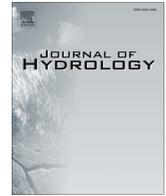




Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Modeling runoff and microbial overland transport with KINEROS2/ STWIR model: Accuracy and uncertainty as affected by source of infiltration parameters



A.K. Guber^{a,*}, Y.A. Pachepsky^b, A.M. Yakirevich^c, D.R. Shelton^b, G. Whelan^d, D.C. Goodrich^e,
C.L. Unkrich^e

^a Michigan State University, Department of Plant, Soil and Microbial Sciences, 1066 Bogue St., Room 584D, East Lansing, MI 48824, United States

^b USDA-ARS, Environmental Microbial and Food Safety Laboratory, 10300 Baltimore Avenue, Building 173, Beltsville, MD 20705, United States

^c Zuckerberg Institute for Water Research, J. Blaustein Institutes for Desert Research, Ben-Gurion University of Negev, Sede Boqer Campus, Beersheba 84990, Israel

^d U.S. Environmental Protection Agency, 960 College Station Road, Athens, GA 30605, United States

^e USDA-ARS, Southwest Watershed Research Center, Tucson, AZ 85719, United States

ARTICLE INFO

Article history:

Received 7 December 2013

Received in revised form 25 June 2014

Accepted 3 August 2014

Available online 12 August 2014

This manuscript was handled by

Konstantine P. Georgakakos, Editor-in-Chief

Keywords:

Bacteria
Runoff
Uncertainty
Modeling
Infiltration parameters

SUMMARY

Infiltration is important to modeling the overland transport of microorganisms in environmental waters. In watershed- and hillslope scale-models, infiltration is commonly described by simple equations relating infiltration rate to soil saturated conductivity and by empirical parameters defining changes in infiltration rate with soil water content. For the microbial transport model KINEROS2/STWIR used in this study, infiltration in unsaturated soil is accounted for by a net capillary drive parameter, G , in the Parlange equation. Scarce experimental data and multiple approaches for estimating parameter G introduce uncertainty, reducing reliability of overland water flow and microbial transport models. Our objectives were to evaluate reliability and robustness of three methods to estimate parameter G and associated accuracy and uncertainty in predicting runoff and fecal coliform (FC) transport. These methods include (i) KINEROS2 fitting to the experimental cumulative runoff data; (ii) estimating solely on soil texture; and (iii) estimating by individual pedotransfer functions (PTFs) and an ensemble of PTFs from basic soil properties. Results show that the most accurate prediction was obtained when the G parameter was fitted to the cumulative runoff. The KINEROS2-recommended parameter slightly overestimated the calibrated value of parameter G and yielded less accurate predictions of runoff, FC concentrations and total FC. The PTFs-estimated parameters systematically deviated from calibrated G values that caused high uncertainty in the KINEROS2/STWIR predictions. Averaging PTF estimates considerably improved model accuracy, reducing the uncertainty of runoff and FC concentration predictions. Overall, ensemble-based PTF estimation of the capillary drive can be efficient for simulations of runoff and bacteria overland transport when a single effective value is used across the study area.

© 2014 Elsevier B.V. All rights reserved.

Abbreviation: TMDL, Total Maximum Daily Load; *E. coli*, *Escherichia coli*; FC, fecal coliforms; CFU, colony forming units; MCU, microbial count units; BD, bulk density; VFS, vegetated filter strip; BC, Brooks–Corey; VG, van Genuchten; PTF, pedotransfer function; OC, organic carbon; AIC, Akaike Information Criterion; NSE, Nash–Sutcliffe efficiency index; MIA, modified index of agreement; STWIR, Simulator of Transport With Infiltration and Runoff; KINEROS2, Kinematic Runoff and Erosion Model; ANN, artificial neural networks.

* Corresponding author. Tel.: +1 517 355 0271x1269.

E-mail addresses: Akguber@msu.edu (A.K. Guber), yakov.pachepsky@ars.usda.gov (Y.A. Pachepsky), alexey@bgu.ac.il (A.M. Yakirevich), whelan.gene@epa.gov (G. Whelan), dave.goodrich@ars.usda.gov (D.C. Goodrich).

1. Introduction

Concerns about surface water safety stimulated the development of predictive models to estimate contaminant concentrations in runoff water during runoff events, and its transport to surface water sources. Among these are COLI (Walker et al., 1990), Hydrologic Simulation Program FORTRAN (HSPF) (Bicknell et al., 1997), Spatially Explicit Delivery Model (SEDMOD) (Fraser, 1999), PROMISE and WATNAT (Medema and Schijven, 2001), Water Assessment Model with ArcView interface (WAMView) (Bottcher and Hiscock, 2001), Loading Simulation Program in C++ (LSPC) (Tetra Tech, Inc.,

2002), Soil and Water Assessment Tool (SWAT) (Sadeghi and Arnold, 2002; Neitsch et al., 2005) and a recently developed add-on module, STWIR (Simulator of Transport With Infiltration and Runoff) (Guber et al., 2010), for the KINEROS2 model (Kinematic Runoff and Erosion Model) (Woolhiser et al., 1990; Semmens et al., 2008; Goodrich et al., 2012; www.tucson.ars.ag.gov/kineros).

The most important component of these models, infiltration, controls rainfall partitioning between overland and subsurface flow and, therefore, partitioning of contaminant fluxes. Among the most frequently used approaches to modeling watershed- and hillslope-scale infiltration into the soil are the Green and Ampt model (Green and Ampt, 1911) implemented in SWAT software, the Philips approach (Philips, 1957) utilized in HSPF and LSPC models, and the runoff curve number (US SCS, 1972) used in COLI. The KINEROS2/STWIR model implements the three-parameter infiltration equation (Parlange et al., 1982) to calculate rate of infiltration into the soil. Different theoretical concepts, scales, and input in various infiltration models introduce substantial uncertainty into simulation results. That is why evaluating uncertainty has become important to microbial water quality predictions in Total Maximum Daily Load (TMDL) modeling (Shirmohammadi et al., 2006).

The importance of soil infiltration capacity for contaminant removal from overland flow has been recognized in designing and evaluating efficiency of vegetation filter strips (VFS). For instance, Misra et al. (1996) determined that infiltration of inflow into VFS was the major factor in reducing herbicide overland transport. Roodsari et al. (2005) found high infiltration capacity of VFS in sandy loam soil significantly reduced FC transport, even on steep slopes. Sullivan et al. (2007) also concluded that infiltration was primarily responsible for attenuation of FC surface transport from pasturelands in manure-treated loamy soils. Fox et al. (2010) and Poletika et al. (2009) noted that pesticide removal efficiency of VFS was proportional to the fraction of infiltrated water (Sabbagh et al., 2009). Fox et al. (2011) studied *E. coli* transport in runoff boxes and found that bacteria mass reduction in the overland flow linearly increased with percent of infiltration. Munoz-Carpena et al. (2010) conducted a global sensitivity and uncertainty analysis in three VFS studies and showed that saturated hydraulic conductivity was the most important variable for predicting infiltration and runoff.

As a general understanding of infiltration's influence on overland contaminant transport has developed, the most attention has been paid to the saturated hydraulic conductivity. Relatively little has been learned about the effects that other parameters have in controlling infiltration into the soil, such as the net capillary drive G that affects soil unsaturated hydraulic conductivity in the Parlange infiltration equation (Parlange et al., 1982). The Parlange infiltration equation used in the KINEROS2/STWIR model has been successfully applied in assessing uncertainty of fecal coliform (FC) overland transport associated with spatially-variable bacteria concentrations in surface-applied manure (Guber et al., 2011) and uncertainty of FC bacteria removal from VFS under overland flow conditions (Guber et al., 2009).

The scarcity of experimental data and multiplicity of the approaches for estimating parameter G introduce uncertainty and reduce reliability of overland water flow and microbial transport models. Kim et al. (2013) developed field-scale values for rainfall-runoff parameters associated with the KINEROS2/STWIR model from 36 small plots distributed throughout a field over four seasons, resulting in 144 runoff experiments. Inherent difficulties included reliance on a nonlinear runoff model, modeling parameters that are difficult to estimate, a highly parameterized model with respect to available data, and a lack of concurrent field-scale monitoring. Typically, small plot-scale (runoff plots) measurements of infiltration during storms produce considerably more

overland flow as compared to hillslope or small catchment scale measurements. As compared to hillslopes or small catchments, small plots have relatively smooth and more uniform surfaces, less heterogeneous rainfall inputs (Sidle et al., 2007; Gomi et al., 2008a, 2008b), less heterogeneous soil physical properties (including effects of interconnected preferential flow that influences infiltration), magnified effects of exposed rocks, roots, and organics, burden effects (along pots) that tend to increase overland runoff, and soil crusts or hydrophobic phenomena which are magnified. This study complements the plot-scale studies by investigating field-scale values for rainfall-runoff parameters associated with the KINEROS2/STWIR model and aims to fill the gap between how data sources and methods of parameter G estimation influence uncertainty of microbial concentration predictions in the surface runoff. Our objectives were to evaluate uncertainty of the net capillary drive associated with the sources and methods of the parameter estimation and examine how this uncertainty translates into accuracy and uncertainty of runoff and microbial overland transport predictions.

2. Materials and methods

2.1. Overland microbial transport model

Overland transport of manure-born microorganisms in this study was computed based on advection–dispersion equations, coupled with the kinematic wave equation. The following processes were included in the microbial model:

- microorganism release from surface applied manure;
- advective–dispersive transport of microorganisms in runoff water;
- infiltration of microorganisms into the soil;
- straining of microorganisms from infiltrating water by plant litter and vegetation layer;
- exchange of microorganisms between runoff water and the mixing zone of soil at the surface; and
- attachment and detachment of microorganisms in the solid phase.

The KINEROS2 (Woolhiser et al., 1990) model was implemented to simulate one-dimensional overland flow within an equivalent rectangle representing an arbitrarily shaped watershed with uniform or curvilinear slope profiles. The model employs a power-law equation to describe the relationship between the ponding depth, H , the storage of water per unit area [L], and flux, q , the water discharge per unit width [$L^{-2} T^{-1}$]:

$$q = \alpha H^m \quad (1)$$

where parameters α and m represent surface properties, i.e. slope, roughness, and flow regime. These parameters were computed based on the Chézy hydraulic resistance equation:

$$\alpha = CS^{1/2} \text{ and } m = 3/2 \quad (2)$$

where S is the surface slope [$L L^{-1}$], and C is the Chézy friction coefficient for overland flow.

Eq. (1) used in conjunction with the mass conservation equation:

$$\frac{\partial H}{\partial t} + \frac{\partial q}{\partial x} = R - I \quad (3)$$

where R is the precipitation rate [$L T^{-1}$], I is the infiltration rate [$L T^{-1}$], t is time [T], and x is the distance along the slope [L].

The infiltration rate is computed in KINEROS2 using the three-parameter infiltration equation (Parlange et al., 1982):

$$I = K_{sat} \left[1 + \frac{\sigma}{\exp(\sigma I^*/B) - 1} \right] \quad (4)$$

where K_{sat} is the saturated hydraulic conductivity [$L T^{-1}$], σ is the dimensionless parameter that represents soil type, I^* is the infiltrated depth (L), $B = (G + h_w)(\theta_s - \theta_i)$, G is the net capillary drive [L], h_w is surface water depth [L], and θ_i and θ_s are the initial and saturated soil water content, respectively [$L^3 L^{-3}$].

The add-on module STWIR developed for the KINEROS2 model employs the convection dispersion equation for bacteria cells in runoff in the form:

$$\frac{\partial HC_r}{\partial t} + \frac{\partial qC_r}{\partial x} = \frac{\partial}{\partial x} \left(a_l q \frac{\partial C_r}{\partial x} \right) - \frac{\partial S_m}{\partial t} - d(k_a C_r - k_d \rho S_s) - (1 - k_s) I C_r \quad (5)$$

where a_l is the dispersivity [L]; d is the thickness of the mixing zone [L], i.e., thickness of the soil surface layer that actively interacts with the overland flow; C_{ir} and C_r are cell concentrations in irrigation and runoff water, respectively, [$M L^{-3}$]; S_m is the cell concentration in the manure applied to the soil surface [$M L^{-2}$]; and S_s is the cell concentration in the solid phase of the soil mixing zone [$M M^{-1}$]; k_a and k_d are the attachment and detachment rates of bacteria at the solid phase [T^{-1}], respectively; ρ is the soil bulk density in the mixing zone [$M L^{-3}$]; and k_s is the bacteria straining coefficient, i.e., the fraction of infiltrated cells which are irreversibly filtered out by plant litter and vegetation layer.

The mass conservation equation of bacteria in the soil mixing zone is:

$$d\rho \frac{\partial S_s}{\partial t} = d(k_a C_r - k_d \rho S_s) + k_f(1 - k_s) I C_r \quad (6)$$

where k_f is the fraction of infiltrated cells which are filtered out within the soil mixing zone.

The net release of bacteria from surface-applied manure is assumed to be proportional to the precipitation rate:

$$\frac{\partial S_m}{\partial t} = -C_m R \quad (7)$$

where C_m is the concentration of released bacteria calculated according to the Bradford and Schijven model (2002):

$$C_m(t) = \frac{C_0 h_m \alpha_m}{R} (1 + \alpha_m \beta_m t)^{-(1+1/\beta_m)} \quad (8)$$

where h_m is the depth of applied manure (L), α_m (T^{-1}) and β_m dimensionless are parameters defining the shape of the release curve, and C_0 is the initial content of bacteria cells in the applied manure ($M L^{-3}$). Applicability of this model has been confirmed for different irrigation rates in field experiments (Guber et al., 2006) where values of α_m were found to be closely related to the irrigation rate R , according to the linear regression equation

$$\alpha_m = 0.036 + 0.860R, \quad R^2 = 0.988 \quad (9)$$

for irrigation rates ranging from 2.51 to 6.93 $cm h^{-1}$.

Since bacteria transport in this study was modeled for a single runoff event that lasted 4.3 h, microorganism die-off in manure, runoff water, soil mixing zone and soil solid phase was disregarded in simulations.

Eqs. (1)–(8) were solved numerically using the implicit finite difference scheme. Initial and boundary conditions were set as described below. The KINEROS2 numerical code (Woolhiser et al., 1990; Semmens et al., 2008; Goodrich et al., 2012) was used to solve the infiltration and runoff flow Eqs. (1)–(4). The boundary conditions of zero ponding depth at $x = 0$ and an initial condition of $H = 0$ for $x > 0$ were used. The front limitation algorithm (Haefner et al., 1997) was applied to solve bacteria transport Eqs. (5)–(7). The Dirichlet boundary condition of zero

concentration and the Neumann boundary condition of zero concentration gradient were set at the inlet and outlet boundaries, respectively, and a constant initial concentration $C_r = 0$ at $x > 0$ was assumed for Eq. (5) with the absence of surface water.

2.2. Field experiment

The study of fecal coliforms (FC) transport in overland flow after manure application was conducted at the OPE3 USDA-ARS research site (Beltsville, Maryland). Detailed descriptions of the study area, soil profile and experiment design can be found in Guber et al. (2011). In brief, a rainfall 21.1 mm deep and lasting 1.5 h generated runoff of 1.4 mm within 4.3 h from a 1.6 ha corn field on May 10 after applying dry bovine manure. An H-flume was installed at the field's edge to record the runoff hydrographs and to collect water samples. The hydrograph was recorded at 1 min interval, while water samples of 350 ml were collected at 5-min and 10-min time intervals during the first and the second hours after the initiation of runoff, respectively. Soil and manure samples were taken randomly from 25 locations across the field before and immediately after the manure application. FC concentrations were determined in the runoff samples, soil and manure extracts by plating 250 μL subsamples onto MacConkeys Agar and counting with an Autoplate 4000 spiral platter (Spiral Biotech, Bethesda, MD). In addition, 21 undisturbed soil samples were randomly taken from the top 10-cm layer across the study area to measure soil bulk density (BD) (Eijkelkamp Agrisearch Equipment BV, Giesbeek, The Netherlands), soil texture (Gee and Bauder, 1986), and organic carbon content (OC) (Nelson and Sommers, 1996).

2.3. Model calibration

The overland flow component of KINEROS2 was calibrated on data from the May 10 runoff event. An 84×190 m rectangular planar overland flow domain with a slope of 2.7% toward the flume was delineated using Arc/Info GIS hydrologic modeling tool watershed to simulate the FC transport with overland flow. Spatially-averaged soil properties and FC concentrations in soil and manure were used as parameters and initial conditions for model calibration. That is, soil characteristics of the simulation domain were regarded as uniform and, despite individual soil property measurements at 21 points, only a single set of soil parameters including a single value of parameter G was used in the simulations.

To calibrate the KINEROS2/STWIR model, we used PEST software (Doherty, 2005) that solved the inverse problem of minimizing the least square errors between the observed and simulated variables. These were the cumulative runoff for KINEROS2 and the FC concentrations in runoff water, normalized by the FC concentration in applied manure for the STWIR model, respectively.

2.4. Estimating the net capillary drive parameter G

Three sources of data were used to estimate the values of parameter G in the infiltration Eq. (4). The first was the measured cumulative runoff hydrograph. Its G value was obtained during the calibration of KINEROS2 with cumulative runoff data described in Section 2.3.

The second source was soil texture class. Its G value was estimated based on the analysis of Rawls et al. (1998). This approach is suggested in the KINEROS2 manual (Woolhiser et al., 1990). Since 20 of 21 sites were of the same soil texture class (sandy loam), sandy loam was an input to determine the value of G parameter. The estimates obtained with this source will be referred to later as KINEROS-recommended.

The third source was basic soil properties: soil texture, bulk density (BD), organic carbon content (OC), and sampling depth. Fourteen pedotransfer functions (PTFs) were used in the study to estimate *G* values from these soil properties (Table 1). Information used as an input to PTFs included: soil texture in all PTFs, BD in 11 PTFs, OC content in eight PTFs, and sampling depth in three PTFs. The PTFs utilized here to estimate the values of the parameter *G* were based on two common water retention equations, Brooks–Corey (1964) (BC) and van Genuchten (1980) (VG).

The first water retention equation is the Brooks–Corey (Brooks and Corey, 1964) model, which establishes a power relationship between soil volumetric water content θ and matric pressure *h* in the form:

$$\frac{\theta - \theta_r}{\phi - \theta_r} = \begin{cases} \left(\frac{h_b}{h}\right)^\lambda, & h > h_b \\ 1, & h \leq h_b \end{cases} \quad (10)$$

where ϕ is the soil porosity, cm³ cm⁻³; θ_r is the residual water content, cm³ cm⁻³; *h_b* is air-entry pressure, cm; and λ is pore size distribution index. The second is the van Genuchten water retention (van Genuchten, 1980) model:

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \frac{1}{[1 + (\alpha h)^n]^m} \quad (11)$$

where θ_s is the saturated water content, cm³ cm⁻³; α is a parameter corresponding approximately to the inverse of the air-entry value, cm⁻¹; and *m* and *n* are empirical shape-defining parameters. Parameter *m* in the van Genuchten model was computed as $m = 1 - 1/n$ in this study.

Seven of the 14 used PTFs were based on the BC model, two PTFs on the VG model, and five PTFs estimated pairs of water content and matric pressure data points to which the VG model was fitted. Overall, the total number of parameter *G* predictors varied from three to five in different PTFs. We intentionally selected the PTFs that use different soil properties and water retention models to avoid possible systematic bias in *G* estimates associated with specific soil properties of the BC or VG model. A pedotransfer function calculator CalcPTF (Guber and Pachepsky, 2010) was used to implement the PTF applications.

The Morel-Seytoux et al. (1996) equation was implemented to compute the net capillary drive *G* from parameters of the Brooks–Corey (BC) and van Genuchten (VG) models. For the BC model, the net capillary drive was:

$$G = h_b \frac{2 + 3\lambda}{1 + 3\lambda} \quad (12)$$

and for the VG model, *G* was:

$$G = \frac{0.046 \text{ m} + 2.07 \text{ m}^2 + 19.5 \text{ m}^3}{\alpha(1 + 4.7 \text{ m} + 16 \text{ m}^2)} \quad (13)$$

A single value of *G* was used for each model simulation; however, three approaches to generating *G* value estimation from the PTFs were implemented. In the first, a single value was computed by averaging soil properties measured in 21 locations, then estimating *G* values based on the soil property averages with each of the 14 PTFs (averaging 1):

$$G1_i = PTF_i(\text{Texture}^*, OC^*, BD^*, \text{Depth}) \quad i = 1, 14 \quad (14)$$

where * denotes average values of soil properties measured in the 21 locations, and *i* denotes the PTFs listed in Table 1. In the second approach, we computed *G* values for each location with each PTF, then averaged the results for each PTF (averaging 2) as:

$$G2_i = \frac{1}{21} \sum_{j=1}^{21} PTF_i(\text{Texture}_j, OC_j, BD_j, \text{Depth}) \quad i = 1, 14 \quad (15)$$

where *j* denotes locations. Estimates of the *G* parameter obtained by the two approaches are referred to later as PTF-estimated. The third approach consisted of obtaining the *G* value for each location using each PTF, then averaging the outcome of all PTFs for each location as:

$$G3_j = \frac{1}{14} \sum_{i=1}^{14} PTF_i(\text{Texture}_j, OC_j, BD_j, \text{Depth}) \quad j = 1, 21 \quad (16)$$

The *G* values estimated by this approach are referred to as ensemble-average.

To evaluate the effect of predictors on accuracy of the PTFs, model discrimination was performed using the Akaike Information Criterion for small sample size, *AIC_c* (Burnham, 2002; Burnham and Anderson, 2004) as:

$$AIC_c = -2 \log(\mathcal{L}(\theta)) + 2K + \frac{2K(K + 1)}{N - K - 1} \quad (17)$$

where $\mathcal{L}(\theta)$ is the likelihood of error function, *N* is the size of the dataset, and *K* is the number of parameters used in the PTFs. Eq. (17) was applied to the *G* value estimations from the soil properties measured in the 21 location with all PTFs.

Table 1
Pedotransfer functions (PTFs) applied to estimate the net capillary drive *G*, along with soil properties (predictors) used in each PTF.

PTF no.	Pedotransfer function	Water retention equation	PTF predictors					
			Sand	Silt	Clay	Organic carbon	Bulk density	Depth
BC-1	Saxton et al. (1986)	BC	+		+			+
BC-2	Campbell and Shiozawa (1992)	BC	+		+			+
BC-3	Rawls and Brakensiek (1985)	BC	+		+			+
BC-4	Williams et al. (1992)	BC	+		+			+
BC-5		BC	+		+	+		
BC-6	Oosterveld and Chang (1980)	BC	+		+			+
BC-7	Mayr and Jarvis (1999)	BC	+	+	+	+		+
VG-1	Wösten et al. (1999)	VG	+	+	+			+
VG-2		VG		+	+	+		+
VG-3	Tomasella and Hodnett (1998)	WH-VG		+	+	+		
VG-4	Rawls et al. (1982) ^a	WH-VG	+	+	+	+		
VG-5	Gupta and Larson (1979)	WH-VG	+	+	+	+		+
VG-6	Rajkai and Várallyay (1992)	WH-VG	+		+	+		+
VG-7	Rawls et al. (1983) ^a	WH-VG	+	+	+	+		+

BC = the Brooks and Corey model parameter estimation.

VG = the van Genuchten model parameter estimation.

WH = estimation of water content at fixed matric pressure.

^a Corrected for organic matter content according to Nemes et al. (2009).

2.5. Runoff and overland microbial transport simulations

The KINEROS2/STWIR model with calibrated transport parameters described in the previous sections was implemented to simulate the runoff and the FC transport with overland flow for the studied rainfall event. Model accuracy in predicting runoff and FC concentrations was evaluated using Nash–Sutcliffe efficiency index (NSE) and the modified index of agreement (MIA), given by Legates and McCabe (1999):

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (P_i - O_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (18)$$

$$\text{MIA} = 1 - \frac{\sum_{i=1}^N |P_i - O_i|}{\sum_{i=1}^N |P_i - \bar{O}| + \sum_{i=1}^N |O_i - \bar{O}|} \quad (19)$$

where P_i and O_i represent simulated and observed values of state variables (cumulative runoff, FC concentration), \bar{O} is the mean of O_i and N is number of observations. The model performance was considered acceptable for $\text{MIA} > 0.75$ (Köhne et al., 2005).

The uncertainty of parameter G and the associated uncertainties in predicting runoff and FC overland transport by the KINEROS2/STWIR model were evaluated for the PTF-estimated and the ensemble-average G values. The measures characterizing uncertainties were standard deviations of the estimated G values (σ_G), simulated total runoff (σ_R), and the total FC transported with runoff (σ_{FC}). The three approaches of obtaining G value from the PTFs allowed us to evaluate two sources of uncertainty: the uncertainty associated with spatial variability of soil properties and the uncertainty associated with selection of the PTFs.

3. Results and discussion

3.1. Model calibration

The KINEROS2 calibration produced the following values of the overland flow parameters: the Chézy friction coefficient $C = 0.29$, saturated hydraulic conductivity $K_{sat} = 4.83 \text{ cm h}^{-1}$, net capillary drive $G = 10.0 \text{ cm}$, and the parameter $\sigma = 1.0$. The values of fitted STWIR parameters obtained via calibration were: dispersivity $a_L = 3.53 \text{ m}$, FC attachment and detachment rates $k_a = 0.613 \text{ h}^{-1}$ and $k_d = 0.005 \text{ h}^{-1}$, FC straining coefficients $k_s = k_f = 1$, and FC release parameter $\beta_m = 0.150$ (Guber et al., 2011).

3.2. Soil properties and estimation of parameter G from soil texture class

Soil properties measured in the 21 locations varied spatially. Even though 20 out of 21 samples belonged to the sandy loam texture class, measured sand content ranged from 49.2% to 73.3%, and silt content from 24.7% to 44.9% (Table 2), soil organic content varied between 1.4% and 2.0%, and soil bulk density ranged from 1.349 g cm^{-3} to 1.576 g cm^{-3} . The recommended value from the KINEROS2 manual for the parameter G was 13 cm for these soil properties.

3.3. Accuracy and uncertainty in estimating parameter G from pedotransfer functions

Both spatial variability in soil properties and differences in the PTFs resulted in highly variable values of G parameter estimates. The PTF-estimates based on the averaged soil properties (averaging 1) ranged over one order of magnitude (Table 3). In general, the PTFs that used the van Genuchten equation generated smaller G values (2.17–7.94 cm) compared to those obtained in

the KINEROS2 calibration (10 cm), as recommended in the KINEROS2 manual (13 cm) and computed using the Brooks and Corey model (5.98–33.76 cm) values. Among the BC estimates, two PTFs (BC-6 and BC-7) underestimated and the other five PTFs (BC-1 through BC-5) overestimated calibrated and KINEROS2-recommended values of the parameter G . Consistently smaller G values from the VG equation, compared to those obtained with the BC-based PTFs, were ultimately caused by parameter h_b , a multiplier in Eq. (12). It can be shown that equivalent h_b values estimated with the VG-based PTFs are smaller than the BC estimates. To demonstrate that, we employed relationships that convert parameters of the VG water retention model to the parameters of the BC model (Morel-Seytoux et al., 1996), we have:

$$\begin{aligned} p &= 1 + 2/m \\ h_b^* &= \frac{p + 3}{2\alpha p(p-1)} \left(\frac{147.8 + 8.1p + 0.092p^2}{55.6 + 7.4p + p^2} \right) \\ \lambda &= m/(1-m) \end{aligned} \quad (20)$$

The computed h_b^* values based on the PTF estimates of the van Genuchten parameters m and α ranged from 1.3 cm to 5.1 cm, and were considerably smaller than the h_b values estimated with the Brooks–Corey model (3.7–21.8 cm) for the averaged soil properties. Morel-Seytoux et al. (1996) have pointed out that the conversion Eq. (20) was developed to produce similar infiltration rates, as a function of time, for both BC and VG models, not necessarily to produce similar water retentions. On the contrary, the PTFs were designed to generate parameters for both models that produce similar relationships between soil water content and capillary pressure. Fig. 1a illustrates differences between h_b values obtained in fits of Eq. (10) to pairs (θ, h) estimated with the PTF-16 (Rawls et al., 1983), and h_b^* converted from parameters m and α fitted to the same data. There is a linear relationship between h_b^* and h_b with high coefficient of determination ($R^2 = 0.959$), a slope of 0.12 and an intercept of 0.15, indicating that converted values of the BC parameter h_b were approximately eight times smaller than those fitted to the pairs (θ, h) . Differences between the BC- and VG-based PTF estimates were not drastic for parameter λ , yet converted values from the VG parameters systematically overestimated the BC parameter λ .

The calibrated value of the parameter G was used as a reference for the PTF estimates since the KINEROS2 fit to the experimental data presumably provided the smallest error of simulated cumulative runoff. For both methods of soil property averaging described in Section 2.4, the most accurate estimates of the parameter G were obtained with BC-5, BC-6 and BC-7, and with VG-4, VG-5 and VG-6 PTFs (Table 3). The ranked accuracy of the PTFs, computed as an absolute deviation of the estimate from the calibrated G value, showed that the lowest ranks (11–14) pertained to the 3-parameter BC-based PTFs, while the highest ranks pertained to the PTFs that used organic carbon content as a predictor (Table 4). This finding corroborates earlier PTF studies. For example, an overall increase in the PTF accuracy with an increase in the number of predictors was observed for soil water retention (Schaap et al., 1998; Twarakavi et al., 2009) and hydraulic conductivity prediction (Schaap and Leij, 1998).

The uncertainty of G estimates, evaluated as a standard deviation of the parameter G , differed for the water retention models. In general, the BC-based PTFs produced more uncertain G values compared to those VG-based. Interestingly, the largest standard deviation values of the net capillary drive (σ_G), and thus more uncertain results, were obtained with the PTFs that did not have organic carbon as a predictor (Tables 3 and 4). This probably is the reason for the overall improved performance of the VG PTFs compared to the BC PTFs since five VG-based, but only two BC-based PTFs, used organic carbon as a predictor.

Table 2

Soil properties in top 10-cm layer measured in 21 randomly selected locations.

Sampling location	Sand	Silt (%)	Clay	Bulk density (g cm ⁻³)	Organic carbon (%)
1	58.0	38.4	3.5	1.428	1.5
2	73.3	24.7	2.0	1.349	1.8
3	63.0	32.7	4.3	1.461	2.0
4	49.2	44.9	5.8	1.360	1.9
5	63.1	32.5	4.3	1.413	2.0
6	51.4	42.6	6.0	1.574	1.6
7	57.1	37.3	5.6	1.506	1.8
8	69.3	29.5	1.2	1.454	1.8
9	60.3	35.1	4.7	1.501	1.7
10	53.6	42.3	4.1	1.373	1.7
11	67.8	27.0	5.2	1.569	1.5
12	52.9	40.6	6.5	1.555	2.0
13	57.7	37.3	5.0	1.509	1.9
14	58.7	32.2	9.0	1.521	1.9
15	58.2	39.7	2.1	1.575	1.5
16	65.2	30.8	4.0	1.439	1.4
17	59.7	36.3	4.0	1.508	1.5
18	52.1	42.0	6.0	1.568	1.5
19	67.0	30.0	2.9	1.549	1.5
20	66.9	29.6	3.5	1.529	1.6
21	53.4	36.3	10.3	1.576	1.5
Average	59.9	35.3	4.8	1.491	1.7
σ	6.62	5.59	2.16	0.074	0.2

 σ stands for the standard deviation of soil properties.**Table 3**Values of net capillary drives G1 and G2 obtained using model calibration recommended by the KINEROS2 manual and estimated using PTFs, with performance statistics (AIC_c , NSE and MIA) of the KINEROS2/STWIR simulations conducted with the estimated G values.

Source of G	Parameter averaging 1					Parameter averaging 2					
	G1 (cm)	NSE		MIA		G2(σ_G) (cm)	AIC_c	NSE		MIA	
		Runoff	FC	Runoff	FC			Runoff	FC	Runoff	FC
KINEROS2 calibration	10.0	0.99	0.77	0.96	0.77						
KINEROS2-manual	13.0	0.94	0.76	0.87	0.76						
PTF-estimates											
BC-1	28.7	0.03	0.20	0.50	0.67	31.6 (13.1)	207	-0.17	0.15	0.48	0.67
BC-2	33.8	-0.32	0.15	0.46	0.67	35.0 (9.4)	209	-0.41	0.15	0.45	0.67
BC-3	24.7	0.30	0.74	0.54	0.74	25.7 (5.5)	189	0.24	0.70	0.53	0.73
BC-4	31.5	-0.16	0.15	0.48	0.67	32.4 (6.5)	204	-0.22	0.15	0.47	0.67
BC-5	13.5	0.93	0.77	0.85	0.77	15.1 (5.5)	155	0.86	0.78	0.79	0.77
BC-6	5.98	0.80	0.70	0.80	0.74	8.87 (8.58)	164	0.97	0.75	0.92	0.75
BC-7	6.24	0.82	0.71	0.81	0.74	6.16 (0.54)	135	0.82	0.71	0.81	0.74
VG-1	2.17	0.08	0.56	0.65	0.72	3.09 (1.50)	157	0.33	0.61	0.69	0.73
VG-2	4.11	0.54	0.65	0.73	0.74	4.30 (0.94)	152	0.58	0.66	0.74	0.74
VG-3	3.35	0.38	0.62	0.70	0.74	1.20 (0.73)	163	-0.28	0.48	0.61	0.70
VG-4	6.30	0.83	0.71	0.82	0.74	7.25 (3.26)	138	0.90	0.72	0.85	0.75
VG-5	7.39	0.91	0.73	0.86	0.75	7.97 (2.45)	126	0.94	0.74	0.88	0.75
VG-6	7.94	0.94	0.74	0.88	0.75	8.82 (5.01)	142	0.97	0.75	0.91	0.75
VG-7	3.20	0.35	0.61	0.69	0.73	3.39 (0.96)	158	0.40	0.62	0.70	0.74

G1, G2 and σ_G are the average and standard deviation of G estimates computed for each PTF, based on soil properties measured in 21 locations; BC = the Brooks and Corey model, VG = the van Genuchten model, NSE is the Nash–Sutcliffe index, MIA is the modified index of agreement, Runoff = results of runoff modeling, FC = results of fecal coliform modeling. Highlighted numbers indicate acceptable model performance.

The number of predictors affected accuracy of the PTFs. Values of AIC_c computed using Eq. (14) (Table 3) and the PTF ranks based on the AIC_c values (Table 4), showed that the most accurate estimates of parameter G were obtained with the 5-parameter PTFs that used organic carbon content as a predictor. Soil bulk density appeared to be another important predictor. The total number of PTFs used in this study is not sufficient to confirm this observation statistically, however, two PTFs (BC-5 and VG-3) that utilized organic carbon content but did not include bulk density, had relatively low AIC_c ranks.

We are not aware of any study of the influence of soil properties on the net capillary drive, however, several have reported the importance of bulk density for predicting saturated hydraulic conductivity (K_{sat}). For example, Rawls et al. (1992) indicated that bulk

density had a major effect on K_{sat} and in later work subdivided K_{sat} estimates for each soil texture class according to bulk density groups (Rawls et al., 1998). Organic matter content and bulk density also appeared most frequently as partitioning variables at the tertiary level of the regression tree developed for K_{sat} prediction by Lilly et al. (2007). Both bulk density and organic matter content were used as predictors of K_{sat} in the regression-based PTFs (Vereecken et al., 1990; Wösten et al., 1999) and in the artificial neural networks (Tamari et al., 1996). The effect of input variables on the accuracy of the VG-based PTFs predicting K_{sat} was examined in the Børgesen et al. (2008) study. These authors showed that the RMSE and AIC values decreased with the addition of bulk density and organic carbon content as predictors for K_{sat} in the PTFs developed for the USDA's soil texture classification. Results of our study

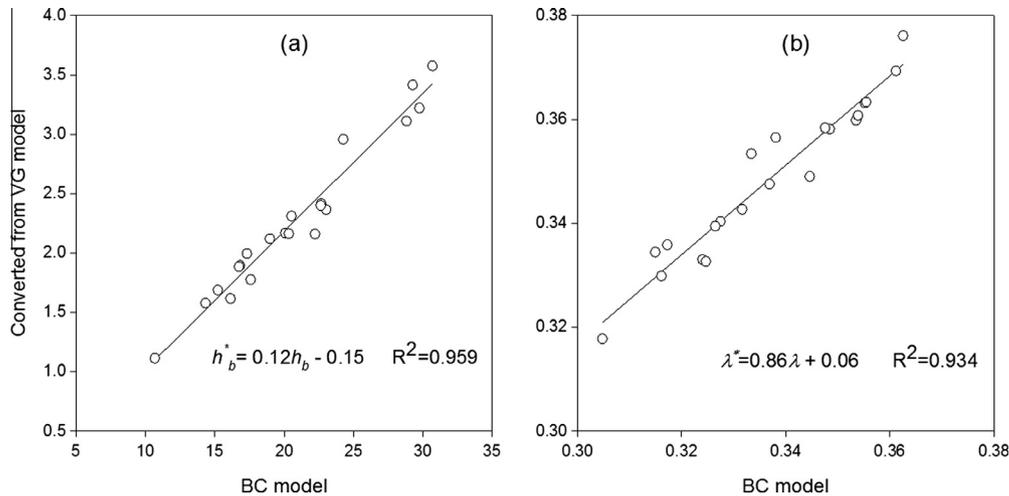


Fig. 1. Relationships between values of h_b (a) and λ (b) parameters obtained in fitting the Brooks–Corey (BC) model to pairs (θ, h) , estimated with the PTF-16 and by converting the van Genuchten (VG) model-fitted parameters to the same data.

Table 4
Rank of PTFs based on accuracy and certainty in estimating the net capillary drive G .

PTF	K	OC	Averaging 1	Averaging 2	σ_G	AIC_c
BC-1	3		12	12	14	13
BC-2	3		14	14	13	14
BC-3	3		11	11	10	11
BC-4	3		13	13	11	12
BC-5	3	+	3	6	9	8
BC-6	4		6	1	12	9
BC-7	5	+	5	5	1	2
VG-1	4		10	9	5	7
VG-2	5	+	7	7	3	5
VG-3	3	+	8	10	2	10
VG-4	5	+	4	4	7	3
VG-5	5	+	2	3	6	1
VG-6	4	+	1	2	8	4
VG-7	5	+	9	8	4	6

K is number of parameters; OC indicates presence of organic carbon in the list of PTF inputs, averaging 1–2 – PTF ranking by the results soil property averaging 1 and averaging 2 procedures; AIC_c – PTF ranking by the value of the Akaike criterion.

indicate that bulk density and organic carbon content can affect both saturated hydraulic conductivity and soil net capillary drive parameters.

Accuracy and uncertainty of the PTF-estimated G values associated with spatial variability of soil properties and with the PTF equations were approximately the same for the two averaging methods. The Kruskal–Wallis One Way Analysis of Variance did not reveal statistically significant differences between the parameter G values estimated by these two methods. Estimated G values ranged from 2.2 cm to 33.8 cm and 1.2 cm to 35.0 cm, with averages of 12.8 cm and 13.6 cm for averaging 1 and averaging 2 methods, respectively (Table 3). This result was surprising, because the parameter G is nonlinearly related to parameters of the BC and VG equations, which are, in turn, nonlinear functions of basic soil properties. Moreover, spatial variability of soil properties led to high variability of the estimated G values, as assessed by σ_G (Table 3), yet the G estimates obtained with the two averaging methods were very similar to each other.

Overall, the ensemble-averaging yielded more accurate estimates of parameter G , compared to averaging 1 and averaging 2 approaches. The ensemble-average estimates ranged from 7.8 cm to 20.1 cm at the 21 studied locations with an average of 13.6 cm (Table 5). The averaged G estimate obtained with the ensemble method was very close to the KINEROS2-estimate, probably

Table 5
Net capillary drive, G_3 , estimates and model performance statistics of the KINEROS2/STWIR simulations conducted with ensemble-average G estimates.

Sampling location	$G_3(\sigma_G)$ (cm)	Nm	NSE		MIA	
			Runoff	FC	Runoff	FC
1	12.0 (10.9)	28	0.97	0.77	0.91	0.76
2	7.8 (7.1)	11	0.93	0.73	0.87	0.75
3	11.0 (10.3)	108	0.99	0.76	0.95	0.76
4	12.8 (11.8)	18	0.95	0.77	0.88	0.77
5	10.0 (9.3)	$1.5 \cdot 10^6$	0.99	0.76	0.96	0.76
6	20.1 (19.0)	4	0.59	0.78	0.60	0.77
7	14.0 (13.1)	11	0.91	0.78	0.83	0.77
8	10.2 (8.6)	$1.1 \cdot 10^3$	0.99	0.76	0.96	0.76
9	12.9 (11.7)	16	0.95	0.77	0.87	0.77
10	12.1 (11.0)	26	0.97	0.77	0.91	0.76
11	12.2 (11.8)	27	0.97	0.77	0.90	0.76
12	18.3 (17.0)	5	0.70	0.78	0.66	0.78
13	14.0 (13.0)	11	0.91	0.78	0.83	0.77
14	13.7 (13.7)	14	0.92	0.77	0.85	0.77
15	17.8 (15.9)	5	0.72	0.78	0.68	0.78
16	11.4 (9.8)	44	0.98	0.77	0.93	0.76
17	13.3 (12.3)	14	0.93	0.77	0.86	0.77
18	19.3 (18.3)	5	0.64	0.78	0.63	0.77
19	12.3 (11.0)	23	0.96	0.77	0.90	0.76
20	11.7 (10.6)	36	0.98	0.77	0.92	0.76
21	19.4 (19.0)	5	0.63	0.78	0.62	0.77

G_3 and σ_G are the average and standard deviation of G estimates for each location, using individual soil properties. NSE and MIA are the Nash–Sutcliffe index and the modified index of agreement, respectively. Nm is the ratio between mean square errors of G values estimated using individual PTFs and mean square error of ensemble-average estimate. The highlighted numbers indicate acceptable model performance.

because most studied soil samples belonged to the same texture class. Interestingly, variability of the G values generated via the ensemble approach was relatively low compared to averaging the PTF estimates, despite an overall greater variability of G values associated with using different PTFs. The σ_G values characterizing spatial variability of the soil properties ranged from 0.5 cm to 13.1 cm for 14 PTFs (Table 3), while individual PTFs produced σ_G values from 7.1 cm to 19.0 cm (Table 5). This indicates that uncertainty associated with the PTF selection dominates the uncertainty associated with the spatial variability of soil properties in this study. Consistent with this observation is that high uncertainty in the estimates obtained using the same set of PTFs has been reported. Guber et al. (2006) estimated soil water retention for five depths within an 8-m trench, based on depth-averaged basic soil

properties, and showed that greater uncertainty in water content estimates was associated with the PTF equations than with the measured soil properties used in the PTFs.

The superior performance of an ensemble prediction over a single estimate observed in this study corroborates the studies of Perrone and Cooper (1993), Breiman (1996), Wagner et al. (2001), and Baker and Ellison (2008). Perrone and Cooper (1993) were probably the first to demonstrate theoretically that the ensemble approach can reduce the mean square error of the multimodel prediction by a factor of total number of models used (N_m). Breiman (1996) proposed and implemented a bootstrap aggregating (bagging) approach for classification trees and demonstrated that the aggregation reduced misclassification rates by up to 77%. The bagging approach has been shown to improve the predictive ability of artificial neural networks for soil hydraulic properties (Minasny et al., 2004). Wagner et al. (2001) examined the accuracy of eight PTFs in estimating K_{sat} and found that the predictions were most accurate when the results of two PTFs, which systematically over- and under-estimated the measured K_{sat} values, were combined in one prediction. Baker and Ellison (2008) showed that, for the HYPRES database, an ensemble with two artificial neural networks yielded more accurate estimates of water content at -24.5 kPa, compared to a single artificial neural network. They concluded that for the data investigated and the methods used, an ensemble greatly improved results compared to the methods using single models, even for a small number of ensemble members.

Following Perrone and Cooper (1993), we estimated the reduction factor N_m that is the ratio of mean square error of G values estimated using the individual PTFs to the mean square error of the ensemble-average estimate. The factor indicates the value of error that can be reduced by replacing the results of individual PTF predictions with the PTF ensemble-average values. The N_m factor computed for the 21 sampling locations used in the 14 PTFs varied over six orders of magnitude with mean $N_m = 28$ that was twice the total number of the pedotransfer functions used in the individual and ensemble predictions. The extremely large N_m values occurred in the sampling locations 5 and 8 where the differences between calibrated and ensemble-estimated values of the parameter G were the smallest (Table 5). Overall high performance of the ensemble-averaging was obviously the result of the systematic over- and under-estimation of the calibrated G values by different PTFs.

3.4. Accuracy and uncertainty in predicting cumulative runoff

The sources of data used to estimate net capillary drive also influenced the accuracy and uncertainty of cumulative runoff simulated with KINEROS2. As expected, the most accurate runoff predictions were achieved with calibrated values of the parameter G . Overall good results were obtained with KINEROS2-recommended parameters (Fig. 2a). These two simulations produced high values of both the efficiency index (NSE) and the modified index of agreement (MIA) (Table 3), indicating a good overall model performance. Since the PTF approaches produced highly variable estimates of the parameter G , they also led to higher uncertainty in the cumulative runoff (Fig. 2b–d). Values of NSE and MIA varied between -1.9 and 0.99 and between 0.37 and 0.96 , respectively, and decreased as the value of G departed from the calibrated one (Fig. 3). Based on the NSE index, the runoff prediction was acceptable when the parameter G ranged from 2.2 cm to 28.7 cm, while, based on the MIA index, it was acceptable when G ranged from 4.7 cm to 15.8 cm. Despite high variability in the soil properties, the KINEROS2 performance was better with G values estimated from the PTFs compared to the KINEROS-recommended value in several instances. Specifically, both performance indices were higher for the G values obtained using VG-6 PTF (averaging 1) and

for BC-6, VG-5 and VG-6 PTFs (averaging 2), compared to the indices obtained using the KINEROS-recommended G value.

The MIA index appeared to be more sensitive to the PTF source of G values in the KINEROS2 simulations than the NSE index. Only six of the 14 PTFs adequately represented the measured runoff hydrograph, according to the MIA values, while 12 PTFs produced positive NSE values and could be accepted in the simulations with spatially-averaged soil properties (averaging 1). Based on the MIA index, accurate runoff predictions were achieved with the PTFs that used both organic carbon and bulk density as the parameter G predictors.

The accuracy and uncertainty of the KINEROS2 results was also affected by the method of G estimation. Ensemble-averaging appeared to be more accurate and robust compared to the individual PTF estimating. The NSE values computed for runoff simulations with the ensemble-average parameters were positive for all locations and the MIA values were greater than 0.75 for 17 of 21 locations (Table 4), indicating 100% and 81% probability of acceptable model performance, respectively. Probabilities of acceptable model performance with the individual PTF-estimated parameters were 86% and 71%, based on the NSE index, and 43% and 43%, based on the MIA index, for averaging 1 and averaging 2 methods, respectively. The accuracy of the runoff simulation with the ensemble-averaged G values did not differ from that of the calibrated model in 10% of the studied cases and was better than the accuracy of the runoff simulation with the KINEROS2-recommended parameter in 52% of the cases. This was due to the ensemble-averaged G values being more similar to the calibrated G values. Those obtained from the KINEROS2-recommended and the PTF estimates deviated more substantially from the calibrated G values.

To the best of our knowledge, efficiency of the individual PTF and the ensemble-average parameter estimating approaches for runoff prediction has not been studied yet. A similar approach to the averaging 1 method used here was implemented by Guber et al. (2006). They utilized the PTFs that we used in the present study to estimate water retention parameters and to predict water content dynamics in an 8-m long trench using HYDRUS-1D software. They found that the averaging 1 method yielded more certain and accurate predictions of soil water contents than the simulations with laboratory-measured soil hydraulic properties. The authors attributed this to better representation of field water retention with multiple PTFs, compared to laboratory measurements. Our study corroborates that earlier finding. One possible reason for a generally better performance of the ensemble approach was using a large number of different PTFs. These PTFs were developed for different soils, based on different datasets, using different regression equations and soil properties as the predictors. As a result, many individual PTFs systematically overestimated or underestimated the calibrated G value. This produced wide tolerance intervals for the cumulative runoff simulated with highly variable parameters (Fig. 2b and c). However, the averaging results of all PTF estimates in the ensemble approach produced more robust and accurate estimates of the parameter G and, consequently, more adequate and certain runoff predictions. We believe that while the specific numeric results will differ for different sets of PTFs, the ensemble approach will perform better overall than other methods, with performance improving as the number of the PTFs used increases. Although this topic is beyond the scope of the present research, it opens an interesting avenue for further studies.

3.5. Accuracy and uncertainty in predicting FC concentrations and total FC amounts transported by runoff

Drastic differences in runoff simulation results did not manifest themselves in the FC concentrations computed with the STWIR

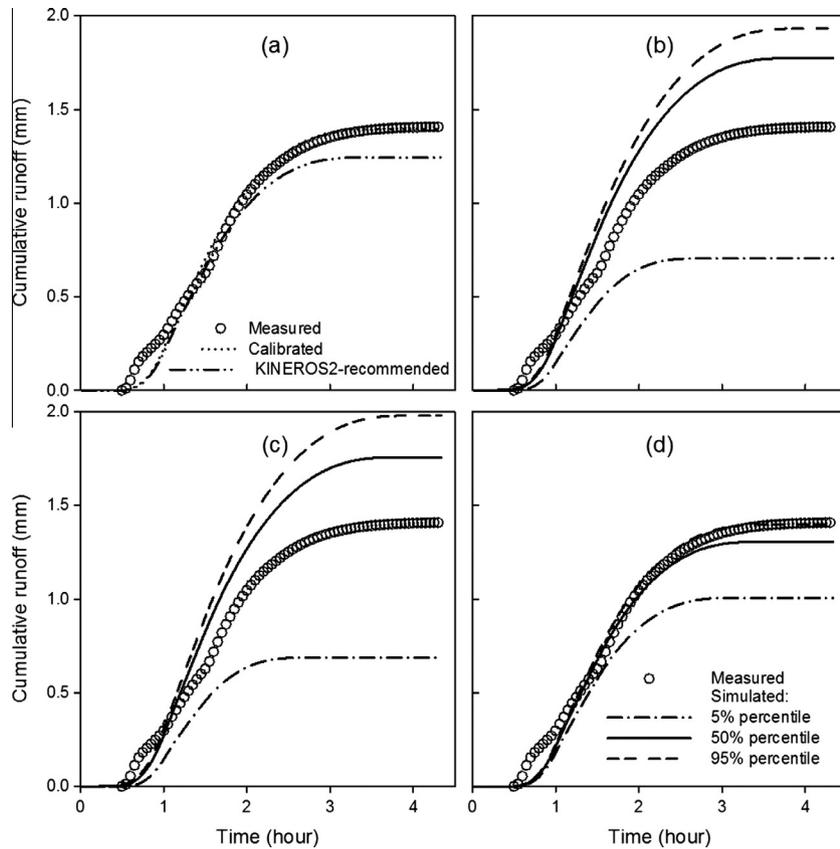


Fig. 2. Results of the KINEROS2 simulations with calibrated and KINEROS2-recommended values of parameter G (a), with PTF-estimated $G1$ -values (b), with PTF-estimated $G2$ -values (c), and with ensemble-averaged $G3$ -values (d).

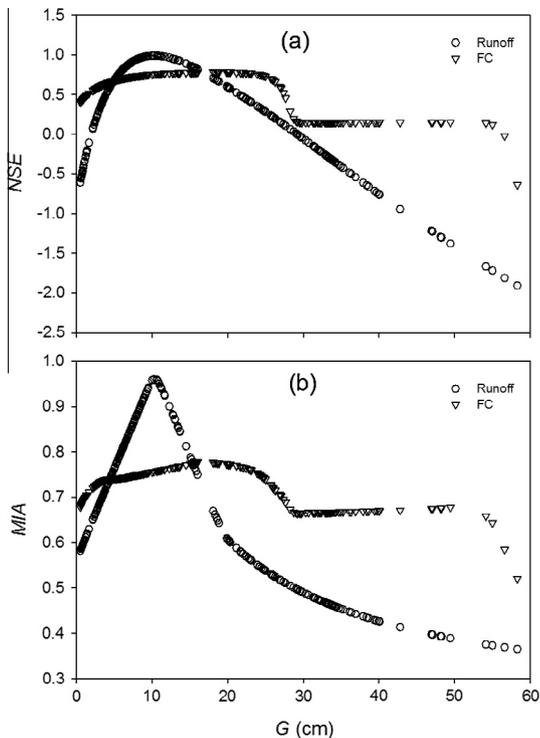


Fig. 3. Values of the Nash–Sutcliffe efficiency index (NSE) and the modified index of agreement (MIA) of cumulative runoff and FC concentrations in runoff water obtained in KINEROS2/STWIR simulations with G -values estimated from measured soil properties using 14 PTFs.

model. The accuracy of the FC concentration predictions was similar for both calibrated and KINEROS2-recommended parameter values (Table 3, Fig. 4a). The 90% tolerance intervals, computed as simulated values corresponding to 5% and 95% probability percentiles, were substantially smaller for the FC concentrations (Fig. 4b–d) than for the cumulative runoff (Fig. 2b–d) in the individual PTF and ensemble approaches. According to the NSE index, the FC predictions were accurate in all simulations, while the MIA values indicated that 43% and 100% of the FC predictions were accurate in the individual PTF and ensemble approaches, respectively (Tables 3 and 5).

In our study, a good overall performance of the STWIR model in predicting the FC concentrations stemmed from lower sensitivity of bacteria concentrations simulated using Eqs. (5–7) to the parameter G , compared to the sensitivity of runoff hydrograph computed using Eqs. (1–4) in the KINEROS2 model. The NSE values for FC were positive when G values were 0.5 cm to 55.0 cm and MIA values were greater than 0.75 when G values were 7.2 cm to 24.2 cm (Fig. 3). Note that the estimated G values ranged from 2.2 to 33.8 cm, 1.2 to 35.0 cm, and 7.8 to 20.1 cm for averaging 1, averaging 2, and the ensemble-averaging methods, respectively. Interestingly, both NSE and MIA indices computed for the FC concentrations did not change when G values were 28.7 cm to 54.6 cm, and monotonously decreased for the simulated runoff (Fig. 3). This can be explained by sampling frequency and different influences of the parameter G on the runoff and the FC concentration. It follows from Eq. (4) that infiltration rate I increases nonlinearly with increasing G value. The increase in I leads to a proportional decrease in runoff rate and, more importantly, to a decrease in runoff duration (Fig. 5a). The runoff rates were not high enough to noticeably affect the FC concentration in our study, but the decrease in the runoff

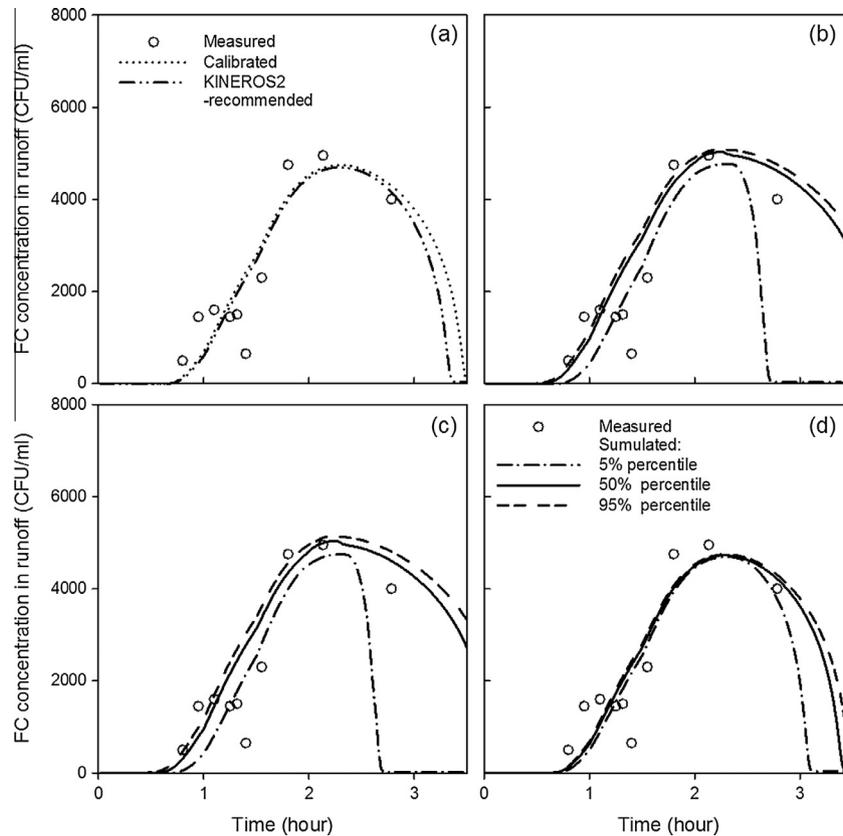


Fig. 4. FC concentrations in runoff obtained in the STWIR simulations with the calibrated and KINEROS2-recommended parameter G (a); with the PTF-averaged $G1$ (b) and averaged $G2$ (c) parameters; and using the ensemble-average values of parameter $G3$ (d).

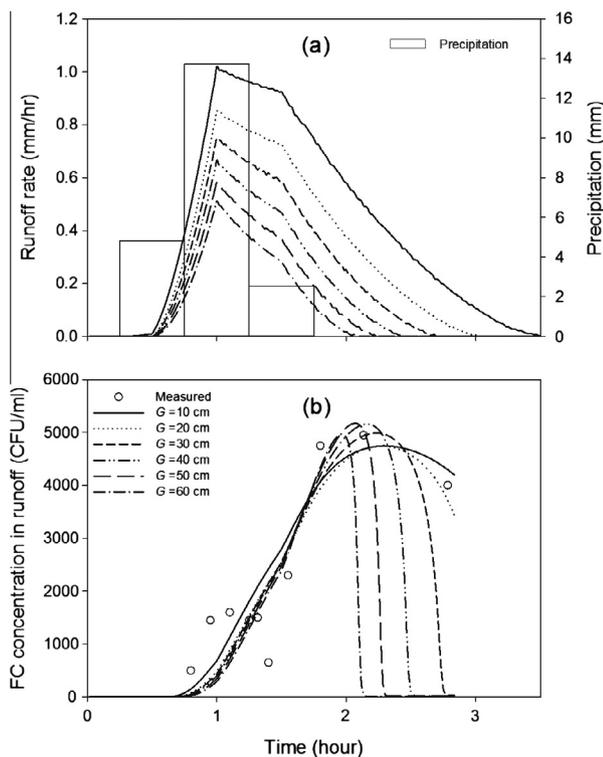


Fig. 5. Effect of the parameter G value on simulated runoff (a) and on FC concentrations in runoff water (b).

duration drastically changed the falling limb of the FC breakthrough curve (Fig. 5b). The peak FC concentrations were approximately the same for all G values, but time to the peak and to zero FC concentrations decreased with the increasing G value. For G values between 30 cm and 50 cm, time to zero concentration in the falling limb of the FC breakthrough curve fell within the time interval between the last two measurements; therefore, zero concentrations did not affect NSE and MIA within this G range (Fig. 2). Relationships between MIA, NSE and G were monotonous since the runoff hydrograph was recorded at 1 min interval in our study. This demonstrated the importance of FC sampling frequency, specifically the tailing part of the bacteria breakthrough curve (BTC) to correctly evaluate model performance with regard to model parameters.

Despite relatively low sensitivity of the STWIR model to the parameter G , the estimation method influenced the model predictive uncertainty. The 90% tolerance interval was considerably narrower, especially in the tailing part of the BTCs simulated with the ensemble-averages parameters, than those estimated individually for each PTF (Fig. 4b–d). This indicates a better overall performance of the former approach in estimating the net capillary drive parameter. The advantage of the ensemble-average prediction was more obvious when the total FC amounts transported with runoff from the KINEROS/STWIR simulations were compared. It yielded far more robust predictions of the total FC, compared to the individual PTF method. The last approach produced the total FC values that varied almost one order of magnitude. Such high variability was caused by overall uncertain predictions of the cumulative runoff and, to a lesser extent, the FC concentrations.

The total number of the runoff-transported FC cells predicted using different methods of the parameter G estimation appeared to be very close to the calibrated model when the mean simulated values were compared (Fig. 6). The relative deviations from the

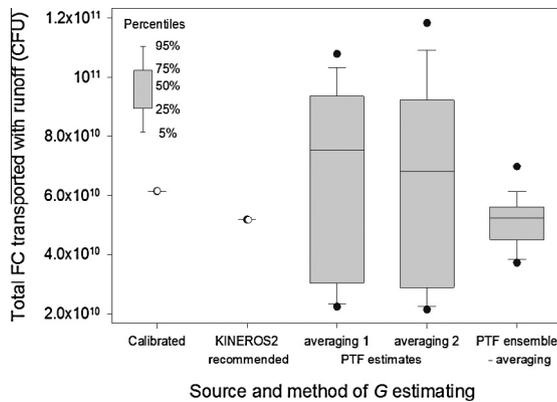


Fig. 6. Total FC transported with runoff obtained in KINEROS2/STWIR simulations with different sources of the parameter G . Close symbols show max and min values, while open symbols show single predicted values.

calibrated prediction were 16%, 23%, 11%, and 15% for the KINEROS2-recommended, averaging 1, averaging 2, and the ensemble-averaged parameters, respectively. This manifested an overall validity of the different averaging approaches for parameter estimation when spatially-variable soil properties and parameters must be represented by a single “effective” parameter value in overland contaminant transport models. The approaches can be particularly useful when soil properties vary beyond one soil texture class and the use of the KINEROS2 estimates becomes nontrivial. All averaging approaches implemented in this study utilize continuous PTFs and additional data such as bulk density, organic carbon content and sampling depth. The additional predictors make parameter estimates within one texture class more accurate.

4. Conclusions

We examined the effect of different approaches for estimating the net capillary drive parameter in the Parlange infiltration equation (Parlange et al., 1982) on the KINEROS2/STWIR predictions of runoff, FC concentration, and total FC transported with runoff. Three parameter estimation methods were compared: (i) fit to the measured cumulative runoff hydrograph (calibration); (ii) estimating based on soil texture class; and (iii) PTF-based estimating from basic soil properties. Overall model accuracy and uncertainty were found to be affected by the accuracy and uncertainty of the parameter G estimation. The most accurate prediction was obtained when the parameter estimates were obtained from model calibration by fitting to the cumulative runoff observations. The KINEROS2-recommended parameter estimation, based on soil texture class, yielded a slightly less accurate prediction of runoff, FC concentrations and total FC.

The spatially-variable soil properties used for the parameter estimation produced highly variable estimates. In general, variability was larger for the individual PTFs than the ensemble-averaging approach. That was explained by higher parameter uncertainty associated with the PTF models than with the spatial variability of soil properties in this study. The smaller uncertainty in G estimations produced by the ensemble method translated into more certain KINEROS2/STWIR predictions of cumulative runoff and total FC in runoff water, however, model uncertainty was output-specific. Uncertainty in predicting FC concentration was much smaller than runoff and total FC predictions, which implies that PTF can provide robust prediction results when FC concentration in runoff water is a concern.

The accuracy of averaged predictions of FC concentration and total FC transported with runoff were approximately the same for all three approaches to estimating the parameter G . This stems

from systematic over- and under-estimations of parameter G values by selected PTFs and is consistent with earlier findings for the ensemble prediction. Results of this study also showed the importance of sampling frequency, specifically in the tailing part of the bacteria breakthrough curve to correctly evaluate model performance with regard to model parameters.

Our study is the first to demonstrate the applicability of parameter ensemble and averaging for the net capillary drive parameter G that is then used in a model to simulate runoff and bacteria overland transport. The overall outcome shows the applicability of individual PTF-based and ensemble-averaging approach for infiltration parameter estimation when a single “effective value” is used as an input in the KINEROS2/STWIR model. We do not intend to generalize beyond the model used in the present study, however, applications to other models will be an interesting avenue for future research.

Acknowledgement

The United States Environmental Protection Agency through its Office of Research and Development collaborated in the research described here. It has been subjected to Agency review and approved for publication.

References

- Baker, L., Ellison, D., 2008. Optimisation of pedotransfer functions using an artificial neural network ensemble method. *Geoderma* 144, 212–224.
- Bicknell, B.R., Imhoff, J.C., Kittle, J.L. Jr., Donigan, A.S. Jr., Johanson, R.C., 1997. Hydrological simulation program—FORTRAN user's manual for version 11. Environmental Protection Agency Report No. EPA/600/R-97/080. U.S. Environmental Protection Agency, Athens, Ga.
- Børgesen, C.D., Iversen, B.V., Jacobsen, O.H., Schaap, M.G., 2008. Pedotransfer functions estimating soil hydraulic properties using different soil parameters. *Hydrol. Process.* 22, 1630–1639.
- Bottcher, A.B., Hiscock, J.G., 2001. WAMView—a GIS/land primary source of bacteria and pathogens on agricultural source based approach to watershed assessment modelling, in: Proc. of the TMDL Science Issues Conf., St. Louis, MO, USA, March 2001. Water Environ. Fed., Frederick, MD, USA, pp. 112–123.
- Bradford, S.A., Schijven, J., 2002. Release of cryptosporidium and giardia from dairy calf manure: impact of solution salinity. *Environ. Sci. Technol.* 36, 3916–3923.
- Breiman, L., 1996. Bagging predictors. *Mach. Learn.* 24, 123–140.
- Brooks, R.H., Corey, A.J., 1964. Hydraulic properties of porous media. *Hydrol. Paper* 3. Colorado State Univ., Fort Collins.
- Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, second ed. Springer-Verlag, New York.
- Burnham, K.P., Anderson, D.R., 2004. Understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304.
- Campbell, G.S., Shiozawa, S., 1992. Prediction of hydraulic properties of soils using particle size distribution and bulk density data. In: van Genuchten et al. (Eds.), Proc. Int. Worksh. on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils, Riverside, CA, 11–13 October 1989. U.S. Salinity Lab., Riverside, CA, pp. 317–328.
- Doherty, J., 2005. PEST: Software for Model-Independent Parameter Estimation. Watermark Numerical Computing, Australia.
- Fox, G.A., Munoz-Carpena, R., Sabbagh, G.J., 2010. Influence of flow concentration on parameter importance and prediction uncertainty of pesticide trapping by vegetative filter strips. *J. Hydrol.* 384, 164–173.
- Fox, G.A., Matlock, E.M., Guzman, J.A., Sahoo, D., Stunkel, K.B., 2011. *Escherichia coli* load reduction from runoff by vegetative filter strips: a laboratory scale study. *J. Environ. Qual.* 40, 980–988.
- Fraser, R.H., 1999. SEDMOD: A GIS-Based Delivery Model for Diffuse Source Pollutants. Ph.D. dissertation. Yale University, New Haven, CT, USA.
- Gee, G.W., Bauder, J.W., 1986. Particle analysis. In: Klute, A. (Ed.), Methods of soil analysis. Part 1, second ed., Agron. Monogr. 9. ASA and SSSA, Madison, WI, pp. 383–412.
- Gomi, T., Sidle, R.C., Miyata, S., Onda, Y., 2008a. Dynamic runoff connectivity of overland flow on steep forested hillslopes: scaling effects of Hortonian overland flow transfer. *Water Resour. Res.* 44, W08411, doi:10.29/2007WR005894.
- Gomi, T., Sidle, R.C., Ueno, M., Miyata, S., Kosugi, K., 2008b. Characteristics of overland flow generation on steep forested hillslopes of central Japan. *J. Hydrol.* 361 (3/4), 275–290.
- Goodrich, D.C., Burns, I.S., Unkrich, C.L., Semmens, D.J., Guertin, D.P., Hernandez, M., Yatheendradas, S., Kennedy, J.R., Levick, L.R., 2012. KINEROS2/AGWA: model use, calibration, and validation. *Trans. ASABE* 55, 1561–1574.
- Green, W.H., Ampt, G., 1911. Studies of soil physics, Part I – The flow of air and water through soils. *J. Agr. Sci.* 4, 1–24.

- Guber, A.K., Pachepsky, Y.A., 2010. Multimodeling with pedotransfer functions. Documentation and User Manual for PTF Calculator (CalcPTF), Version 2.0. Environmental Microbial and Food Safety Laboratory, Beltsville Agricultural Research Center, USDA-ARS, Beltsville, MD <<https://www.msu.edu/~akguber/software.html>>.
- Guber, A.K., Pachepsky, Y.A., van Genuchten, M.Th., Rawls, W.J., Jacques, D., Šimunek, J., Cady, R.E., Nicholson, T.J., 2006. Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. *Vadose Zone J.* 5, 234–247.
- Guber, A.K., Yakirevich, A.M., Sadeghi, A.M., Pachepsky, Y.A., Shelton, D.R., 2009. Uncertainty evaluation of coliform bacteria removal from vegetated filter strip under overland flow condition. *J. Environ. Qual.* 38, 1636–1644.
- Guber, A.K., Pachepsky, Y.A., Yakirevich, A.M., Shelton, D.R., Sadeghi, A.M., Goodrich, D.C., Unkrich, C.L., 2010. STWIR, A Microorganism Transport with Infiltration and Runoff add-on Module for the KINEROS2 Runoff and Erosion Model: Documentation and User Manual, USDA-ARS, Beltsville <<https://www.msu.edu/~akguber/software.html>>.
- Guber, A.K., Pachepsky, Y.A., Yakirevich, A.M., Shelton, D.R., Sadeghi, A.M., Goodrich, D., Unkrich, K., 2011. Uncertainty in modeling of Faecal coliform overland transport associated with manure application in Maryland. *Hydrol. Process.* 25, 2393–2404.
- Gupta, S.C., Larson, W.E., 1979. Estimating soil water retention characteristics from particle-size distribution, organic matter percent, and bulk density. *Water Resour. Res.* 15, 1633–1635.
- Haefner, F., Boy, S., Wagner, S., Behr, A., Piskarev, V., Palatnik, V., 1997. The 'front limitation' algorithm a new and fast finite-difference method for groundwater pollution problems. *J. Contam. Hydrol.* 27, 43–61.
- Kim, K., Whelan, G., Purucker, S.T., Bohrmann, T.F., Cytterski, M., Molina, M., Gu, Y., Pachepsky, Y., Guber, A., Franklin, D., 2013. Rainfall-runoff model parameter estimation and uncertainty evaluation on small plots. *Hydrol. Process.* <http://dx.doi.org/10.1002/hyp.10001>.
- Köhne, J.M., Mohanty, B.P., Šimunek, J., 2005. Inverse dual-permeability modeling of preferential water flow in a soil column and implications for field-scale solute transport. *Vadose Zone J.* 5, 59–76.
- Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35, 233–241.
- Lilly, A., Nemes, A., Rawls, W.J., Pachepsky, Y.A., 2007. Probabilistic approach to the identification of input variables to estimate hydraulic conductivity. *Soil Sci. Soc. Am. J.* 72, 16–24.
- Mayr, T., Jarvis, N.J., 1999. Pedotransfer functions to estimate soil water retention parameters for a modified Brooks–Corey type model. *Geoderma* 91, 1–9.
- Medema, G.J., Schijven, J.F., 2001. Modelling the sewage discharge and dispersion of *Cryptosporidium* and *Giardia* in surface water. *Water Res.* 35, 4307–4316.
- Minasny, B., Hoppmans, J.W., Harter, T., Eching, S.O., Tuli, A., Denton, M.A., 2004. Neural networks prediction of soil hydraulic functions for alluvial soils using multistep outflow data. *Soil Sci. Soc. Am. J.* 68, 417–429.
- Misra, A.K., Baker, J.L., Mickelson, S.K., Shang, H., 1996. Contributing area and concentration effects of herbicide removal by vegetative filter strips. *Trans. ASAE* 39, 2105–2111.
- Morel-Seytoux, H.J., Meyer, P.D., Nachabe, M., Touma, J., van Genuchten, M.T., Lenhard, R.J., 1996. Parameter equivalence for Brooks–Corey and van Genuchten soil characteristics: preserving the effective capillary drive. *Water Resour. Res.* 32, 1251–1258.
- Munoz-Carpena, R., Fox, G.A., Sabbagh, G.J., 2010. Parameter importance and uncertainty in predicting runoff pesticide reduction with filter strips. *J. Environ. Qual.* 39, 630–641.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Srinivasan, R., Williams, J.R., 2005. Soil and Water Assessment Tool Theoretical Documentation, Version 2005. Grassland, Soil and Water Research Laboratory, ARS-USDA, Temple, TX.
- Nelson, E.W., Sommers, L.E., 1996. Total carbon, organic carbon, and organic matter. In: Sparks, D.L. (Ed.), *Methods of Soil Analysis: Chemical Methods*. Part 3. Soil Sci. Soc. of Am., Madison WI.
- Nemes, A., Timlin, D.J., Pachepsky, Y.A., Rawls, W.J., 2009. Evaluation of the Rawls et al. (1982). Pedotransfer functions for their applicability at the US National Scale. *Soil Sci. Soc. Am. J.* 73, 1638–1645.
- Oosterveld, M., Chang, C., 1980. Empirical relations between laboratory determinations of soil texture and moisture retention. *Can. Agric. Eng.* 21, 149–151.
- Parlange, J.-Y., Lisle, I., Braddock, R.D., Smith, R.E., 1982. The three parameter infiltration equation. *Soil Sci.* 133, 337–341.
- Perrone, M.P., Cooper, L.N., 1993. When networks disagree: ensemble method for neural networks. In: Mammone, R.J. (Ed.), *Neural Networks for Speech and Image Processing*. Chapman-Hall, London, pp. 126–142.
- Philips, J.R., 1957. The theory of infiltration: the infiltration equation and its solution. *Soil Sci.* 83, 345–375.
- Poletika, N.N., Coody, P.N., Fox, G.A., Sabbagh, G.J., Dolder, S.C., White, J., 2009. Chlorpyrifos and atrazine removal from runoff by vegetated filter strips: experiments and predictive modeling. *J. Environ. Qual.* 38, 1042–1052.
- Rajkai, K., Várallyay, Gy. 1992. Estimating soil water retention from simpler properties by regression techniques. In: van Genuchten et al. (Eds.), *Proc. Int. Worksh. on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, Riverside, CA, 11–13 October 1989, U.S. Salinity Lab., Riverside, CA, pp. 417–426.
- Rawls, W.J., Brakensiek, D.L., 1985. Prediction of soil water properties for hydrologic modeling. In: Jones, E.B., Ward, T.J. (Eds.), *Proc. Symp. Watershed Management in the Eighties*, Denver, CO, 30 April–1 May 1985. Am. Soc. Civil Eng., New York, pp. 293–299.
- Rawls, W.J., Brakensiek, D.L., Saxton, K.E., 1982. Estimation of soil water properties. *Trans. ASAE* 25, 1316–1320.
- Rawls, W.J., Brakensiek, D.L., Soni, B., 1983. Agricultural management effects on soil water processes. Part I. Soil water retention and Green–Ampt parameters. *Trans. ASAE* 26, 1747–1752.
- Rawls, W.J., Ahuja, L.R., Brakensiek, D.L., Shirmohammadi, A., 1992. Infiltration and soil water movement. In: Maidment, D.R. (Ed.), *Handbook of Hydrology*. McGraw-Hill Inc., New York, NY.
- Rawls, W.J., Gimenez, D., Grossman, R., 1998. Use of soil texture, bulk density, and slope of the water retention curve to predict saturated hydraulic conductivity. *Trans. ASAE* 41, 983–988.
- Rawls, W.J., Gimenez, D., Grossman, R., 1998. Use of soil texture, bulk density and the slope of the water retention curve to predict saturated hydraulic conductivity. *Trans. ASAE* 4, 983–988.
- Roodsari, R.M., Shelton, D.R., Shirmohammadi, A., Pachepsky, Y.A., Sadeghi, A.M., Starr, J.L., 2005. Fecal coliform transport as affected by surface condition. *Trans. ASAE* 48, 1055–1061.
- Sabbagh, G.J., Fox, G.A., Kamanzi, A., Roepke, B., Tang, J.-Z., 2009. Effectiveness of vegetative filter strips in reducing pesticide loading: quantifying pesticide trapping efficiency. *J. Environ. Qual.* 38, 762–771.
- Sadeghi, A., Arnold, J.G., 2002. A SWAT/microbial submodel for predicting pathogen loadings in surface and groundwater at watershed and basin scales. In: *Total Maximum Daily Load (TMDL) Environmental Regulations*, Proceedings of the Conference, Fort Worth, TX, USA, March 11–13, pp. 56–63.
- Saxton, K.E., Rawls, W.J., Romberger, J.S., Papendick, R.I., 1986. Estimating generalized soil–water characteristics from texture. *Soil Sci. Soc. Am. J.* 50, 1031–1036.
- Schaap, M.G., Leij, F.J., 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. *Soil Tillage Res.* 47, 37–42.
- Schaap, M.G., Leij, F.J., van Genuchten, M.Th., 1998. Neural network analysis for hierarchical prediction of soil hydraulic properties. *Soil Sci. Soc. Am. J.* 62, 847–855.
- Semmens, D.J., Goodrich, D.C., Unkrich, C.L., Smith, R.E., Woolhiser, D.A., Miller, S.N., 2008. KINEROS2 and the AGWA modeling framework. In: *Wheat, H., Sorooshian, S., Sharma, K.D. (Eds.), Hydrological Modelling in Arid and Semi-Arid Areas*. Cambridge University Press, London, pp. 49–69, Chapter 5.
- Shirmohammadi, A., Chaubey, I., Harmel, R.D., Bosh, D.D., Muñoz-Carpena, R., Dharmasri, C., Sexton, A., Arabi, M., Wolfe, M.L., Frankenberger, J., Graff, C., Sohrabi, T.M., 2006. Uncertainty in TMDL models. *Trans. ASAE* 49, 1033–1049.
- Sidle, R.C., Hirano, T., Gomi, T., Terajima, T., 2007. Hortonian overland flow from Japanese forest plantations – an aberration, the 'real thing', or something in between? *Hydrol. Process.* 21, 3237–3247.
- Sullivan, T.J., Moore, J.A., Thomas, D.R., Mallery, E., Snyder, K.U., Wustenberg, M., Wustenberg, J., Mackey, S.D., Moore, D.L., 2007. Efficacy of vegetated buffers in preventing transport of fecal coliform bacteria from pasturelands. *Environ. Manage.* 40, 958–965.
- Tamari, S., Wösten, J.H.M., Ruiz-Suárez, J.C., 1996. Testing an artificial neural network for predicting soil hydraulic conductivity. *Soil Sci. Soc. Am. J.* 60, 1732–1741.
- Tetra Tech, Inc., US Environmental Protection Agency (USEPA), 2002. *The Loading Simulation Program in C++ (LSPC) Watershed Modeling System—Users' Manual*, Pasadena, CA, USA.
- Tomasella, J., Hodnett, M.G., 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Sci.* 163, 190–202.
- Twarakavi, N.K.C., Šimunek, J., Schaap, M.G., 2009. Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Sci. Soc. Am. J.* 73, 1443–1452.
- US SCS, 1972: *National Engineering Handbook*, Section 4, Hydrology, U.S. Department of Agriculture, Washington, DC.
- van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892–898.
- Vereecken, H., Maes, J., Feyen, J., 1990. Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Sci.* 149, 1–12.
- Wagner, B., Tarnawski, V.R., Hennings, V., Müller, U., Wessolek, G., Plagge, R., 2001. Evaluation of pedotransfer functions for unsaturated soil hydraulic conductivity using an independent data set. *Geoderma* 102, 275–297.
- Walker, S.E., Mostaghimi, S., Dillaha, T.A., Woeste, F.E., 1990. Modeling animal waste management practices: impacts on bacteria levels in runoff from agricultural lands. *Trans. ASAE* 33, 807–817.
- Williams, J., Ross, P., Bristow, K., 1992. Prediction of the Campbell water retention function from texture, structure, and organic matter. In: van Genuchten et al. (Eds.), *Proc. Int. Worksh. on Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, Riverside, CA, 11–13 October 1989, U.S. Salinity Lab., Riverside, CA, pp. 427–442.
- Woolhiser, D.A., Smith, R.E., Goodrich, D.C., 1990. KINEROS, A Kinematic Runoff and Erosion Model: Documentation and User Manual. ARS-77. USDA, ARS, Washington, DC.
- Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90, 169–185.