative of $\eta_1(\theta)$ (*i. e.* $\mathrm{d}\eta_1(\theta)/\mathrm{d}\theta$), and using different numerical methods also influence the computational accuracy. In order to improve the accuracy of estimation, we used optimization algorithm in the previous two processes. We used Philip model (Index 1, Index 2) to represent PIM using different methods. When index 1 is Y (or N), it will represent using (or not using) optimization algorithm to fit function $\eta_1(\theta)$ in sorptivity calculation. When index 2 is Y (or N), it will represent using (or not using) optimization algorithm to fit function $\eta_1(\theta)$ in $\mathrm{d}\eta_1(\theta)/\mathrm{d}\theta$ calculation.

Description of CGGA

In order to improve the computational precision, function $\eta_1(\theta)$ and $\mathrm{d}\eta_1(\theta)/\mathrm{d}\theta$ are estimated by complex non-linear expressions (*i. e.* Chebyshev polynomial with order 20 in this study). The CGGA is used to reduce the computational amount and improve the accuracy, in which chaos mapping is applied to generate initial population and new chaos mutation and Hooke-Jeeves evolution operation are used in the subsequent steps.

We simplify the non-linear optimization problem in this study:

$$\begin{aligned} &\mathrm{Min}[f(X)] \\ &s.\,t. \quad a_j \le x(j) \le b_j \end{aligned} \tag{10}$$

where $X = \{x(j),\ j = 1,\ldots,p\}$, $x(j)$ is the parameter to be optimized, f – the objective function, and the range of the j^{th} parameter $x(j)$ – the interval $[a_j,\ b_j]$. The details of this algorithm are described by Yang *et al.* [20], the parameters of CGGA are all the same with Yang *et al.* [20], and the simplified computational procedures are in fig. 1.

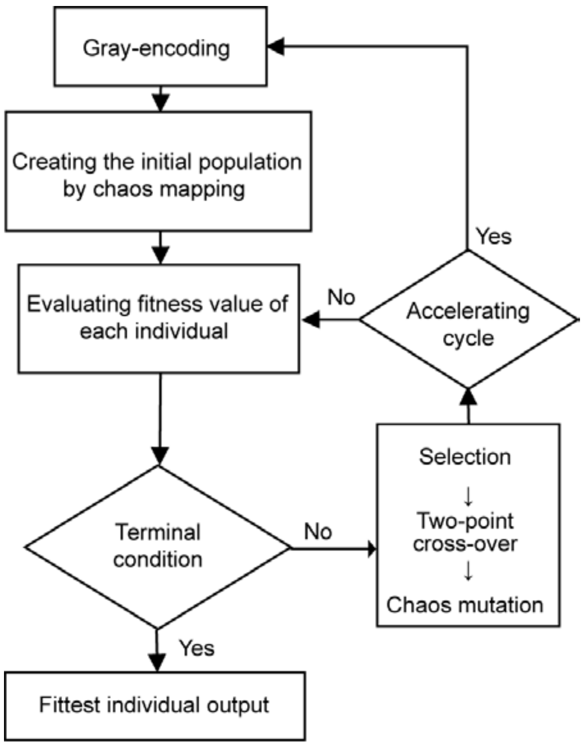

Figure 1. Computational procedures of CGGA

Experiment and data

The USDA textural triangle and the software RETC [25] were used to estimate the soil physical parameters of van Genuchten model, and the well-known HYDRUS-1D software [26] was applied to numerically stimulate the 1-D infiltration process. We did not stimulate soils with high sand content, so seven different soils in the USDA textural triangle were chosen to calculate in this study. Table 1 lists the various parameters of seven corresponding soils. The first column shows the soil texture of our numerical experiment, and the detailed information of different soil textures are shown in the next three columns. The van Genuchten model's parameters were predicted by RETC (bulk density of 1.5 g/cm^3). The last column of tab. 1 shows the total depths of simulated soils. In the HYDRUS-1D, all the soils were equally divided into one hundred parts with 0.16 as the initial condition and the saturation soil water content as the upper boundary value. The simulated time was 5 hour, and time step was 10 seconds.

Table 1. The parameters of seven soils for numerical experiments

Soil texture	Sand [%]	Silt [%]	Clay [%]	van Genuchten model's parameters					Stimulated depth in Hydrus-1D [cm]
				K_s (10^{-3})	θ_r	θ_s	α	n	
Silt	7.39	87.38	5.23	0.35394	0.0506	0.4076	0.0071	1.6217	80
Loam	41.16	40.54	18.30	0.12025	0.0559	0.3759	0.0108	1.4913	60
Silt loam	21.45	65.29	13.26	0.19954	0.0542	0.3751	0.0058	1.6355	80
Clay loam	32.50	34.00	33.50	0.06829	0.0793	0.4141	0.0127	1.3960	20
Silty clay loam	10.00	56.50	33.50	0.06192	0.0846	0.4290	0.0084	1.4873	20
Silty clay	6.66	46.67	46.67	0.05162	0.0938	0.4485	0.0125	1.3538	15
Clay	19.57	17.60	62.83	0.09144	0.0968	0.4446	0.0193	1.2180	15

Results and discussion

The percent bias (PBIAS) and Nash and Sutcliffe efficiency (NSE) [27] index were selected to evaluate the performance of PIM with different computational methods. The two statistics are calculated:

$$\mathrm{PBIAS} = \frac{\sum_{j=1}^{N}(Y_j^{\mathrm{ref}} - Y_j^{\mathrm{sim}})}{\sum_{j=1}^{N} Y_j^{\mathrm{ref}}} 100\% \tag{11}$$

$$\mathrm{NSE} = \frac{1 - \sum_{j=1}^{N}(Y_j^{\mathrm{ref}} - Y_j^{\mathrm{sim}})^2}{\sum_{j=1}^{N}(Y_j^{\mathrm{ref}} - \overline{Y})^2} \tag{12}$$

where Y_j^{ref} and Y_j^{sim} are the j^{th} reference and simulated values, $\overline{Y}$ – the mean of the reference data, and N – the total number. If the NSR > 0.5 and the absolute value of PBIAS is less than 25%, the model's performance can be identified as satisfactory [28].

Figure 2 shows the comparisons among PIM with different calculation methods and referenced cumulative infiltration (*i. e.*, Hydrus 1-D simulation results) of seven chosen soils. The reference solutions are plotted in black line with 10 secons time interval. Philip model (*Y*, *Y*), Philip model (*Y*, *N*), Philip model (*N*, *Y*), and Philip model (*N*, *N*) are plotted in different symbols showing in fig. 2 with 50 seconds time interval.

Results in fig. 2 showed that Philip model (*Y*, *Y*) performed best. Philip model (*Y*, *N*) and Philip model (*N*, *Y*) had similar performance and showed different advantages in different soils, which indicated two different accuracy improvement methods had different effect on different soils. The results indicated that if clay content was higher, the infiltration calculated by different methods were closer to the reference data. One possible explanation was the sorption process was the major part of infiltration [16] in the researched time scale, which means sorptivity took a large part in the infiltration process at least for the chosen soils.

Table 2 shows the PBIAS and NSE values of PIM with different numerical calculation methods. The evaluation results showed that optimization algorithm contributed a lot to improve the accuracy of PIM, and the Philip model (*Y*, *Y*) had the best results in all the

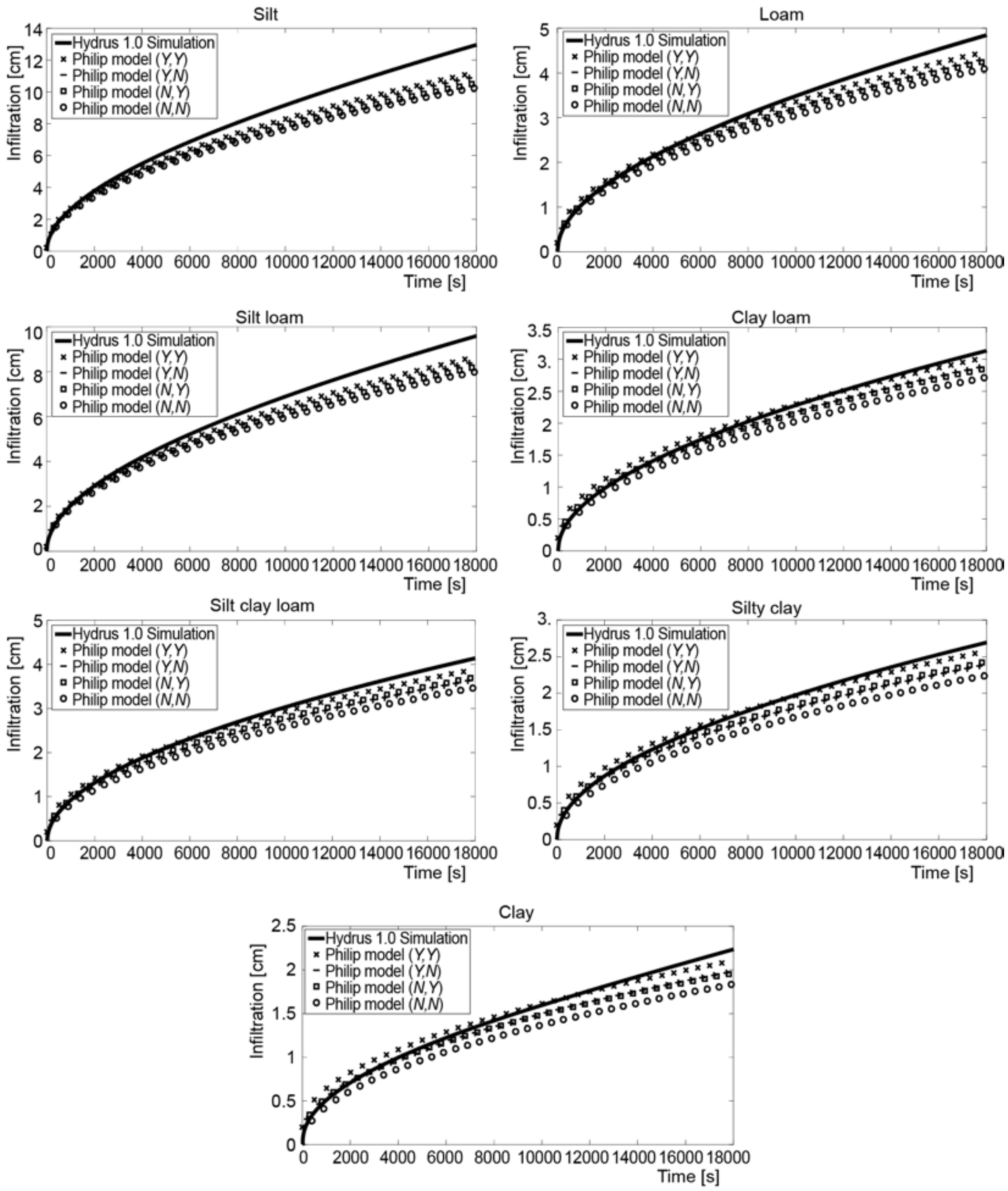

Figure 2. Comparisons of PIM with different calculation methods on cumulative infiltration

researched soils. The traditional numerical method, which was trapezoidal integration in this study, made Philip model (*N*, *N*) the worst performance. Philip model (*Y*, *N*) performed better than Philip model (*N*, *Y*) for silt, loam, silt loam, and clay loam, which implied that fitting function $\eta_1(\theta)$ for sorptivity calculation can improve the prediction accuracy of sorptivity. Because sorption controlled the infiltration process in the researched time scale, the four soils had better results with Philip model (*Y*, *N*) than Philip model (*N*, *Y*). Although sorption was still the major part of the infiltration process for silty clay loam, silty clay, and clay, the sorptivity is so small that it can be influenced by the values of function $\eta_2(\theta)$. So Philip model (*N*, *Y*) had better performance than Philip model (*N*, *Y*) for these three soils with high clay content. Compared with Philip model (*Y*, *Y*) and Philip model (*N*, *N*), using optimization technique can increase NSE from 0.82 to 0.97, and decrease PBIAS from 14% to 2%.

Table 2. Evaluation results of different soils and methods

Soil texture	Philip model (*Y*, *Y*)		Philip model (*Y*, *N*)		Philip model (*N*, *Y*)		Philip model (*N*, *N*)	
	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE
Silt	0.09	0.91	0.12	0.87	0.14	0.81	0.17	0.76
Loam	0.03	0.98	0.08	0.94	0.08	0.93	0.13	0.85
Silt loam	0.06	0.95	0.09	0.92	0.11	0.89	0.14	0.83
Clay loam	–0.02	0.99	0.05	0.97	0.05	0.96	0.11	0.88
Silty clay loam	0.02	0.98	0.09	0.91	0.08	0.94	0.15	0.81
Silty clay	0	0.99	0.08	0.92	0.07	0.95	0.15	0.79
Clay	–0.02	0.98	0.07	0.94	0.07	0.94	0.15	0.80
Mean value	0.02	0.97	0.08	0.92	0.09	0.92	0.14	0.82

Conclusion

This study revisited PIM and introduced CGGA into it. The algorithm was used in fitting function $\eta_1(\theta)$ and its derivative, which made the function relationship more precise, changing from discrete form to continuous form. Combining the USDA soil texture triangle, RETC and Hydrus-1D, we got the soil physical parameters and referenced cumulative infiltration for seven different soils. Using the referenced soil physical parameters and infiltration data, the comparisons among PIM with different numerical calculations were made. The comparisons showed that Philip model (*Y*, *Y*) with optimization algorithm performed best, and Philip model (*N*, *N*) with trapezoidal integration performed worst. Evaluation results indicated that using optimization technique can increase NSE from 0.82 to 0.97, and decrease PBIAS from 14% to 2%. We could conclude that using the discrete relationship of $\eta_1(\theta)$ in Philip model's numerical calculation underestimated the parameters, which could be corrected by optimization algorithm. Chebyshev polynomial with order 20 was used as fitting function to fit function $\eta_1(\theta)$ in this study, but the relationship near saturated and residue water content was not correct, which need to explore a more appropriate function for representing function $\eta_1(\theta)$ in the future.

Acknowledgment

This work was supported by the National Key Research and Development Program of China (No. 2017YFC0506603), the Key Program of National Natural Science Foundation of China (No. 41530635), and the General Program of National Natural Science Foundation of China (No. 51679007, 41330529).

References

[1] Green, W. H., *et al.*, Studies on Soils Physics: The Flow of Air and Water Through Soils, *J. Agric. Sci.*, *4* (1911), 1, pp. 1-24
[2] Yang, X. H., *et al.*, Vulnerability of Assessing Water Resources Based on the Improved Set Pair Analysis, *Thermal Science*, *18* (2014), 5, pp. 1531-1535

[3] Horton, R. E., Analysis of Runoff Plat Experiments with Varying Infiltration Capacity, *Trans. Am. Geophys. Union*, *20* (1939), 4, pp. 693-711
[4] Yang, X. H., *et al.*, Improved Gray-Encoded Evolution Algorithm Based on Chaos Cluster for Parameter Optimization of Moisture Movement, *Thermal Science*, *21* (2017), 4, pp. 1581-1585
[5] Philip, J. R., The Theory of Infiltration, *Adv. Hydrosci*., *5* (1969), pp. 215-296
[6] Parlange, J. Y., *et al.*, The Three-Parameter Infiltration Equation, *Soil Sci*., *133* (1982), 6, pp. 337-341
[7] Parlange, J. Y., *et al.*, Infiltration Under Ponded Conditions: 1. Optimal Analytical Solution and Comparison with Experimental Observations, *Soil Sci*., *139* (1985), 4, pp. 305-311
[8] Smith, R. E., *et al.*, *Infiltration Theory for Hydrologic Applications*, *Water Resour. Monogr*., AGU, Washington, D. C., 2002
[9] Serrano, S. E., Modeling Infiltration with Approximate Solutions to Richard's Equation, *J. Hydrol. Eng*., *9* (2004), 5, pp. 421-432
[10] Basha, H. A., Infiltration Models for Semi-Infinite Soil Profiles, *Water Resour. Res*., *47* (2011), 8, pp. 192-198
[11] Richards, L. A., Capillary Conduction of Liquids in Porous Mediums, *Physics*, *1* (1931), 5, pp. 318-333
[12] Philip, J. R., The Theory of Infiltration: 1. The Infiltration Equation and its Solution, *Soil Sci*., *83* (1957a), pp. 345-357
[13] Philip, J. R., The Theory of Infiltration: 2. The Profile of Infinity, *Soil Sci*., *83* (1957b), pp. 435-448
[14] Philip, J. R., The Theory of Infiltration: 4. Sorptivity and Algebraic Infiltration Equation, *Soil Sci*., *84* (1957c), pp. 257-264
[15] Philip, J. R. Numerical Solution of Equations of the Diffusion Type with Diffusivity Concentration Dependent, II. *Aust. J. Pyhs*., *10* (1957d), pp. 29-42
[16] Brutsaert, W., *Hydrology : An Introduction*, Cambridge Univ. Press, Cambridge, UK, 2005
[17] Horne, A., *et al.*, Optimization Tools for Environmental Water Decisions: A Review of Strengths, Weaknesses, and Opportunities to Improve Adoption, *Environmental Modelling & Software*, *84* (2016), Oct., pp. 326-338
[18] Yang, X. H., *et al.*, GHHAGA for Environmental Systems Optimization, *Journal of Environmental Informatics*, *5* (2005), 1, pp. 36-41
[19] Yang, X. H., *et al.*, Hierarchy Evaluation of Water Resources Vulnerability under Climate Change in Beijing, *Natural Hazards*, *84* (2016), Suppl. 1, pp. S63-S76
[20] Yang, X. H., *et al.*, Chaos Gray-Coded Genetic Algorithm and its Application for Pollution Source Identifications in Convection-Diffusion Equation, *Communications in Nonlinear Science and Numerical Simulation*, *13* (2008), 8, pp. 1676-1688
[21] Buckingham, E., Studies on the Movement of Soil Moisture, US Department of Agriculture, Bureau of Soils, Bull. No. *38*, Washington, D. C., 1907
[22] Van Genuchten, M. T., A Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Soil Sci. Soc. Am. J*., *44* (1980), 5, pp. 892-898
[23] Mualem, Y., A New Model for Predicting the Hydraulic Conductivity of Unsaturated Porous Media, *Water Resour. Res*., *12* (1976), 3, pp. 513-522
[24] Philip, J. R., Numerical Solution of Equations of the Diffusion Type with Diffusivity Concentration Dependent, *Faraday Soc. Trans.*, *51* (1955), pp. 885-892
[25] Van Genuchten, M. T., *et al.*, The RETC Code for Quantifying the Hydraulic Functions of Unsaturated Soils, Rep. EPA 600/2-91/065, *85*, U. S. Environ. Prot. Agency, Washington, D. C., 1991
[26] Simunek, J., *et al.*, The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-saturated Media-Version 4.08, Dep. of Environ. Sci., University of California, Riverside, Cal., USA, 2009
[27] Nash, J. E., *et al.*, River Flow Forecasting through Conceptual Models Part I – A Discussion of Principles, *J. Hydrol*., *10* (1970), pp. 282-290
[28] Ogden, F. L., *et al.*, Validation of Finite Water-Content Vadose Zone Dynamics Method Using Column Experiments with a Moving Water Table and Applied Surface Flux, *Water Resour. Res*., *51* (2015), 5, pp. 3108-3125

Paper submitted: December 5, 2017
Paper revised: December 12, 2017
Paper accepted: December 13, 2017