European Journal of Soil Science

European Journal of Soil Science, March 2016, **67**, 135–147 doi: 10.1111/ejss.12322

Estimation of saturated hydraulic conductivity from double-ring infiltrometer measurements

M. Fatehnia[a], K. Tawfiq[a] & M. Ye[b]

[a]*Department of Civil and Environmental Engineering, FAMU-FSU College of Engineering, 2525 Pottsdamer Street, Tallahassee, FL 32310, USA, and* [b]*Department of Scientific Computing, Florida State University, 400 Dirac Science Library, Tallahassee, FL 32306, USA*

Summary

This research aims to determine soil vertical saturated hydraulic conductivity (K_s) *in situ* from the measured steady infiltration rate (I), initial soil properties and double-ring infiltrometer (DRI) test data. Characterizing the effects of these variables on the measured steady infiltration rate will enable more accurate prediction of K_s. We measured the effects of the ring diameter, head of ponding, ring depth, initial effective saturation and soil macroscopic capillary length on measured steady infiltration rates. We did this by simulating 864 DRI tests with the finite element program HYDRUS-2D and by conducting 39 full-scale *in situ* DRI tests, 30 Mini-Disk infiltrometer experiments and four Guelph Permeameter tests. The M5′ model trees and genetic programming (GP) methods were applied to the data to establish formulae to predict the K_s of sandy to sandy-clay soils. The nine field DRI tests were used to verify the computer models. We determined the accuracy of the methods with 30% of the simulated DRI data to compare I/K_S values of the finite element models with estimates from the suggested formulae. We also used the suggested formulae to predict the K_s values of 30 field DRI experiments and compared them with values measured by Guelph Permeameter tests. Compared with the GP method, the M5′ model was better at predicting K_S, with a correlation coefficient of 0.862 and root mean square error (RMSE) of 0.282 cm s^{-1}. In addition, the latter method estimated K_s values of the field experiments more accurately, with an RMSE of 0.00346 cm s^{-1}.

Highlights

- Effects of initial soil properties and DRI test data on steady infiltration rate are studied.
- M5′ model algorithm and genetic programming method are applied on the data to formulate the equations.
- Compared to the GP method, the M5′ model was better at predicting K_s.
- The equation from M5′ tree model is robust for estimating *in situ* K_s with the DRI.

Introduction

In spite of the similarity in units of infiltration rate (I) and saturated hydraulic conductivity (K_S) (cm s^{-1}) of soil, there is a distinct difference between these two measurements. In soil science, infiltration rate is defined as the rate at which soil is able to absorb rainfall or irrigation water and is governed by gravity and capillary action. The K_s is the steady-state 1-D infiltration velocity when water is supplied at the soil surface under zero ponded conditions and the hydraulic gradient is equal to one (Fatehnia *et al.*, 2014a). Infiltration rate and K_s can be related directly only when the hydraulic boundary conditions, such as hydraulic gradient and the extent of lateral flow of water, are known (ASTM, 2009). Accurate *in situ* measurement of K_S is essential for various agricultural, environmental and geotechnical applications, such as the design of drainage systems and measurement of seepage from reservoirs, detention ponds or wastewater lagoons (Fatehnia, 2015).

Single or double ring infiltrometers are commonly used for *in situ* measurement of infiltration rate; the double ring infiltrometer is the preferred choice. The use of the outer ring helps to reduce the error that may result from lateral flow in soil (Bouwer, 1986). The DRI test as described by the American Society for Testing and Materials, (ASTM D3385) consists of open inner and outer cylinders that should be inserted manually into the ground and

Correspondence: M. Fatehnia. E-mail: mfatehnia@fsu.edu

Received 17 November 2014; revised version accepted 17 December 2015

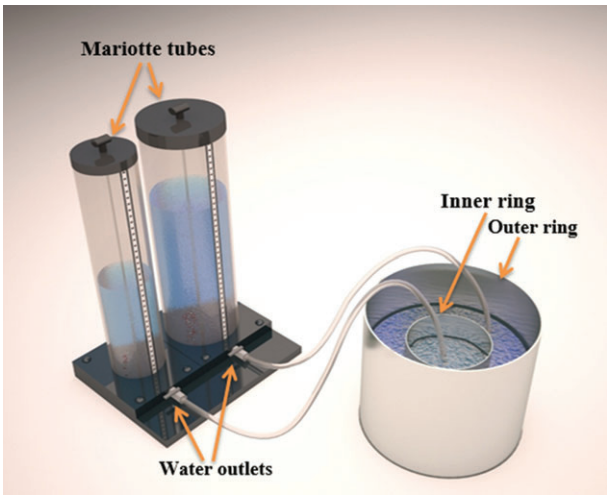

Figure 1 Schematic sketch of the double-ring infiltrometer.

be partially filled with a constant head of water (ASTM, 2009) (Figure 1). The infiltration rate is determined by measuring the volume of liquid added to the inner ring to maintain a constant level. Because the DRI test process is time-consuming, several attempts have been made to automate it (Fatehnia *et al.*, 2016).

The dependence of K_S measurements on the diameter of the cylinders used has been studied by several researchers and several dimensions have been proposed for the minimum diameter required for test cylinders. Swartzendruber & Olsen (1961) reported that radii of 60 and 50 cm for the outer and inner rings, respectively, were the most satisfactory for various conditions in a sandy soil. Ahuja *et al.* (1976) reported that when an outer ring of 90-cm diameter was used with an inner ring of 30-cm diameter, the lateral flow was almost eliminated. Further, he reported that even when an outer ring of 60-cm diameter was used, its effect on the final infiltration rate was negligible. Bouwer (1986) suggested that a diameter of at least 100 cm should be used for accurate K_S measurements. Youngs (1987) concluded that results were consistent from site to site when the ring size was at least 15 cm. Gregory *et al.* (2005) concluded that for a constant head test in sandy soil generally found in north and central Florida, a double-ring infiltrometer with 15-cm inner and 30-cm outer diameters would be suitable. After 28 DRI tests and a series of numerical experiments with the two-dimensional model HYDRUS-2D (Šimůnek *et al.*, 1999), Lai & Ren (2007) concluded that an inner ring diameter of at least 80 cm is needed to minimize the effects of lateral divergence because of capillary gradients. Lai *et al.* (2010) conducted 7224 numerical simulations from which they concluded that inner ring diameters greater than 80 cm are needed to obtain reliable *in situ* measurements of K_S.

These studies aimed to provide more accurate estimates of K_S with DRI by evaluating the effect of ring size on the measured infiltration rate. Their main focus was to determine the minimum required ring diameter for a specific soil type. Disregard for the effect of other factors, such as soil properties or the set-up to measure infiltration rate, was the main reason for inconsistencies in the ring diameters recommended in different studies.

Other researchers evaluated the effect of soil and test variables on steady infiltration rate, rather than ring diameter. Reynolds *et al.* (2002) suggested a steady-state solution originally developed for single-ring infiltrometers (Reynolds & Elrick, 1990). Their proposed solution accounts explicitly for the effects of ring diameter, depth of ring insertion, head of ponded water and a soil structure dependent parameter (alpha*) that expresses the relative importance of gravitational to capillary forces during infiltration. Gregory *et al.* (2005) concluded that soil bulk density had an important effect on the final infiltration rate from the test. Chowdary *et al.* (2006) studied the effectiveness of DRI in reducing lateral flow. They used a dimensional analysis technique and developed relationships for cumulative infiltration based on the diameter of the infiltrometer, head of ponding, depth of penetration, elapsed time, K_S and initial moisture content. Lai & Ren (2007) realized during their experiments that a large initial water content weakens the capillary effect, which decreases the lateral flow contribution to total infiltration.

The objective of this current research was to provide an accurate measurement of K_S with DRI by studying the effects of the initial soil properties and set-up conditions on measured infiltration rate using the M5′ model algorithm and genetic programming. Initially, we did 864 computer simulations and nine *in situ* DRI experiments to validate the computer models. Then, the M5′ model algorithm and genetic programming were applied to the data to develop formulae to estimate K_S in sandy to sandy-clay soil. Six independent variables of measured steady infiltration rate, initial water content, inner ring diameter, soil macroscopic capillary length, head of ponding and depth of ring insertion were used to estimate K_S. Accuracy of the proposed formulae was determined by calculating the correlation coefficient and root mean square error (RMSE). Finally, the proposed formulae were applied to 30 DRI experiments and the results were compared with the K_S values measured by the Guelph permeameter experiment.

Materials and methods

Properties that affect infiltration rate with the DRI

We investigated the effects of the following five properties on measurements of the rate of steady-state infiltration in the DRI experiment.

Macroscopic capillary length (λ). For ponded infiltration, the macroscopic capillary length, λ_c (cm), is represented by White & Sully (1987) as:

$$\lambda_c = \phi_m / \left(K_{\psi_0} - K_{\psi_i}\right), \quad (1)$$

where K_ψ (cm s^{-1}) is the hydraulic conductivity at matric potential ψ (cm), ψ_0 is a matric potential at or near saturation and ψ_i is a more negative pore-water pressure head that corresponds to water content θ_i (cm^3 cm^{-3}) such that $\psi_0 > \psi_i$. In this form, λ_c (cm) may be interpreted as the mean height of a capillary rise above the water table (Raats & Gardner, 1971) and ϕ_m (cm^2 s^{-1}) was defined by

Gardner (1958) as:

$$\phi_m = \int_{\psi_i}^{0} K_\psi d\psi \qquad -\infty \le \psi_i \le 0. \tag{2}$$

Macroscopic capillary length expresses the relative importance of gravitational to capillary forces during infiltration. This variable changes with soil type and can be effective for quantifying the unsaturated hydraulic conductivity function; it can be measured by a mini-disk infiltrometer. This is a manually operated portable tension infiltrometer that can measure unsaturated hydraulic conductivity at tensions between −0.05 and −0.6 kPa. It consists of an upper chamber to control the suctions and a lower chamber to store infiltration water (Fatehnia *et al.*, 2014a).

Inner ring diameter (d_i). Different ring diameters can be used to assess their effect on the rate of steady infiltration. In practice, smaller rings are easier to use and handle. They also require less volume of water to run the test. However, they have less contact with the ground surface, which causes variation in the results for heterogeneous soil.

Head of ponding (H). The effect of head of ponding on infiltration rate with DRI has been studied previously (Reynolds *et al.*, 2002; Chowdary *et al.*, 2006). The head of water in the rings is related directly to the hydraulic gradient of the infiltration system and affects the infiltration rate.

Ring insertion depth (D). Reynolds *et al.* (2002) and Chowdary *et al.* (2006) selected ring insertion depth as one of the variables that affects DRI results. Guisheng *et al.* (2011) stated that accurate measurement of the vertical infiltration rate with DRI requires the insertion depth to be larger than the ASTM-specified value. They recommended depths of between 19 and 22 cm for acceptable accuracy of the measurements of rate of infiltration. However, it is difficult practically to insert rings to such depths in some types of soil. Therefore, it is important to study the effect of ring depth on infiltration rate.

Initial effective soil saturation (S). Lai & Ren (2007) found that large initial water content weakens the capillary effect and decreases the contribution of lateral flow to total infiltration. An increase in water content affects the infiltration rate by minimizing horizontal flow and also by reducing the time needed to reach the steady state. To compare soil moisture in different types of soil, water content measurements were standardized. Effective soil saturation (S) is a dimensionless variable defined by van Genuchten (1980) as:

$$S = \frac{\theta - \theta_r}{\theta_s - \theta_r}, \tag{3}$$

where θ is the volumetric water content, θ_r is the residual water content and θ_s is the saturated water content, which is equivalent to porosity.

Field experiments and laboratory measurements

Field experiments were required to verify and calibrate the computer models. They were also needed to evaluate robustness and computational efficiency of the final equations to estimate K_S. Nine field DRI tests were carried out at the FAMU-FSU College of Engineering campus in Tallahassee, Florida, to verify the computer models. Subsequently, after the final equations to estimate K_S had been determined, 30 DRI field experiments with various initial conditions were carried out at four test sites to assess the accuracy of the formulae with real field data. To specify the soil classification of the sites, granulometric analysis was carried out on soil samples taken from different locations of each site. Soil classification of the material was based on ASTM D 422-02 (ASTM, 2007) and United States Department of Agriculture (USDA) soil classification systems. However, all sites were classified as Ferric Acrisol according to Food and Agriculture Organization of the United Nations (FAO) soil classification systems (FAO, 1998). Results of the preliminary laboratory tests showed that the soil was similar at each of the test locations.

The first set of 30 DRI tests was also carried out at the FAMU-FSU College of Engineering Campus. The test site has a gentle slope of 4–6° and the soil is an Orangeburg fine sandy loam. The second test site was the Florida State University Reservation camping area with light brown well-sorted sand. The site has a gentle slope of 2–4°. The third site was in the vicinity of Tallahassee international airport. The surface soil included mottled dark and light brown medium to fine sand. The last set of tests was carried out in Greenway, Tallahassee. The surface material is mottled light brown and reddish brown loamy sand. The soil surface has a slope of 2–4°.

The three sizes of rings used for the 30 field experiments were 15-cm inner and 30-cm outer, 20-cm inner and 40-cm outer, and 30-cm inner and 60-cm outer ring diameters. Experiments were carried out at three different times. Initially the soil was relatively dry with an effective saturation of around 0.3, then the soil was saturated immediately after precipitation; the final measurements were made under semi-saturated conditions with an effective saturation of around 0.6. Before every infiltration test, initial gravimetric moisture content of the soil adjacent to each test site was measured by standard laboratory procedures (ASTM, 2010). The residual water content was also determined by oven-drying undisturbed cores of soil of a known volume to estimate bulk density. This information was used to measure the effective saturation. Steady-state infiltration was assumed to have been reached when discharge changes were < 0.5% over a 5-minute interval. After the DRI experiments, the macroscopic capillary length of the soil adjacent to each location was measured with a mini-disk infiltrometer. The procedure proposed by Zhang (1998) was used to measure macroscopic capillary length with two suction values of −0.05 and −0.4 kPa (Fatehnia *et al.*, 2014a) (Figure 2a). Finally, to measure K_S at each site, the Guelph permeameter method was used (Figure 2b).

(a) (b)

Figure 2 Mini-disk infiltrometer and Guelph permeameter test.

Computer simulations

Computer simulations were selected as an alternative to field tests because of the restrictions on doing a large number of field DRI experiments to obtain data. We could adjust the relevant variables in each simulation and observe their respective effects on measured infiltration rate. The HYDRUS-2D finite element code (Šimůnek *et al.*, 1999), which can simulate the two- and three-dimensional movement of water, heat and multiple solutes in variably saturated media, was used for simulations. This program allows the user to simulate a wide range of boundary conditions and irrigation regimes. HYDRUS has been applied previously for simulating infiltration under a double-ring infiltrometer (Lai & Ren, 2007; Lazarovitch *et al.*, 2007; Lai *et al.*, 2010; Fatehnia, 2015) and infiltration inside a ring infiltrometer (Léger *et al.*, 2013). The program uses the Richards equation (Richards, 1931) for simulating variably-saturated flow (Šimůnek *et al.*, 2006):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x_i}\left[K\left(K_{ij}^{\mathrm{A}}\frac{\partial \psi}{\partial x_j} + K_{iz}^{\mathrm{A}}\right)\right] - S, \tag{4}$$

where θ is the volumetric water content ($\mathrm{cm^3\ cm^{-3}}$), S is a sink term ($\mathrm{cm^3\ cm^{-3}\ s^{-1}}$), ψ is the soil water matric head (cm), x_i ($j = 1, 2$) are the spatial coordinates (cm), t is time (s), K_{ij}^{A} are components of a dimensionless anisotropy tensor K^{A} in the horizontal plane and K_{iz}^{A} those in the vertical plane, and K is the unsaturated hydraulic conductivity function ($\mathrm{cm\ s^{-1}}$) given by:

$$K(\psi, x, y, z) = K_{\mathrm{s}}(x, y, z)\, K_{\mathrm{r}}(\psi, x, y, z), \tag{5}$$

where K_{r} is the relative hydraulic conductivity and K_{s} is the saturated hydraulic conductivity ($\mathrm{cm\ s^{-1}}$).

The van Genuchten–Mualem soil hydraulic properties model (Mualem, 1976; van Genuchten, 1980) was selected for the numerical simulations of the 2-D-axisymmetric domain of the DRI test. The size of the transport domain was selected such that the outer boundaries did not affect the flow field within the domain. The transport domain was discretized into an unstructured triangular finite element mesh. The area immediately surrounding the infiltration rings had considerably smaller mesh sizes, whereas they became larger as distance from the rings increased (Figure 3).

Computer models were calibrated by simulating the initial conditions of nine field experiments and by comparing the infiltration rate of simulations with the values measured in the field. Similar to the field tests, infiltration rates of the models were determined by measuring the volumetric water flux across the upper borders along the inner ring. The average difference of 9.32% between the simulations and measurements was evidence of the repeatability of the models (Table 1).

Three values for head of ponding, three values for ring depth, four values for effective soil saturation, four values for macroscopic capillary length and six values for ring diameter were evaluated (Table 2). A total of 864 different conditions in DRI tests were modelled to investigate the effects of initial conditions on infiltration rate. For modelling of the different properties, we chose a range of values and intervals that covered those of the soil and of the DRI test in reality. The hydraulic variables of the different soil types used in the models were determined by the Rosetta Lite program

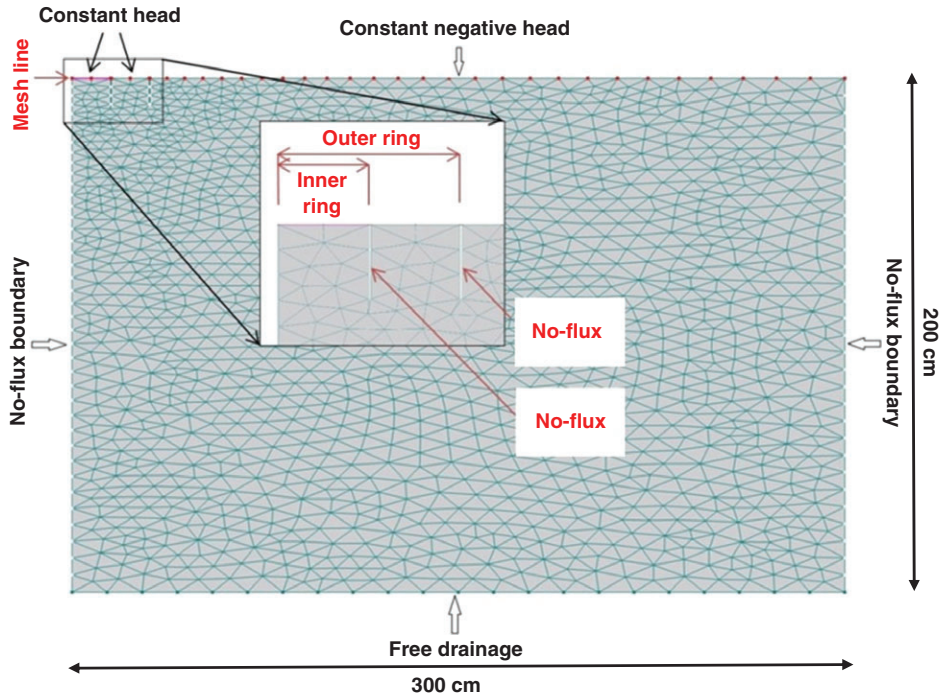

Figure 3 The 2-D-axisymmetric HYDRUS simulation domain and boundary conditions.

Table 1 Comparison between the field and simulated DRI experiments

Inner ring diameter / cm	Effective soil saturation	Soil macroscopic capillary length / cm	Measured infiltration rate in field / cm s^{-1}	Measured infiltration rate in simulations / cm s^{-1}
15	0.3	5.0	1.54E−02	1.34E−02
15	0.6	5.0	1.29E−02	1.33E−02
15	1.0	5.0	1.02E−02	1.24E−02
20	0.3	5.5	1.35E−02	1.27E−02
20	0.6	5.0	1.17E−02	1.26E−02
20	1.0	5.5	1.04E−02	1.17E−02
30	0.3	5.0	1.36E−02	1.22E−02
30	0.6	6.0	1.16E−02	1.22E−02
30	1.0	6.0	1.08E−02	1.13E−02

Water ponding heads and ring insertion depths were 10 and 15 cm, respectively.

(Schaap *et al.*, 2001), which is embedded in the HYDRUS software package (Šimůnek *et al.*, 2006), with particle-size distribution and bulk density of the topsoil as inputs. The values of K_S for each simulated test were predefined in the finite element models, whereas the steady infiltration rates were measured from the results of modelling. Ring depths and diameters were modelled graphically; values of macroscopic capillary length were assigned as $1/\alpha$ in the van Genuchten–Mualem soil-water retention function and head of ponding and effective saturation values were specified by assigning specific boundary conditions and initial moisture content values, respectively. All simulated models had the ratio of outer to inner ring diameters equal to 2.

Table 2 Values of double ring infiltrometer test variables used for 864 simulations

Variable	Modelled variable values
H: head of ponding in the rings / cm	5 10 15
D: depth of the ring in the soil / cm	5 10 15
S: effective soil saturation	0 0.6 0.8 1
λ: macroscopic capillary length / cm	7 13 17 33
d_i: inner ring diameter / cm	15 20 30 40 50 60

The four values of macroscopic capillary length represented different soil types. Based on the soil hydraulic variables of Schaap *et al.* (2001), macroscopic capillary lengths of 7, 13, 17 and 33 cm corresponded to sand, sandy loam, sandy clay loam and sandy clay soil, respectively. In this way, the proposed formula can be used for sandy to sandy clay soil types. Because soil is heterogeneous and spatially variable, K_S can change by more than an order of magnitude within a short distance. Homogeneity of the soil within the measured area is one of the common assumptions for determining K_S in DRI tests (Lai & Ren, 2007); therefore, the soil was assumed to be homogeneous for modelling.

Data analysis

Saturated hydraulic conductivities and their respective measured steady infiltration rates, together with the heads of water, ring depths, initial effective soil saturations, macroscopic capillary lengths and ring diameters of the 864 simulated DRI tests, were entered into the Weka program (Witten & Frank, 2005) and the

MATLAB genetic programming toolbox (Silva & Almeida, 2003) to predict the relation between K_S and the rest of the variables with the M5′ model tree and genetic programming methods, respectively.

The M5′ model tree

The M5′ model tree was the first method used to formulate an equation to estimate K_S in the DRI test. Model trees are machine-learning techniques that divide the property space into subspaces and formulate a multivariate linear regression model in each of them. Model trees, like regression trees, are efficient for large sets of data. The M5 model tree was first introduced by Quinlan (1992). The method was later improved to M5′ by Wang & Witten (1997) to handle enumerated attributes and missing values in the data. This model has been applied previously in various geo-environmental areas, including modelling water level discharge (Bhattacharya & Solomatine, 2005), sediment transport rate (Mafi *et al.*, 2013) or water movement in sandy soil (Samadianfard *et al.*, 2014).

The M5 algorithm consists of three processes: splitting of the input space, creating the tree and extracting the information from the tree (Figure 4). During the splitting process, which uses linear regression models and minimizes the errors between the measurements and predictions, the input space is divided into several regions (Figure 4a). The information obtained from the first process is used to form the tree. A schematic diagram of this process looks like an inverted tree in which the root is on top and the leaves are at the bottom (Figure 4b). In the next step, based on the relations between data and models established in step one, nodal points are established to create a tree-like structure that links a particular model with a particular dataset. Finally in the third step, which is known as the derivation process, a new data record is introduced at the root of the tree (Figure 4c). The newly introduced data record finds its way down by passing through the nodes. When a leaf is reached, the related prediction of the introduced data record is obtained and the data record is classified on the basis of the class appointed to that leaf.

Compared with the M5 algorithm, the M5′ algorithm consists of four steps: (i) forming a tree with a splitting criterion, (ii) development of a linear regression function at each node, (iii) pruning the tree to avoid the problem of over-fitting and (iv) smoothing the tree to compensate for the sharp discontinuities caused by the splitting. Standard deviation reduction (SDR) is used to split the input space and construct the regression tree with the M5′ algorithm. The SDR is the maximum expected reduction in output errors after branching and is given as:

$$\mathrm{SDR} = \sigma(T) - \sum_i \frac{T_i}{|T|} \times \sigma\left(T_i\right), \tag{6}$$

where T is the set of data points that reach the node, T_i is the data point that results from splitting at the node and falls into one sub-space according to the chosen splitting parameter and σ is the standard deviation (Wang & Witten, 1997). Standard deviation is considered as a measure of error for the data points that fall into a subspace. The M5′ algorithm constructs a linear multiple regression model for each inner node after forming the tree. The data related to that node and all the tested attributes in the sub-tree rooted at that node are used to form the model. After that, if the SDR for the linear model in the root of a sub-tree is smaller or equal to the expected error for the sub-tree, the tree is pruned from the leaves. After pruning, there might be discontinuities between the pruned leaves and adjacent linear models, which can be smoothed by combining the estimated value with the predicted one as follows:

$$P' = (np + kq) / (n + k), \tag{7}$$

where P' is the prediction passed up to the next higher node, p is the prediction passed to this node from below, q is the value predicted by the model at this node, n is the number of training instances that reach the node below and k is the smoothing constant. The smoothing process substantially improves the accuracy of predictions (Wang & Witten, 1997).

The model trees form linear models only; therefore, a common practice to account for possible non-linear relations is to use a logarithmic transformation. In this research, a logarithmic transformation of the data was used to develop a model to predict K_S. The WEKA program was applied to set up the M5′ model based on the existing 864 values. The model was calibrated with 70% of the values and was verified with the rest of the data to measure the accuracy of the prediction method.

Genetic programming

Genetic programming (GP) was used as the second method to formulate an equation to estimate K_S with the DRI test information; it is an extension of genetic algorithms (GAs). It is an example of artificial intelligence inspired by biological evolution and an evolutionary algorithm-based methodology to find computer programs that perform a given computational task (Fatehnia *et al.*, 2014b). Genetic programming has been used previously to estimate K_S (Parasuraman *et al.*, 2007). It was initially proposed by Koza (1992) as a domain-independent problem-solving approach in which computer programs composed of functions and terminals are evolved to solve, or approximately solve, problems by generating a structured representation of the data that imitates the biological evolution of living organisms and emulates naturally occurring genetic operations. The modelling steps used in GP start with the creation of an initial population of computer models (also called chromosomes) that are composed of a randomly selected set of functions and terminals defined by the user to suit a certain problem. The functions and terminals are arranged in a treelike structure to form a computer model that contains a root node, branches of functional nodes and terminals (Figure 5).

Functions and terminals used in GP can be standard arithmetic operations, Boolean logic functions, numerical constants, trigonometric functions, logical constants, variables and user-defined operators (Sette & Boullart, 2001). The first step of the analysis was to determine a set of functions that we considered to represent the nature of the problem or data. Next, the ‘fitness function’ was

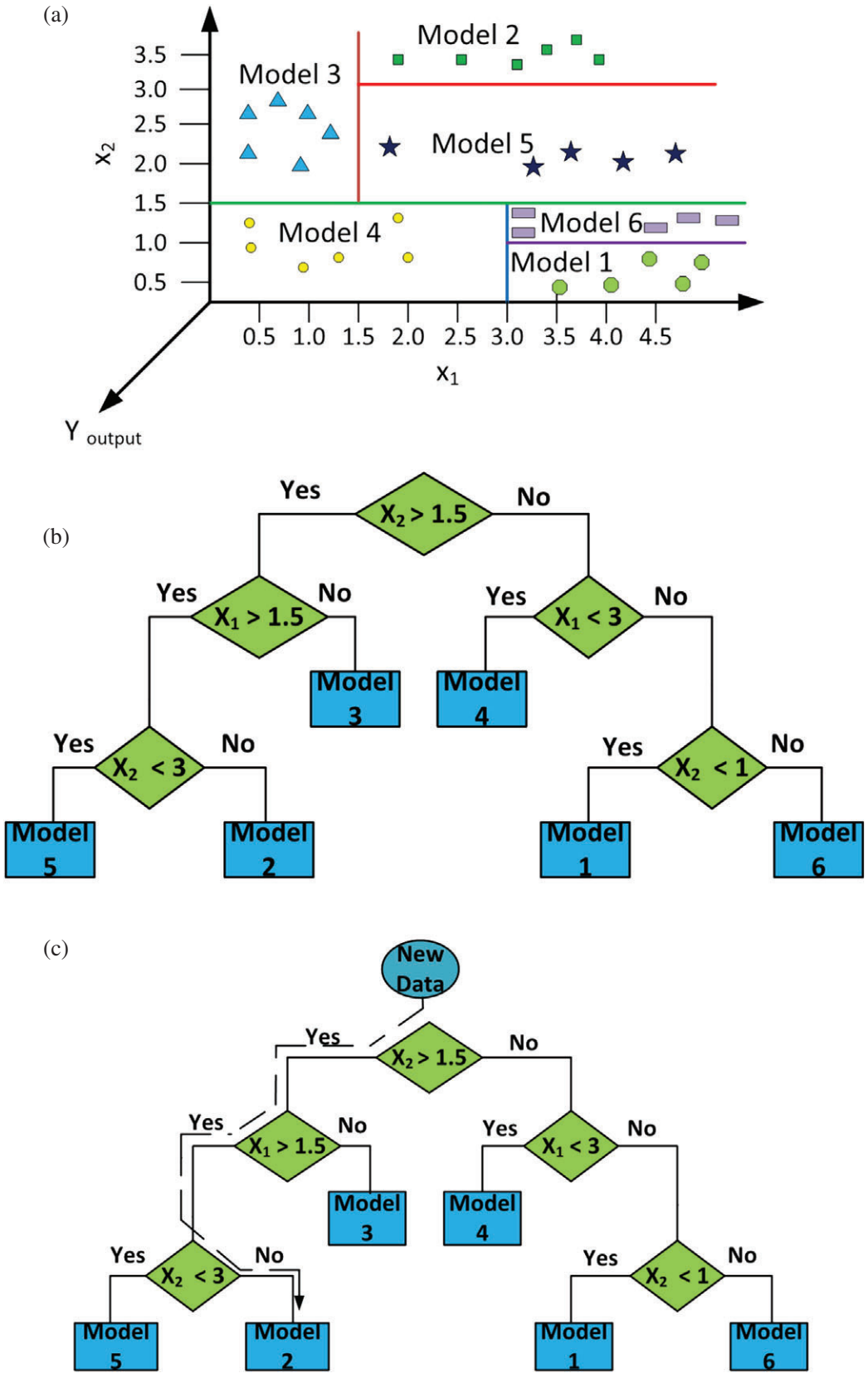

Figure 4 Schematic diagram of an *M5′* algorithm: (a) splitting of the input space, (b) forming the tree and (c) prediction of a new data value.

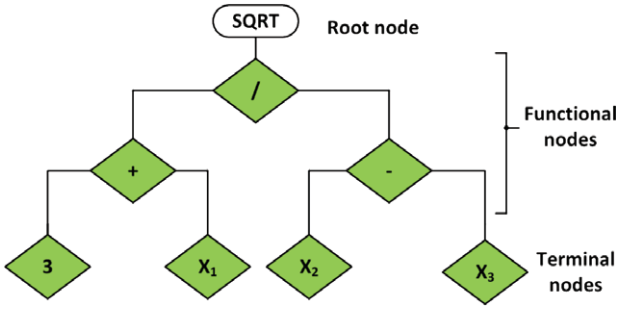

Figure 5 Typical example of GP tree representation for the function $[(3+X_1)/(X_2-X_3)]^2$.

assigned. In this research, the fit of the model was evaluated by the error between the output provided by the model and the desired actual output (Nikoo *et al.*, 2011). The fit, f_i, of an individual chromosome i was measured with the following expression:

$$f_i = \sum_{j=1}^{C_t} \left(M - \left| C_{(i,j)} - T_j \right| \right) , \quad (8)$$

where M is the range of selection, $C_{(i,j)}$ is the value returned by the individual chromosome i for its fit to j (out of C_t cases of fit) and T_j is the target value for the fit of j.

In GP, a new population is created by applying reproduction, crossover and mutation. These operations are applied to certain proportions of the computer models in the existing population, and the models are selected according to how well they fit. Reproduction is the copying of a computer model from an existing population into the new population without any change. Crossover is the genetic recombining of randomly chosen parts of two computer models. Mutation is the replacement of a randomly selected functional or terminal node with others from the same function or terminal set. The evolutionary process of evaluating the fit of an existing population and producing a new population is continued until a termination criterion is met, which can be either an acceptable error or a maximum number of generations. Finally, the best computer model is generated by GP using the fitness function adopted.

In this research, the analysis was carried out with the GPlab, a GP toolbox for MATLAB developed by Silva & Almeida (2003). We used a lexicographic parsimony pressure (lexictour) procedure to select parents for reproduction in the GP analysis. The population size and maximum generation were 50 and 100, respectively. The maximum tree depth was set to 20 and the maximum number of generations was assumed to be the termination criterion. The best individual that was found after running the program 50 times was stored and converted to a mathematical formula.

Results and discussion

The effect of each variable on the measured steady infiltration rate was analysed. The fixed values for other variables were selected randomly to observe the effect of changing a variable at a given time. An increase in ring diameter for all four soil materials resulted in a decrease in infiltration rate (Figure 6). Deeper rings had smaller infiltration rates than the shallower ones. This can be attributed to the reduction in horizontal flow by inserting the rings deeper into the soil (Figure 7). Infiltration rates were larger, as expected, for higher

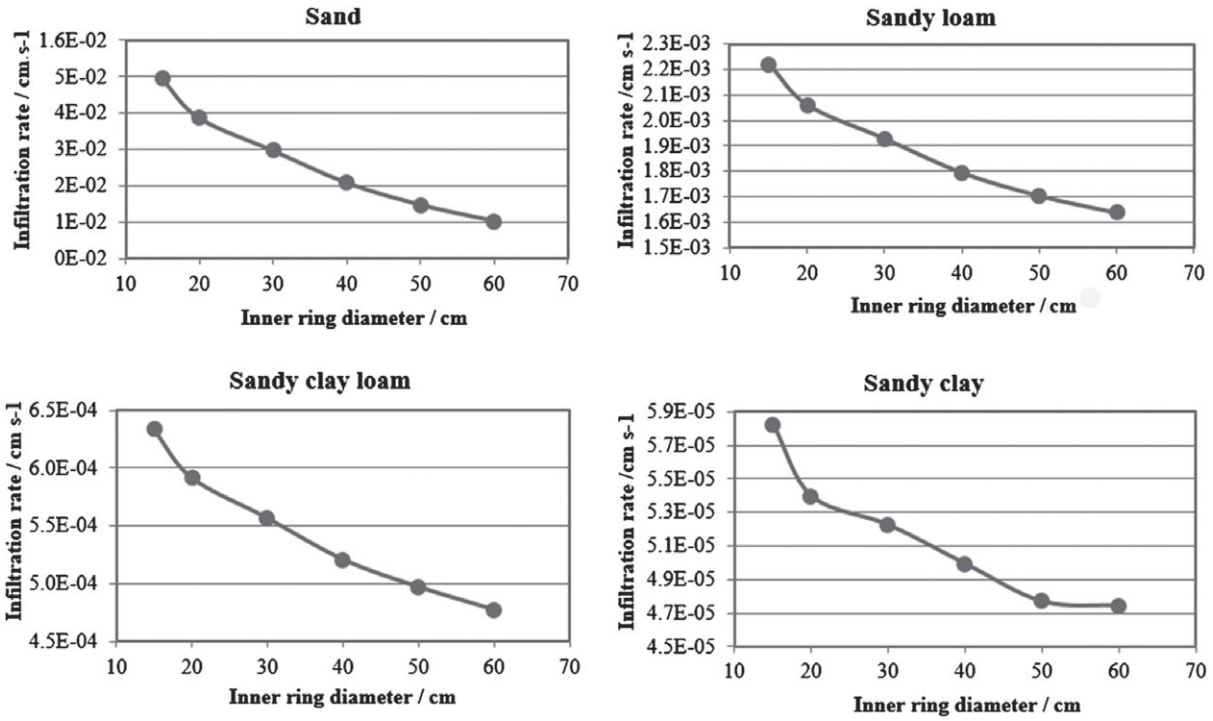

Figure 6 Effect of ring size on I for four different soil types with $H = 10$ cm, $D = 10$ cm and $S = 0.8$.

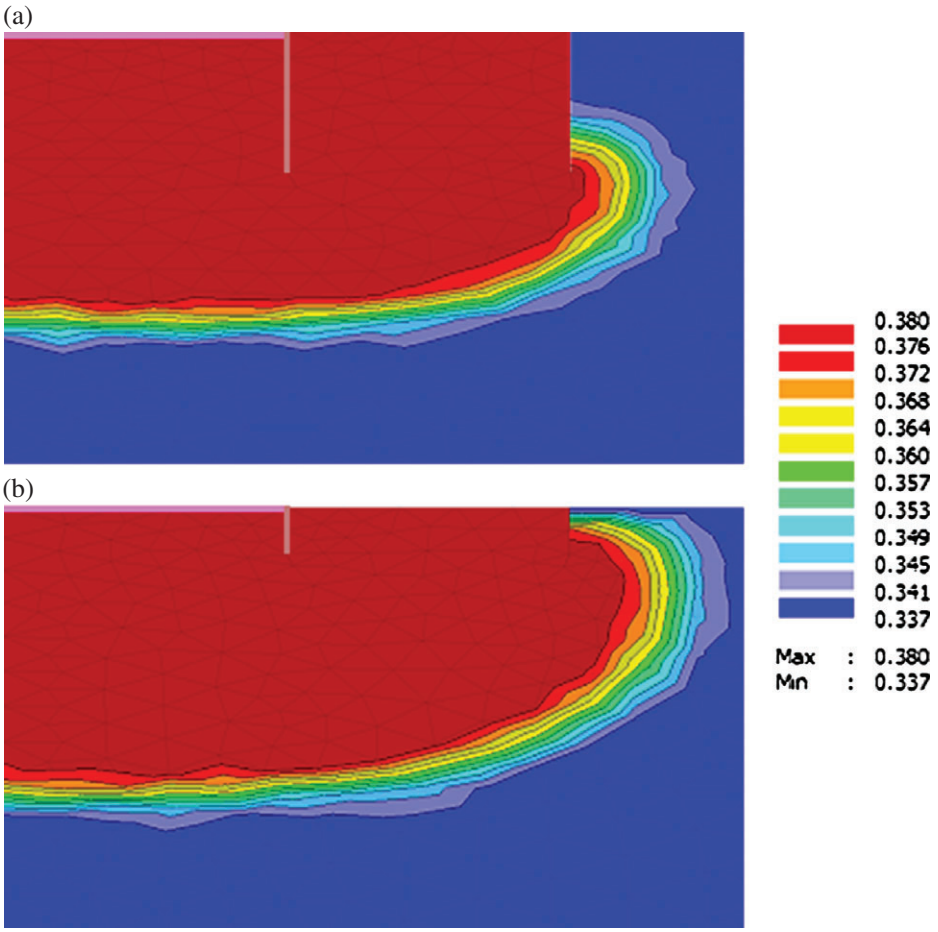

Figure 7 Comparison of the water contents at the end of the experiment in two similar DRI tests with (a) 15-cm and (b) 5-cm ring depths.

ponding heads with a greater hydraulic gradient. Reduced infiltration rates with increased soil water content could be attributed to larger capillary forces in drier soil. This was observed in both field tests and simulations. Finally, by comparing the values for steady infiltration rate of four simulated soil types, larger infiltration rates were observed in soil with smaller values for macroscopic capillary length. Few DRI studies have taken the effect of soil type on K_S measurements into consideration (Reynolds *et al.*, 2002).

The M5′ model tree

With the use of the M5′ model tree, a double-criterion formula was derived by the program (Equation (9)). This formula can be used to predict K_S in the DRI test. It should be noted that in deriving this formula, the diameter of the outer ring was assumed to be twice that of the inner ring. The equation is given as:

$$\begin{cases} I/K_s = 1 + 1.10451 \times B^{0.53} & \text{For } \lambda < 15 \text{ cm}, \\ I/K_s = 1 + 0.7243 \times B^{0.5174} & \text{For } \lambda \geq 15 \text{ cm}, \end{cases}$$
$$B = \frac{H \times \lambda}{(S+1) \times d_i \times D}, \qquad (9)$$

where I is the steady infiltration rate (cm s^{-1}), K_S is saturated hydraulic conductivity (cm s^{-1}), S is effective saturation, λ is macroscopic capillary length (cm), H is head of ponding in the ring (cm), d_i is the inner ring diameter (cm) and D is the ring insertion depth (cm).

The accuracy of Equation (9) was assessed with the 30% validation data, which were selected randomly by the program. The correlation coefficient, mean absolute error, root mean squared error (RMSE), relative absolute error and root relative squared error of the equation were determined by comparing the K_S values from the simulations with the estimates of K_S from Equation (9) (Table 3). Estimated I/K_s values from Equation (9) were also plotted against the simulated values (Figure 8). This plot shows that for values of

Table 3 Assessment of accuracy of estimates from Equations (9) and (10)

Measures of error of prediction	Equation (9)	Equation (10)
Correlation coefficient	8.618E−01	8.065E−01
Mean absolute error / cm s^{-1}	2.243E−01	2.954E−01
Root mean squared error / cm s^{-1}	2.823E−01	3.413E−01
Relative absolute error / %	49.59	59.18
Root relative squared error / %	51.49	61.04

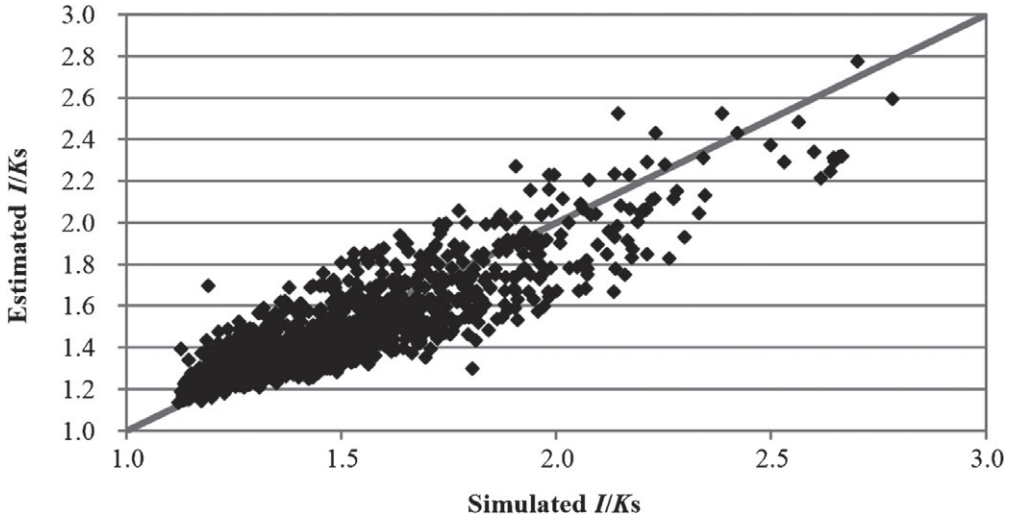

Figure 8 Scatter plot of 864 simulated I/K_s values against I/K_s values predicted from Equation (9) ($R^2 = 0.75$).

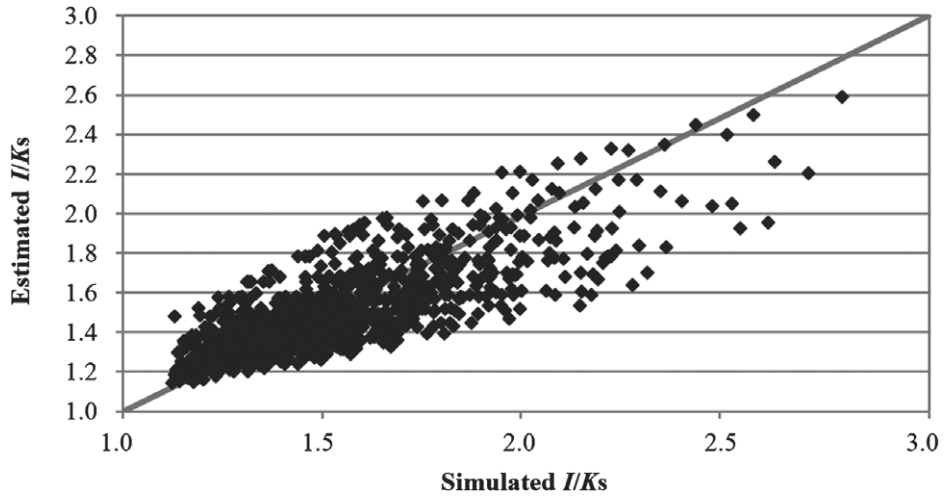

Figure 9 Scatter plot of 864 simulated I/K_s values against I/K_s values predicted from Equation (10) ($R^2 = 0.65$).

$I/K_s < 1.75$ the estimates from Equation (9) and finite element simulations are similar, but for values of $I/K_s > 1.75$ Equation (9) gives smaller estimated values than the simulations.

Genetic programming

For the GP method, similar to the M5′ model, 70% of the 864 values were used for prediction by the equation and the remaining 30% were used to assess the accuracy of estimation. Equation (10), which uses only head of ponding in the rings, inner ring diameter, effective saturation and steady infiltration rate as the independent variables to predict the dependent variable K_s, was identified as the most appropriate equation to estimate saturated hydraulic conductivity by the GP method. In Equation (10) the diameter of the outer ring was assumed to be twice that of the inner ring:

$$I/K_s = 1 + 1.21889 \times \left(\frac{H}{S + d_i}\right)^{0.70814}. \quad (10)$$

Accuracy of the method was determined from the error values obtained by comparing simulated K_s values with estimates from Equation (10) (Table 3). The simulated I/K_s values were also compared with the estimates from Equation (10) (Figure 9). Figure 9 shows that for I/K_s values > 1.4 Equation (10) gives smaller estimates than the finite element simulations.

Figures 8 and 9 show that the values derived from Equation (9) are more concentrated around the optimal line than the values derived from Equation (10). This indicates that Equation (9) selected by the M5′ method gives more accurate predictions than the genetic programming formula using the input variables.

Analysis of field data

To check the applicability of Equations (9) and (10) for estimating *in situ* saturated hydraulic conductivities, the K_s values of 30 field DRI experiments at four sites were predicted with these

Table 4 Comparison of the Guelph permeameter measured K_s with estimates of it from Equations (9) and (10)

Test location	d_i / cm	H / cm	D / cm	S	Measured steady infiltration rate / cm s^{-1}	λ / cm	Guelph permeameter measured K_s / cm s^{-1}	Equation (9) predicted K_s / cm s^{-1}	Equation (10) predicted K_s / cm s^{-1}
FAMU-FSU College of	15	5	15	0.3	1.13E−02	5.0	0.92E−02	8.69E−03	6.00E−03
Engineering, Tallahassee,	15	10	10	0.6	1.51E−02			1.02E−02	5.51E−03
Florida	15	15	5	1.0	1.57E−02			8.90E−03	4.40E−03
	20	5	5	1.0	1.13E−02			8.27E−03	6.62E−03
	20	10	10	0.6	1.39E−02			9.84E−03	5.72E−03
	20	15	15	0.3	1.46E−02			9.99E−03	4.61E−03
	20	15	15	0.6	1.37E−02			9.70E−03	4.35E−03
	30	5	10	0.6	1.11E−02			9.02E−03	7.21E−03
	30	10	5	1.0	1.26E−02			8.83E−03	6.08E−03
	30	15	10	1.0	1.36E−02			9.95E−03	5.22E−03
Florida State University	15	5	15	1.0	1.04E−02	6.0	1.43E−02	8.23E−03	5.60E−03
Reservation Camp,	15	10	10	0.3	1.51E−02			9.49E−03	5.46E−03
Tallahassee, Florida	15	15	5	0.6	2.14E−02			1.10E−02	5.92E−03
	20	5	5	0.3	1.29E−02			8.56E−03	7.49E−03
	20	10	10	1.0	1.26E−02			8.97E−03	5.23E−03
	30	5	15	1.0	0.98E−02			8.29E−03	6.38E−03
	30	10	5	0.3	1.41E−02			8.86E−03	6.75E−03
	30	15	15	0.6	1.38E−02			1.01E−02	5.26E−03
Tallahassee International Airport,	15	5	5	0.3	1.39E−02	8.5	1.32E−02	8.12E−03	7.38E−03
Tallahassee, Florida	15	15	15	0.6	1.49E−02			9.10E−03	4.12E−03
	20	5	15	0.6	1.09E−02			8.35E−03	6.35E−03
	20	15	5	1.0	1.72E−02			9.20E−03	5.52E−03
	30	5	10	1.0	1.05E−02			8.26E−03	6.84E−03
	30	15	15	0.3	1.41E−02			9.45E−03	5.36E−03
Greenway, Tallahassee, Florida	15	5	5	0.6	2.10E−02	10.0	1.12E−02	1.24E−02	1.12E−02
	15	15	15	1.0	2.14E−02			1.32E−02	6.00E−03
	20	5	15	0.6	1.57E−02			1.18E−02	9.15E−03
	20	15	10	0.6	2.41E−02			1.39E−02	7.66E−03
	30	5	5	0.6	1.72E−02			1.16E−02	1.12E−02
	30	15	10	1.0	2.04E−02			1.33E−02	7.83E−03

equations. Predicted values were compared with the K_s values derived from Guelph permeameter tests. The RMSE values for estimates from Equations (9) and (10) were 3.46E−03 cm s^{-1} and 5.98E−03 cm s^{-1}, respectively (Table 4). A comparison of the accuracies of the two equations and also of the error values (Table 3) shows that Equation (9) derived with the M5′ model tree gave the better results. This equation predicts K_s from all five selected variables, whereas Equation (10) uses steady infiltration rate, head of ponding in the ring, effective saturation and inner ring diameter only as the effective variables.

Conclusions

The main purpose of this research was to provide accurate estimates of K_s while measuring *in situ* steady infiltration rate with DRI. Equation (9) from the M5′ tree model resulted in more accurate predictions of saturated hydraulic conductivity than the GP method; the correlation coefficients were 0.862 and 0.806, respectively, and the RMSE values were 0.282 and 0.341 cm s^{-1}, respectively. This formula also performed better than the equation selected by the GP method in estimating K_s values from the 30 field DRI experiments; the RMSEs were 3.46E-03 and 5.98E-03 cm s^{-1}, respectively. The measured errors confirmed the robustness and acceptable performance of the equation from the M5′ tree model for estimating *in situ* saturated hydraulic conductivity with the DRI. Further increases in the accuracy of predicting K_s are possible by increasing the number of *in situ* experiments and by applying different optimization techniques. Similar research is recommended to provide estimates of K_s for types of soil not considered here.

References

Ahuja, L.R., El-Swaify, S.A. & Rahman, A. 1976. Measuring hydrologic properties of soil with a double-ring infiltrometer and multiple-depth tensiometers. *Soil Science Society of America Journal*, **40**, 494–499.

ASTM 2007. Standard D422-63(2007)e2. *Standard Test Method for Particle-Size Analysis of Soils.* American Society for Testing and Materials, West Conshohocken, PA, doi: 10.1520/D0422-63R07E02, www.astm.org [accessed on 15 November 2014].

ASTM 2009. Standard D3385-09. *Standard Test Method for Infiltration Rate of Soils in Field Using Double-Ring Infiltrometer*. ASTM International, West Conshohocken, PA, doi: 10.1520/D3385-09, www.astm.org [accessed on 15 November 2014].

ASTM 2010. Standard D2216-10. *Standard Test Method for Laboratory Determination of Water (Moisture) Content of Soil and Rock by Mass*. ASTM International, West Conshohocken, PA, doi: 10.1520/D2216-10, www.astm.org [accessed on 15 November 2014].

Bhattacharya, B. & Solomatine, D.P. 2005. Neural networks and M5 model trees in modelling water level-discharge relationship. *Neurocomputing*, **63**, 381–396.

Bouwer, H. 1986. Intake rate: cylinder infiltrometer. In: *Methods of Soil Analysis Part 1: Physical and Mineralogical Methods* (ed. A. Klute), pp. 825–844. American Society of Agronomy and Soil Science Society of America, Madison, WI.

Chowdary, V.M., Rao, M.D. & Jaiswal, C.S. 2006. Study of infiltration process under different experimental conditions. *Agricultural Water Management*, **83**, 69–78.

FAO 1998. *World Reference Base for Soil Resources*. World Soil Resources Report No 84, Food and Agriculture Organization of the United Nations, Rome.

Fatehnia, M. 2015. *Automated method for determining infiltration rate in soils.* Electronic thesis, Treatises and Dissertations. Paper 9327, Florida State University, Tallahassee, FL.

Fatehnia, M., Tawfiq, K. & Abichou, T. 2014a. Comparison of the methods of hydraulic conductivity estimation from mini disk infiltrometer. *Electronic Journal of Geotechnical Engineering*, **19**, 1047–1063.

Fatehnia, M., Tawfiq, K., Hataf, N. & Ozguven, E.E. 2014b. New method for predicting the ultimate bearing capacity of driven piles by using flap number. *KSCE Journal of Civil Engineering*, **19**, 611–620.

Fatehnia, M., Paran, S., Kish, S. & Tawfiq, K. 2016. Automating double ring infiltrometer with an Arduino microcontroller. *Geoderma*, **262**, 133–139.

Gardner, W.R. 1958. Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Science*, **85**, 228–232.

Gregory, J.H., Dukes, M.D., Miller, G.L. & Jones, P.H. 2005. Analysis of double-ring infiltration techniques and development of a simple automatic water delivery system. *Applied Turfgrass Science*, **2**, 1–7.

Guisheng, F., Yonghong, H. & Min, S. 2011. Experimental study on the reasonable inbuilt-ring depth of soil one-dimensional infiltration experiment in field. In: *Computer and Computing Technologies in Agriculture V* (eds D. Li & Y. Chen), pp. 427–436. Springer, Heidelberg.

Koza, J.R. 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge, MA.

Lai, J. & Ren, L. 2007. Assessing the size dependency of measured hydraulic conductivity using double-ring infiltrometers and numerical simulation. *Soil Science Society of America Journal*, **71**, 1667–1675.

Lai, J., Luo, Y. & Ren, L. 2010. Buffer index effects on hydraulic conductivity measurements using numerical simulations of double-ring infiltration. *Soil Science Society of America Journal*, **74**, 1526–1536.

Lazarovitch, N., Ben-Gal, A., Šimůnek, J. & Shani, U. 2007. Uniqueness of soil hydraulic parameters determined by a combined wooding inverse approach. *Soil Science Society of America Journal*, **71**, 860–865.

Léger, E., Saintenoy, A. & Coquet, Y. 2013. Estimating saturated hydraulic conductivity from surface Ground-penetrating radar monitoring of infiltration. In: *Proceedings of 4th International Conference on HYDRUS Software Applications to Subsurface Flow and Contaminant Transport Problems* (ed. J. Šimůnek, M. Th. van Genuchten & R. Kodešová). 1–11. Department of Soil Science and Geology, Czech University of Life Sciences, Prague.

Mafi, S., Yeganeh-Bakhtiary, A. & Kazeminezhad, M.H. 2013. Prediction formula for longshore sediment transport rate with M5' algorithm. In: *Proceedings of the 12th International Coastal Symposium* (eds D.C. Conley, G. Masselink, P.E. Russell & T.J. O'Hare), pp. 2149–2154. Coastal Processes Research Group, Plymouth.

Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, **12**, 513–522.

Nikoo, M.R., Kerachian, R., Niksokhan, M.H. & Hatami Bahman Beiglou, P. 2011. A game theoretic model for trading pollution discharge permits in river systems. *International Journal of Environmental Science & Development*, **2**, 162–166.

Parasuraman, K., Elshorbagy, A. & Si, B.C. 2007. Estimating saturated hydraulic conductivity using genetic programming. *Soil Science Society of America Journal*, **71**, 1676–1684.

Quinlan, J.R. 1992. Learning with continuous classes. In: *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence* (eds A. Adams & L. Sterling), pp. 343–348. World Scientific, Singapore.

Raats, P.A.C. & Gardner, W.R. 1971. Comparison of empirical relationships between pressure head and hydraulic conductivity and some observations on radially symmetric flow. *Water Resources Research*, **7**, 921–928.

Reynolds, W.D. & Elrick, D.E. 1990. Ponded infiltration from a single ring: I. Analysis of steady flow. *Soil Science Society of America Journal*, **54**, 1233–1241.

Reynolds, W.D., Elrick, D.E. & Youngs, E.G. 2002. Single-ring and double-or-concentric-ring infiltrometers. In: *Methods of Soil Analysis, Part 4, Physical Methods*, Number 5 in the SSSA Book Series (ed. J.H. Dane &. G.C. Topp) 821–826. Soil Science Society of America Journal, Inc., Madison, WI.

Richards, L.A. 1931. Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, **1**, 318–333.

Samadianfard, S., Nazemi, A.H. & Sadraddini, A.A. 2014. M5 model tree and gene expression programming based modeling of sandy soil water movement under surface drip irrigation. *Agriculture Science Developments*, **3**, 178–190.

Schaap, M.G., Leij, F.J. & van Genuchten, M.T. 2001. Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of Hydrology*, **251**, 163–176.

Sette, S. & Boullart, L. 2001. Genetic programming: principles and applications. *Journal of Engineering Applications of Artificial Intelligence*, **14**, 727–736.

Silva, S. & Almeida, J. 2003. Dynamic maximum tree depth – a simple technique for avoiding bloat in tree-based GP. In: *Proceedings of Genetic and Evolutionary Computation* (eds E. Cant´u-Paz, J.A. Foster, K. Deb, L.D. Davis, R. Roy, U.M. O'Reilly, *et al.*), pp. 1776–1778. Springer, Berlin.

Šimůnek, J., Šejna, M. & van Genuchten, M.T. 1999. *HYDRUS 2D Software for Simulating Water Flow and Solute Transport in Two Dimensional Variably Saturated Media.* Version 2.0. International Groundwater Modeling Center, Colorado School of Mines, Golden, CO.

Šimůnek, J., Šejna, M. & van Genuchten, M.T. 2006. *The HYDRUS Software Package for Simulating Two- and Three-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media, User manual.* Version 1.0. PC Progress, Prague.

Swartzendruber, D. & Olsen, T.C. 1961. Model study of the double ring infiltrometers as affected by depth of wetting and particle size. *Soil Science*, **92**, 219–255.

Van Genuchten, M.T. 1980. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, **44**, 892–898.

Wang, Y. & Witten, I.H. 1997. Induction of model trees for predicting continuous classes. In: *Proceedings of Poster Papers of the European Conference on Machine Learning* (eds M. van Someren & G. Widmer), pp. 128–137. University of Economics, Faculty of Informatics and Statistics, Prague.

Witten, I.H. & Frank, E. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, CA.

White, I. & Sully, M.J. 1987. Macroscopic and microscopic capillary length and time scales from field infiltration. *Water Resources Research*, **23**, 1514–1522.

Youngs, E.G. 1987. Estimating hydraulic conductivity values from ring infiltrometer measurements. *Journal of Soil Science*, **38**, 623–632.

Zhang, R. 1998. Estimating soil hydraulic conductivity and macroscopic capillary length from the disk infiltrometer. *Soil Science Society of America Journal*, **62**, 1513–1521.