Computers and Electronics in Agriculture 122 (2016) 168–175

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

A new predictive model for furrow irrigation infiltration using gene expression programming

Mohamed A. Yassin [a,*], A.A. Alazba [a,b], Mohamed A. Mattar [b,c]

[a] Alamoudi Water Research Chair, King Saud University, Riyadh, Saudi Arabia
[b] Agricultural Engineering Department, King Saud University, Riyadh, Saudi Arabia
[c] Agricultural Engineering Research Institute (AEnRI), Agricultural Research Center, Giza, Egypt

ARTICLE INFO

Article history:
Received 23 October 2015
Received in revised form 30 December 2015
Accepted 21 January 2016
Available online 12 February 2016

Keywords:
Gene expression programming
Furrow irrigation
Infiltrated water volume

ABSTRACT

This study investigates the ability of gene expression programming (GEP) in modeling of the infiltrated water volume (Z) under furrow irrigation. Field data were collected in the literature study for modeling Z covering wide range of opportunity time. Five variables were used as input parameters; inflow rate (Q_o), furrow length (L), waterfront advance time at the end of the furrow (T_L), infiltration opportunity time (T_o) and cross-sectional area of the inflow (A_o). The following statistical parameters that coefficient of determination (R^2), overall index of the model performance (OI), root mean square errors (RMSE) and mean absolute errors (MAE) are used as comparing criteria for the evaluation of the models' performances. The best value of the statistical parameters which range in training, testing and validation processes as the following (R^2 = 95–97%; OI = 94–97%; RMSE = 0.013–0.009 $m^3\ m^{-1}$; and MAE = 0.011–0.007 $m^3\ m^{-1}$) implies that the GEP model provides an excellent fit for the measured data. A comparison is made between the estimates provided by the GEP and the two-point method. The comparison results reveal that the GEP models are superior to two-point method. Furthermore, the remarkable advantage of GEP was that it resulted in an explicit equation for the estimation of the Z under furrow irrigation.

1. Introduction

Infiltration characteristics of soil are one of the most important parameters in the design, assessment and management of furrow irrigation (Esfandiari and Maheshwari, 1997). Estimation of soil infiltration is a major problem in irrigation studies due to proper selection of the technique used to determine the parameters of the infiltration models, the use of empirical models and its dependence on soil moisture, soil characteristics and surface roughness. Thus, the techniques used to determine soil infiltration characteristics must be appropriate for the purpose of the study (Walker and Busman, 1990).

It is necessary to utilize the mathematical models for simulation of surface irrigation because of reducing costs and decrease of time in analysis of indexes including application efficiency and distribution uniformity (Mahdizadeh Khasraghi et al., 2015). Several empirical equations have been developed to calculate infiltration that is a function of time through a surface irrigation event. The parameters of these equations are derived by fitting them to the actual cumulative infiltration data. The equations are then used to estimate cumulative infiltration and infiltration rates (Ravi and Williams, 1998). Some commonly used infiltration equations, which have no apparent physical basis, are the Horton, Kostiakov and Modified Kostiakov equations.

The mathematical models of surface irrigation are important for the evaluation and design purposes may be classified into four main categories. These models are the hydrodynamic (HD), the zero inertia (ZI), the kinematic wave (KW), and the volume balance (VB). Valipour and Montazar (2012a, 2012b, 2012c), and Valipour (2012) compared the HD, ZI, and KW models to optimize infiltration parameters in furrow irrigation systems. The authors concluded that performance of the HD and ZI was similar and better than the KW model in all irrigation events. Recently, Valipour et al. (2015) observed from more than one hundred data review that the priority of irrigation methods to simulate using HD and other models is border, basin, and furrow irrigation. It is border, furrow, and basin for KW and VB models. Finally, this priority is basin, border, and furrow for ZI model.

Various approaches have been used to estimate the parameters of empirical infiltration functions from VB and irrigation evaluation data. Therefore, it is important to analyze the problems of irrigation management and design (Strelkoff et al., 2009). Alazba (1999) stated that "Several models with various solutions have

* Corresponding author.
E-mail addresses: myassin1@ksu.edu.sa, athema2000@gmail.com (M.A. Yassin).

http://dx.doi.org/10.1016/j.compag.2016.01.035
0168-1699/

been made [for example, Lewis and Milne (1938), Hall (1956), Fok and Bishop (1965), Chen (1966), Bassett (1972), Kincaid et al. (1972), Sakkas and Strelkoff (1974), Katopodes and Strelkoff (1977), Strelkoff and Clemmens (1981), Elliott et al. (1982), Walker and Skogerboe (1987), Schmitz and Seus (1990), Valiantzas (1993, 1997), Alazba and Strelkoff (1994)]. Most of these studies are based on either VB or the ZI. While VB represents the simple and less accurate models, ZI represents the complex and more accurate ones. Engineers usually prefer to solve an engineering problem, particularly for practical and routine tasks, with simple models. Therefore, the VB is commonly used in surface irrigation design, evaluation, and management, because the sophisticated models require extensive programming and high computer cost due to the long execution time". Elliott and Walker (1982) developed a simplified technique which a well-known two-point method that uses only two points from the advance phase, usually at mid-distance and at the downstream end of the field. They assumed that the advance equation can be described by a power function. Holzapfel et al. (2004) evaluated four different methods to determine for two furrow sizes narrow (0.4 m top width) and wide (0.6 m top width). The results showed that that the two point method performed much better than the other methods when applied to wide furrow and only slightly worse when used in narrow furrows. Bautista et al. (2009) reported that the two-point method is one of the best known techniques to determine empirical infiltration parameters from surface irrigation evaluation data and mass balance, mainly because of its limited data requirements and mathematical simplicity. Ebrahimian et al. (2010) concluded that the two-point method had good performance in prediction of the infiltration for both furrow and border irrigation. There are several other similar methods based on the VB (Norum and Gray, 1970; Wu, 1971; Lal and Pandya, 1972; El-Shafei, 1980; Oyonarte and Mateos, 2003; Holzapfel et al., 2004), but the main difficulty with them is the use of particular forms of the infiltration equation. In many field situations, the form of a particular infiltration equation may not fit the field data, and therefore these methods may also not be suitable.

Recent technological progress has generated extremely accurate and reliable computer-aided modeling tools. An efficient paradigm in this area is a pattern recognition system that is effectively capable of learning from experience (Gandomi et al., 2012). Gene expression programming (GEP) is one of these intelligent systems, which are able to map input–output relationships without any understanding of the physical process involved. GEP was invented by Ferreira (2001b) and is the natural development of genetic algorithms and genetic programming. GEP has been applied in fields as diverse as artificial intelligence, artificial life, engineering and science, financial markets, industrial, chemical and biological processes and mechanical models. It has been used to solve problems such as symbolic regression, multi-agent strategies, time series prediction, circuit design and evolutionary neural networks (Samadianfard, 2012). GEP has been used in a number of hydrological and hydraulic modeling problems. Guven and Aytek (2009) used a GEP approach to model the stage–discharge relationship and compared the results with conventional methods. They found that the explicit algebraic formulations resulting from the GEP approach gave the best results. In a similar study, Azamathulla et al. (2011) developed mathematical models to estimate the stage–discharge relationship for the Pahang River based on GP and GEP techniques. Marti et al. (2013) evaluated the performance of the artificial neural networks (ANN), GEP and Multi Linear Regression (MLR) for estimating dissolved oxygen at sand filter outlet using data from 769 experimental filtration cycles. The results indicated that the GEP model tended to provide the most accurate estimations, followed by ANN and, lastly, by MLR models. Yassin et al. (2016) compared GEP, ANN to estimate daily reference evapotranspiration under arid conditions in Saudi Arabia.

Ravi and Williams (1998) reported that the unsaturated or the vadose zone is a key element of the hydrological cycle, directly influencing infiltration, storm runoff, evapotranspiration, interflow, and aquifer recharge. Understanding the nature of water movement in the vadose zone and its quantification is essential to solving a variety of problems. Therefore, evaluation of infiltration parameters for a field is difficult. Numerous models are available for performing simulations related to the movement of water. Often these models use over-simplified estimates of infiltration, which have little basis in reality and do not reflect actual field conditions well. However, the practical application of these infiltration models has not been adequately addressed. Therefore, the objectives of this study were (1) to use the GEP technique to build a predictive model for the volume of infiltrated water in a furrow and to find a generalized solution for infiltration that can be applied to different soils and furrow conditions and (2) to compare field estimated water infiltrated volume with both of the results of a recently completed GEP model and the two-point method using a VB model as described by Walker (1989).

2. Materials and methods

2.1. Database sets

The available database was used for development the GEP model was obtained from published literature. A total of 159 data points were collected from six studies (Valiantzas et al., 2001; Alvarez, 2003;Holzapfel et al., 2004; Playán et al., 2004; Mateos and Oyonarte, 2005; Sepaskhah and Shaabani, 2007). The studies were conducted in different locations differing in soil types and in furrow geometries. Many discharge rates were tested at each location. Two separate experiments were conducted in the field. The first experiment studied furrow infiltration using an infiltrometer to estimate infiltration parameters of the empirical Kostiakov equation. The second experiment consisted of measurements including land leveling conditions, furrow discharges, furrow cross-sections; advance time, recession time and hydraulic roughness. Stations were marked at certain distances from the furrow head. In the advancing phase, as the inflow water entered the furrows and reached each station, the time of reach, water depth, surface water width, flow cross-section and wetted perimeter were determined. After termination of the inflow, the time of water disappearance at each station was recorded to determine the recession times. Then, the infiltration opportunity time along the furrow length at each station was calculated as the time difference between when water disappeared and when it first started to advance at the same point along the furrow. Both experiments were conducted on the same field and performed under similar conditions. The infiltration parameters established from the first experiment were used for estimating the furrow infiltrated water volume (Z) in the second experiment. Table 1 presents details of the sites and values of the process variables and outputs. Data sets for development of the GEP model were prepared and included inflow rate (Q_o), furrow length (L), waterfront advance time at the end of the furrow (T_L), infiltration opportunity time (T_o) and cross-sectional area of the inflow (A_o) as input variables, and Z as the output variable. The descriptive statistics of the data used in this study is given in Table 2.

2.2. Gene expression programming

Gene expression programming (GEP) is a new evolutionary artificial intelligence technique developed by Ferreira (2001a).

Table 1
Summary of the experimental data used and analysis of variance for the variables.

Soil	Variables						Location	References
	Q (l s^{-1})	L (m)	T_L (min)	T_o (min)	A_o (cm^2)	Z (m^3 m^{-1})		
Loam to clay loam	0.85	400	124.3	100–1000	15.8	0.01–0.06	USA	Valiantzas et al. (2001)
Silty clay loam	1.5	360	208	100–1000	29.7	0.04–0.21		
Sandy loam	2	360	400	100–1000	36.6	0.05–0.33		
Vertic Gleysol[a]	3	240	93	2.4–93	114.4	0.02–0.08	Cuba	Alvarez (2003)
Haplic Acrisol[a]	3	333	244	3.4–247.1	132.9	0.03–0.16		
Rodic Ferrasol[a]	4	333	140	1.1–134.1	171.4	0.01–0.1		
Eutric Vertisol[a]	6.6–7.5	380	94–97	1.5–97.1	507.1–558.2	0.01–0.1		
Clay loam	1.07	110	72	49.2	42.95	0.05	Chile	Holzapfel et al. (2004)
Loam (calcareous)	1.03–3.01	35–48	12.5–16	5–30.3	126.5–301.3	0.01–0.06	Spain	Playán et al. (2004)
Silty	0.92–1.28	234	66–167	364–245	60.6–94.3	0.05–0.19	Spain	Mateos and Oyonarte (2005)
Clay loam	1.24	100	40.7	68–99.3	222.7	0.03–0.05	Iran	Sepaskhah and Shaabani (2007)
P value	2.46×10^{-6}	3.65×10^{-6}	3.12×10^{-28}	1.61×10^{-26}	5.36×10^{-2}			

[a] According to FAO–UNESCO (1988) soil classification.

Table 2
Descriptive statistics of the variables used in the model development.

Parameter	Input Variables														
	Training					Testing					Validation				
	Q_{in} (l s^{-1})	L (m)	A_o (cm^2)	T_L (min)	T_o (min)	Q_{in} (l s^{-1})	L (m)	A_o (cm^2)	T_L (min)	T_o (min)	Q_{in} (l s^{-1})	L (m)	A_o (cm^2)	T_L (min)	T_o (min)
Mean	2.89	278.18	189.0	131.9	239.6	2.89	304.4	160.3	148.2	265.0	2.62	227.4	172.7	111.2	174.1
Standard Error	0.21	10.68	17.15	9.32	25.18	0.34	14.10	27.57	15.25	47.90	0.40	27.9	32.92	19.18	51.31
Median	2	333	126	97	87	3	333	114.41	140.0	134.10	3	237	123.6	93	83
Standard Deviation	2.24	113.12	181.55	98.73	266.50	1.91	78.47	153.54	84.93	266.72	1.61	111.8	131.71	76.75	205.2
Variance	5.02	12,798	32,961	9748	71,024	3.64	6158	23,577	7213	71,140	2.59	12,517	17,348	5890.1	42,124
Skewness	1.09	−0.947	1.15	1.37	1.09	0.956	−1.13	1.688	1.733	0.965	1.85	−0.54	1.83	0.44	0.90
Kurtosis	−0.19	−0.29	−0.13	1.55	0.26	0.197	2.03	1.896	3.567	0.421	5.08	−0.78	4.06	−0.74	−1.10
Maximum	7.50	400	558.19	400	1000	7.50	400	558.19	400	1000	7.50	380	558.19	244	545
Minimum	0.85	35	15.82	12.50	1.12	0.85	48	15.82	12.50	2.40	0.92	48	29.71	12.50	5.38

According to Ferreira (2001a,b) the primary difference between GEP and its predecessors, genetic algorithms (GAs) and genetic programming (GP), stems from the nature of the individuals: in GAs, the individuals are linear strings of fixed length (chromosomes). In GP, the individuals are nonlinear entities of different sizes and shapes (parse trees). In GEP, the individuals are encoded as linear strings of fixed length (chromosomes) that are expressed as nonlinear entities of different sizes and shapes.

GEP uses chromosomes, which are usually composed of more than one gene of equal length, and expression trees or programmes, which are the expressions of the genetic information encoded in the chromosomes (Ferreira, 2006). The chromosomes are composed of multiple genes, each gene encoding a smaller sub-programme. In GEP, the linear chromosomes represent the genotype and the branched expression trees represent the phenotype (Ferreira, 2001b). Fig. 1 illustrates the general GEP modeling procedure.

GEP is a complete genotype/phenotype system in which the genotype is totally separate from the phenotype. In contrast, in GP, the genotype and phenotype constitute one entangled mess, more formally referred to as a simple replicator system. As a result, GEP's genotype/phenotype system surpasses the GP system by a factor of 100–60,000 (Ferreira, 2001a,b). GEP models encode their information in linear chromosomes, which are later translated or expressed in expression trees. These computer programmes are usually developed to solve a particular problem and are selected according to their ability to solve that problem (Guven and Aytek, 2009).

2.2.1. Developing the GEP model

To develop the GEP model, the available data (159 data points) was distributed randomly into a training set (70% of the datasets), a testing set (20% of the data sets), and a validation set (10% of the data sets). The training and testing sets therefore contain 112 and 31 data points, respectively; while the validation set has the remaining 16 data points. In the training stage, the GEP is trained to get the best performance with minimum errors. The testing step was carried out to calculate statistical measures of goodness of fit (Kalogirou, 2001; Yang et al., 2003). The five input parameters used were Q_o, L, T_L, T_o and A_o. The Z was the output parameter. GEP model development consisted of five major steps (Ferreira, 2001a,b):

(1) Select the fitness function. The fitness (fi) of an individual program (i) is measured by:

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

where M is the selection range, $C(i,j)$ is the value returned by the individual chromosome i for fitness case j (out of C_t fitness cases) and T_j is the target value for fitness case j. If $|C_{(i,j)} - T_j|$ (the precision) $\leqslant 0.01$, then the precision is 0 and $f_i = f_{max} = C_t M$. The advantage of this fitness function is that the system can find the optimal solution by itself.

(2) Choose the set of terminals (T) and the set of functions (F) to create the chromosomes. For instance, the terminal set includes the following variables: Q_o, L, T_L, T_o, and A_o. The choice of functions depends on the user. In this study, basic arithmetic operators (+, −, ×, ÷) and some basic mathematical functions ($\sqrt{}$, exp) were used to get the optimum GEP model, as listed in Table 3.

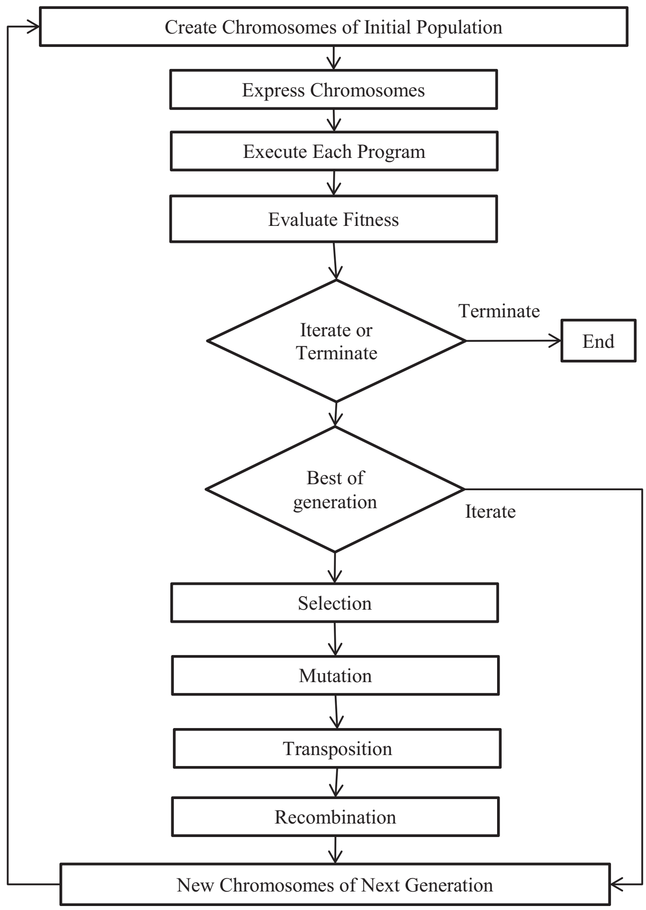

Fig. 1. Flow chart of the gene expression algorithm (Ferreira, 2001b).

Table 3
Parameters of the optimized GEP model.

Parameter	Description of parameter	Parameter setting
P1	Number of chromosomes	30
P2	Number of genes	3
P3	Function set	+, −, ×, ÷, sqrt, exp, X^2
P4	Linking function	Addition
P5	Fitness function	Mean squared error
P6	Mutation rate	0.00138
P7	Inversion rate	0.00546
P8	One-point recombination rate	0.00277
P9	Two-point recombination rate	0.00277
P10	Gene recombination rate	0.00277
P11	Gene transposition rate	0.00277
P12	Number of constants	7

(3) Choose the chromosomal architecture. A single gene and two head length was initially used. The number of genes and heads were increased one after another during each run and the training and testing performance of each model was monitored.
(4) Choose the linking function. Only addition or multiplication linking functions could be chosen for algebraic sub-trees.
(5) Select the set of GEP operators from mutation, transposition and recombination. This process was repeated for a pre-specified number of generations or until a solution was found.

In the present work, the GeneXpro program is used to predict the volume of infiltrated water in a furrow and to find a generalized solution for infiltration that can be applied to different soils and furrow conditions.

2.3. Two-point method

The VB is based on the principle of mass conservation, which is expressed by the following integral form (Lewis and Milne, 1938):

$$Q_o t_x = A_o x + \int_0^x Z(t_x - t_s) ds \tag{2}$$

where Q_o = inflow rate per furrow (m^3/min); t_x = time from the start of inflow (min); A_o = cross-sectional area of inflow (m^2); x = advance distance along the field (m); Z = cumulative infiltration volume per unit length as a function of opportunity time $(t_x - t_s)$ (m^3 m^{-1}); and t_s = advance time (min) to point s (m). The two-point method is based on the VB in which the measurement of advance times at two points, preferably one in the middle and the other at the end of a furrow, is used to calculate values of the infiltration parameters (Christiansen et al., 1966; Elliott and Walker, 1982; Burt et al., 1982; Blair and Smerdon, 1988). Eq. (2) may be written as:

$$Q_o t_x = \sigma_y A_o x + \sigma_z Z_o x \tag{3}$$

where Z_o = infiltrated area at the inlet (m^3 min^{-1}); σ_y = surface water profile shape factor; and σ_z = subsurface water profile shape factor. Two of the simplest and most commonly used approximations for infiltration are the Kostiakov equation which can be written in general terms for furrow irrigation as (Walker et al., 2006):

$$Z = kt^a \tag{4}$$

and the modified Kostiakov equation

$$Z = k'\tau^{a'} + f_o \tau \tag{5}$$

where Z = cumulative infiltration in units of volume per unit length of furrow (m^3 min^{-1}); τ = elapsed time of infiltration (min); f_o = basic infiltration rate (m^3 m^{-1} min^{-1}); and k, a, k' and a' = empirical coefficients which must be determined experimentally and which vary with the type of soil and its condition. k and k' have units of m^3 m^{-1} min^{-a} and m^3 m^{-1} min$^{-a'}$, respectively.

According to infiltration theory, the exponents a and a' in Eqs. (2) and (3) should lie between 0 and 1, with most observed values lying between 0.2 and 0.9 (Blair and Reddell, 1983; Serralheiro, 1988). Substituting the value of Zoin Eq. (3) by Eq. (5) gives:

$$Q_o t_x = \sigma_y A_o x + \sigma_z k' t_x^{a'} x + \frac{f_o t_x x}{(1+r)} \tag{6}$$

If Eq. (2) is used instead of Z_o, the third term in Eq. (6) will be eliminated. The basic infiltration rate, f_o, needs to be determined independently. When the field evaluation includes measurements of inflow and outflow, f_o can be determined as:

$$f_o = \frac{Q_o - Q_{tw}}{L} \tag{7}$$

where Q_{tw} = tailwater outflow per furrow (m^3 min^{-1}). The value of σ_y is usually assumed to be constant during the advance and is in the range of 0.7–0.9 (Strelkoff and Katopodes, 1977). Walker and Skogerboe (1987) have proposed that σ_y = 0.77. Moreover, the average area of the surface stream is constant from the beginning of irrigation until the advance halts before or at the end of the field. The water depth at the inlet of the system is usually assumed to be the normal depth, which can be computed using Manning's equation as follows (Walker, 1989):

$$A_o = \left(\frac{Q_o n}{p_1 S_o^{0.5}}\right)^{1/p_2} \tag{8}$$

where n = manning roughness coefficient; S_0 = field slope (m m^{-1}); p_1 and p_2 = parameters depending on the furrow geometry. For determining σ_z, many authors have used the basic assumption that the advance trajectory does not have a concise mathematical description, but can be reasonably well approximated with the simple power function (Elliott and Walker, 1982; Walker and Skogerboe, 1987; Scaloppi et al., 1995):

$$x = pt_x^r \tag{9}$$

where p and r = fitting parameters without physical interpretation. Elliott and Walker (1982) made several comparisons of Eq. (11) with more elaborate relationships and methods of fitting and concluded that the best results are achieved by a two-point fitting of the equation. So, Kiefer (1965) derived the following expression:

$$\sigma_z = \frac{a + r(1-a) + 1}{(1+a)(1+r)} \tag{10}$$

Eq. (6) can be written for the two common advance points which are the mid-distance of the furrow ($0.5L$, $t_{0.5L}$) and the end of the furrow (L, t_L), to provide a simultaneous solution for k' and a' as follows:

$$a' = \frac{\log(V_L/V_{0.5L})}{\log(t_L/t_{0.5L})} \tag{11}$$

and

$$k' = \frac{V_L}{\sigma_z t_L^{a'}} \tag{12}$$

In which

$$V_{0.5L} = \frac{Q_0 t_{0.5L}}{0.5L} - \sigma_y A_0 - \frac{f_0 t_{0.5L}}{(1+r)} \tag{13}$$

and

$$V_L = \frac{Q_0 t_L}{L} - \sigma_y A_0 - \frac{f_0 t_L}{(1+r)} \tag{14}$$

2.4. Performance criteria

The infiltrated water volume obtained from the GEP technique and two-point method were evaluated by computing four standard statistical performance evaluation criteria. The statistical performance evaluation criteria that were used to reflect the goodness of simulation can be expressed as:

$$R^2 = \frac{\left(\sum_{i=1}^{n}(E_i - \overline{E})(C_i - \overline{C})\right)^2}{\sum_{i=1}^{n}(E_i - \overline{E})^2 \cdot \sum_{i=1}^{n}(C_i - \overline{C})^2} \tag{15}$$

$$\mathrm{OI} = \frac{1}{2}\left[1 - \frac{\mathrm{RMSE}}{E_{\max} - E_{\min}} + \frac{\sum_{i=1}^{n}(E_i - C_i)^2}{\sum_{i=1}^{n}(E_i - \overline{E})^2}\right] \tag{16}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(E_i - C_i)^2}{n}} \tag{17}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|E_i - C_i|}{n} \tag{18}$$

where E_i = value of $\mathrm{ET_{ref}}$ estimated by the PMG; C_i = corresponding value calculated by mathematical $\mathrm{ET_{ref}}$ models; n = number of observations; $\overline{E}$ = average of the estimated values; and $\overline{C}$ = average of the calculated values. $E_{\max}$ = the maximum experimental value. $E_{\min}$ = the minimum experimental value.

The coefficient of determination (R^2) measures the degree of correlation between the estimated and calculated values, where values approaching 1.0 indicate a good correlation. The root mean square error (RMSE) expresses the error in the same units that describe the variable (Legates and McCabe, 1999). The lower the RMSE, the better the matching. The overall index of the model performance (OI) combines the normalized RMSE and the model efficiency value. An OI value of 1.0 indicates a perfect fit between a model's estimated and calculated values (Alazba et al., 2012; Mattar et al., 2015; Mattar and Alamoud, 2015). The mean absolute error (MAE) is the average value of the absolute differences between the estimated and calculated values. A low MAE implies good model performance.

3. Results and discussion

3.1. GEP-based formulation for infiltrated water volume under furrow irrigation

Formulations Z under furrow irrigation the best result by the GEP algorithm are as given below:

$$Z = \frac{Q_{in}^2}{(A + (-5.75 \times \sqrt{L}))^2} + \frac{(7.8 \times Q_{in}) + (-0.134 \times T_o)}{(A \times Q_{in}) + (T_L + 193.27)} + \frac{T_o}{(-0.0257 \times T_L \times A) + 1939 + L}$$

In which Q_{in} is in (l s^{-1}), L is in (m), T_L is in (min), T_o is in (min) and A_o is in (cm^2). The expression tree of the above formulation is shown in Fig. 2. A comparison of the observed values versus values predicted by GEP is shown in Figs. 3 and 4. As shown in this figure, the proposed equation can be separated into three independent components (subprograms or genes) linked by an addition function. Each subprogram represents an individual aspect of the problem such that a meaningful overall solution is developed (Ferreira (2001a,b)). Therefore, each of the evolved subprograms contains important information about the physiology of the final model. Each gene expressed in the final equation is responsible for resolving a particular facet of the problem. Such information provides an opportunity for further scientific discussion at genetic and chromosomal level (Ferreira (2001a,b)).

3.2. Performance analysis and model validity

The experimental data taken from literature are subdivided into three sets which are training, testing and validation. These sets are used to evaluate the generalization capacity of the GEP-based model. All of the results obtained from experimental studies and predicted by using the training and testing processes of GEP model are given in Figs. 3 and 4. As it is visible in Figs. 3 and 4 the values obtained from the training and testing processes are very close to the experimental results. The result of training and testing in Table 4 shows that the GEP model is capable of generalizing between input and output variables with reasonably good predictions. The high values of R^2 (97.2%) and OI (97%) and the low values of RMSE (0.01 m^3 m^{-1}) and MAE (0.008 m^3 m^{-1}) in training process, Furthermore, A similar results in testing process which are The high values of R^2 (95.8%) and OI (94.5%) and the low values of RMSE (0.01 m^3 m^{-1}) and MAE (0.011 m^3 m^{-1}) show that the GEP model is suitable and can predict the infiltrated water volume under furrow irrigation and the values very close to the experimental values. However, no rational model to predict the all systems of surface irrigation (Furrows, basins, and borders) has been developed yet that would encompass the influencing variables considered in this study. Therefore, it is not possible to generalize the developed GEP furrow model for all systems of surface irrigation.

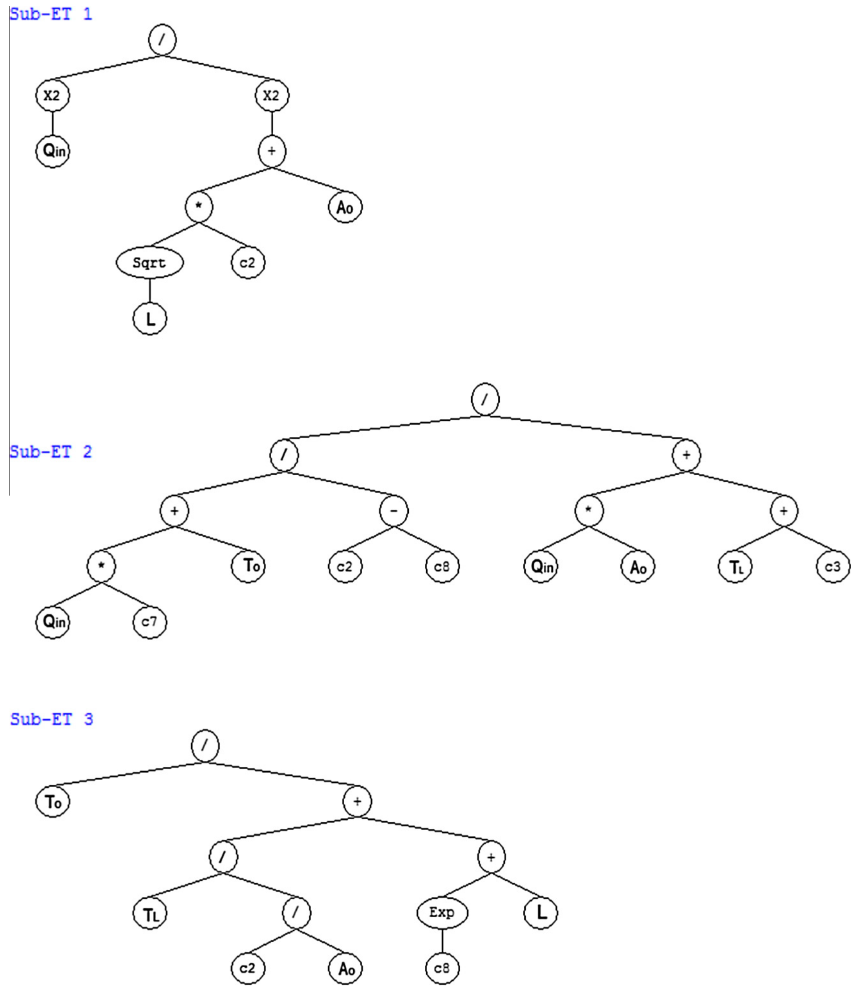

Fig. 2. Flow chart of the Gene Expression Algorithm (Ferreira, 2001b).

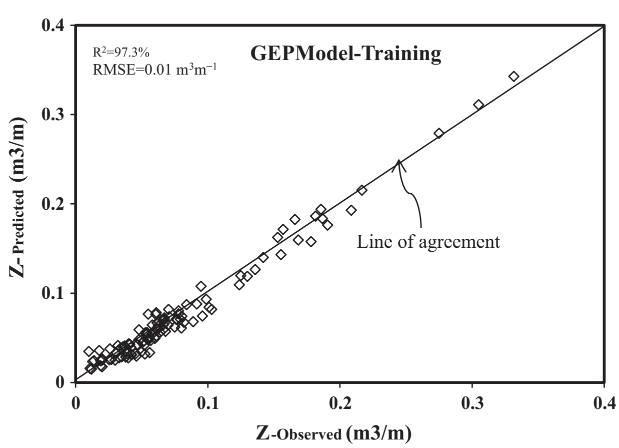

Fig. 3. Comparison of observed and GEP-predicted values for the training data set.

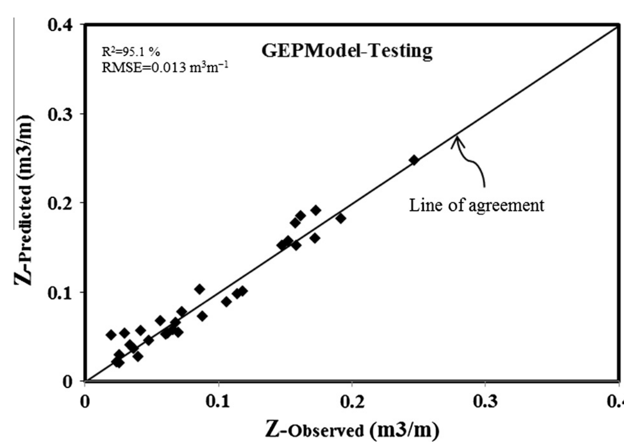

Fig. 4. Comparison of observed and GEP-predicted values for the testing data set.

3.3. Comparison between the GEP model and the two-point method

The validation results of the GEP model was compared with the two-point method Table 4 and Fig. 5. It can be seen from the table and the figure that the GEP model performs better than the two-point method. Statistical analysis of the data shows a close relationship between the observed and the simulated series; the determination coefficient R^2 of GEP model reached 96%. While the two-point method had the a R^2 value that was about 3.2% less accurate than GEP model. Similarly, The OI value for the GEP model was closer to one (95%) than its value (92%) for the two-point method. Furthermore, the value of RMSE for the two-point method (0.013 $m^3\,m^{-1}$) was almost 1.4 times that of the value (0.0092 $m^3\,m^{-1}$) for the GEP model and the MAE value for the two-point method (0.0094 $m^3\,m^{-1}$) was almost 1.3 times that of the value (0.0072 $m^3\,m^{-1}$) the GEP model. Table 4 and Fig. 5 show

Table 4
Statistical performance of the optimized GEP model and the two-point method during training, testing and validation.

Criteria	Training		Testing		Validation	
	GEP	Two-point	GEP	Two-point	GEP	Two-point
R^2 (%)	0.972	0.968	0.957	0.950	0.961	0.931
RMSE ($m^3\ m^{-1}$)	0.010	0.012	0.013	0.014	0.009	0.012
OI (%)	0.970	0.959	0.944	0.937	0.949	0.920
MAE ($m^3\ m^{-1}$)	0.008	0.009	0.011	0.012	0.007	0.009

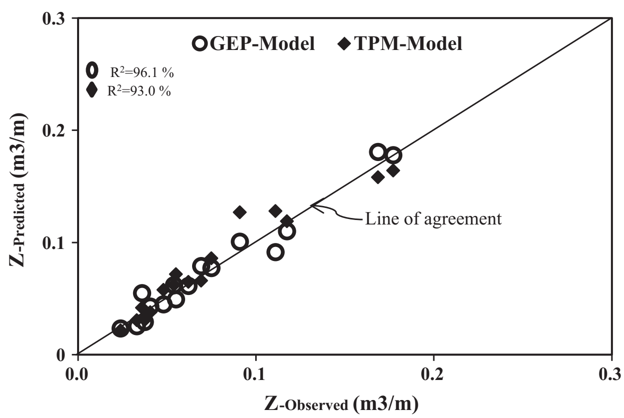

Fig. 5. Comparison of observed and predicted values for both the GEP model and the two-point method using the validation data set.

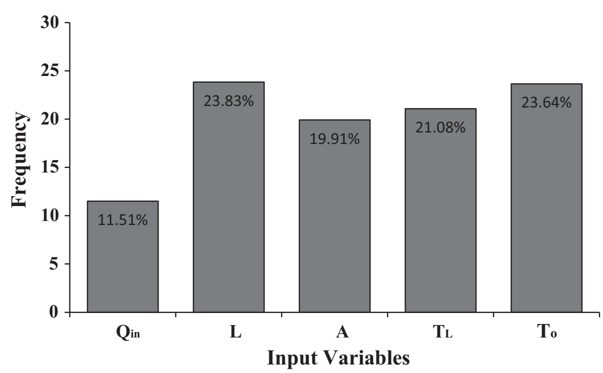

Fig. 6. Contributions of the input variables in the GEP model.

that some values predicted from the two-point method were inaccurate, though most of the predictions were acceptable when using the GEP model. This indicates that the two-point method is not a reliable forecasting procedure. Therefore, the GEP model is a good alternative to the two-point method to some extent. This agrees with (Shiri et al., 2012) who stated that the main advantage of GEP models over other models (e.g., the adaptive neuro-fuzzy inference system) is their ability to explicitly express the relationship between the dependent and independent variables. The GEP models are explicit and simple such that they can be used, by anyone not necessarily being familiar with GEP, in a spreadsheet on a very simple PC, even on a hand-held calculator (Landeras et al., 2012).

3.4. Contribution analysis

A sensitivity analysis was carried out to determine the contribution of the closely concerned predictor variables (Q_o, L, T_L, T_o, A_o) affecting the Z under furrow irrigation. A frequency value ranging from zero to one where value approaching 1.0 refers that this variable has been appeared in 100% of the best thirty programs evolved by GEP. This is a common approach in the GP-based analyses (Mollahasani et al., 2011). The frequency values of the predictor variables are presented in Fig. 6. As it is seen, T_o and L are same more sensitive (23.83%) to Z, followed by T_L (21.08%). Moreover, Q_o had less accurate (11.51%) than other infiltration parameters.

4. Conclusions

The Z under furrow irrigation is a complex phenomenon. In this study, soft computing technique, namely GEP is applied to estimate this phenomenon directly. The five input variables were Q_o, L, T_L, T_o and A_o. The data used to develop the GEP model collected from the experimental previous work. The GEP model was trained on 70% of the available data, tested using the remaining 20% and validated using the remaining 10%. Assess the performance of the new expression was conducted by comparing the predictions from the GEP model and the two point method with the experimental results. These comparisons showed that agreement between the predicted and observed data was reasonable for the two-point method but better for the GEP model. The best values of R^2 and OI are 96% and 95% and the minimum values of RMSE and MAE are 0.0092 and 0.0072 $m^3\ m^{-1}$ respectively, in validation phase of GEP. These results showed that GEP model are capable of predicting suitable results for the Z under furrow irrigation.

Acknowledgement

The project was financially supported by King Saud University, Vice Deanship of Research Chairs.

References

Alazba, A.A., Strelkoff, T., 1994. Correct form of Hall technique for border irrigation advance. J. Irrig. Drain. Div., ASCE 120 (2), 292–307.
Alazba, A.A., 1999. Explicit volume balance model solution. J. Irrig. Drain. Div. ASCE 125, 273–279.
Alazba, A., Mattar, M.A., ElNesr, M.N., Amin, M.T., 2012. Field assessment of friction head loss and friction correction factor equations. J. Irrig. Drain. Eng., ASCE 138 (2), 166–176.
Alvarez, J.A.R., 2003. Estimation of advance and infiltration equations in furrow irrigation for untested discharges. Agric. Water Manage. 60, 227–239.
Azamathulla, H.Md., Ghani, A.A., Leow, C.S., Chang, K.C., Zakaria, N.A., 2011. Gene-expression programming for the development of a stage–discharge curve of the Pahang River. Water Resour. Manage. 25, 2901–2916.
Bautista, E., Clemmens, A.J., Strelkoff, T.S., 2009. Structured application of the two-point method for infiltration parameter estimation in surface irrigation. J. Irrig. Drain. Eng., ASCE 135 (5), 566–578.
Bassett, D.L., 1972. A mathematical model of water advance in border irrigation. Trans. ASAE 15 (5), 992–995.
Blair, A.W., Reddell, D.L., 1983. Evaluation of empirical infiltration equations for blocked furrow infiltrometers. In: Chieago, ASAE Winter Meeting. Paper 83-2521.
Blair, A.W., Smerdon, E.T., 1988. Infiltration from irrigation advance data. II. Experimental. J. Irrig. Drain. Eng., ASCE 114, 18–30.
Burt, C.M., Robb, G.A., Hanon, A., 1982. Rapid evaluation of furrow irrigation efficiencies. In: ASAE Paper No. 82. ASAE, St. Joseph, MI, pp. 25–37.
Chen, C.L., 1966. Discussion of 'A solution of the irrigation advance problem.' by Otto Wilke and Ernest T. Smerdon. J. Irrig. Drain. Div., ASCE 92 (2), 97–101.
Christiansen, J.E., Bishop, A.A., Kiefer, F.W., Fok, Yu-Si, 1966. Evaluation of intake rate constants as related to advance of water in surface irrigation. Trans. ASAE 5, 671–674.
Ebrahimian, H., Liaghat, A., Ghanbarian-Alavijeh, B., Abbasi, F., 2010. Evaluation of various quick methods for estimating furrow and border infiltration parameters. Irrig. Sci. 28 (6), 479–488.
Elliott, R.L., Walker, W.R., Skogerboe, G.V., 1982. Zero-inertia modeling of furrow irrigation advance. J. Irrig. Drain. Div. ASCE 108 (3), 179–195.
Elliott, R.L., Walker, W.R., 1982. Field evaluation of furrow infiltration and advance functions. Trans. ASAE 25 (2), 396–400.
El-Shafei, Y.Z., 1980. General solution for the evaluation of advance of water in furrow irrigation: theory and practice. Z. Kulturtech. Flurbereinig. 21, 8–19 (in English).

Esfandiari, M., Maheshwari, B.L., 1997. Application of the optimization method for estimating infiltration characteristics in furrow irrigation and its comparison with other methods. Agric. Water Manage. 34, 169–185.
Ferreira, C., 2001a. Gene expression programming in problem solving. In: 6th Online World Conference on Soft Computing in Industrial Applications (Invited Tutorial).
Ferreira, C., 2001b. Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst. 13 (2), 87–129.
Ferreira, C., 2006. Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence, second ed. Springer-Verlag, Germany.
Fok, Y.S., Bishop, A.A., 1965. Analysis of water advance in surface irrigation. J. Irrig. Drain. Div. ASCE 91 (1), 99–116.
Gandomi, A.H., Babanajad, S.K., Alavi, A.H., Farnam, Y., 2012. Novel approach to strength modeling of concrete under triaxial compression. J. Mater. Civ. Eng. 24 (9), 1132–1143.
Guven, A., Aytek, A., 2009. A new approach for stage–discharge relationship: gene-expression programming. J. Hydrol. Eng. 14 (8), 812–820.
Hall, W.A., 1956. Estimating irrigation border flow. Agric. Eng. 37 (1), 263–265.
Holzapfel, E.A., Jara, J., Zu~niga, C., Mari~no, M.A., Paredes, J., Billib, M., 2004. Infiltration parameters for furrow irrigation. Agric. Water Manage. 68, 19–32.
Kalogirou, S.A., 2001. Artificial neural networks in the renewable energy systems applications: a review. Renew. Sustain. Energy Rev. 5, 373–401.
Katopodes, N., Strelkoff, T., 1977. Dimensionless solutions of border-irrigation advance. J. Irrig. Drain. Div. ASCE 103 (4), 401–417.
Kiefer, F.W., 1965. Average Depth of Absorbed Water in Surface Irrigation. Special Publication. Department of Civil Engineering, Utah State University, Logan.
Kincaid, D.C., Heermann, D.F., Kruse, E.G., 1972. Hydrodynamics of border irrigation advance. Trans. ASAE 15 (4), 674–680.
Lal, R., Pandya, A.C., 1972. Volume balance method for computing infiltration rates in surface irrigation. Trans. ASAE 15 (1), 69–72.
Landeras, G., Lopez, J.J., Kisi, O., Shiri, J., 2012. Comparison of Gene Expression Programming with neuro-fuzzy and neural network computing techniques in estimating daily incoming solar radiation in the Basque Country (Northern Spain). Energy Convers. Manage. 62, 1–13.
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 (1), 233–241.
Lewis, M.R., Milne, W.E., 1938. Analysis of border irrigation. Agric. Eng. 19, 267–272.
Mahdizadeh Khasraghi, M., Gholami Sefidkouhi, M.A., Valipour, M., 2015. Simulation of open- and closed-end border irrigation systems using SIRMOD. Arch. Agron. Soil Sci. 61 (7), 929–941.
Marti, P., Shiri, J., Duran-Ros, M., Arbat, G., Ramirez de Cartagena, F., Puig Bargués, J., 2013. Artificial neural networks vs. Gene Expression Programming for estimating outlet dissolved oxygen in micro-irrigation sand filters fed with effluents. Comput. Electron. Agric. 99, 176–185.
Mateos, L., Oyonarte, N.A., 2005. A spreadsheet model to evaluate sloping furrow irrigation accounting for infiltration variability. Agric. Water Manage. 76, 62–75.
Mattar, M.A., Alamoud, A.I., 2015. Artificial neural networks for estimating thehydraulic performance of labyrinth-channel emitters. Comput. Electron. Agric. 114 (5), 189–201.
Mattar, M.A., Alazba, A.A., Zin El-Abedin, T.K., 2015. Forecasting furrow irrigationinfiltration using artificial neural networks. Agric. Water Manage. 148, 63–71.
Mollahasani, A., Alavi, A.H., Gandomi, A.H., 2011. Empirical modeling of plate load test moduli of soil via gene expression programming. Comput. Geotech. 38 (2), 281–286.
Norum, D.E., Gray, D.M., 1970. Infiltration equations from rate of advance data. J. Irrig. Drain. Eng., ASCE 96 (2), 111–119.
Oyonarte, N.A., Mateos, L., 2003. Accounting for soil variability in the evaluation of furrow irrigation. Trans. ASAE 45 (6), 85–94.
Playán, E., Rodriguez, A.J., Garcia-Navarro, P., 2004. Simulation model for level furrows. I: Analysis of field experiments. J. Irrig. Drain. Eng., ASCE 130 (2), 106–112.
Ravi, V., Williams, J.R., 1998. Estimation of infiltration rate in the Vadose zone: compilation of simple mathematical models. In: Vol. I. EPA/600/R-87/128a. United States Environmental Protection Agency, 26.
Sakkas, J.G., Strelkoff, T., 1974. Hydrodynamics of surface irrigation—advance phase. J. Irrig. Drain. Div. ASCE 100 (1), 31–48.
Samadianfard, S., 2012. Gene expression programming analysis of implicit Colebrook-White equation in turbulent flow friction factor calculation. J. Petrol. Sci. Eng. 92–93, 48–55.
Scaloppi, E.J., Merkley, G.P., Willardson, L.S., 1995. Intake parameters from advance and wetting phases of surface irrigation. J. Irrig. Drain. Eng., ASCE 121 (1), 57–70.
Schmitz, G.H., Seus, G.J., 1990. Mathematical zero-inertia modeling of surface irrigation: advance in borders. J. Irrig. Drain. Div., ASCE 116 (5), 603–615.
Sepaskhah, A.R., Shaabani, M.K., 2007. Infiltration and hydraulic behavior of an anguiform furrow in heavy texture soils of Iran. Biosyst. Eng. 98, 248–256.
Serralheiro, R.P., 1988. Um estudo da rega por sulcos num solo Argiluviado (A Study of Furrow Irrigation on a Luvisol Soil). University of Evora, Evora (Doctoral dissertation presented to the University of Evora).
Shiri, J., Kisi, O., Landeras, G., Lopez, J.J., Nazemi, A.H., Stuyt, L., 2012. Daily reference evapotranspiration modeling by using genetic programming approach in the Basque Country (Northern Spain). J. Hydrol., 414–415
Strelkoff, T.S., Katopodes, N.D., 1977. Border-irrigation hydraulics with zero-inertia. J. Irrig. Drain. Div., ASCE 103 (3), 325–342.
Strelkoff, T., Clemmens, A.J., 1981. Dimensionless stream advance in sloping borders. J. Irrig. Drain. Div., ASCE 107 (4), 361–382.
Strelkoff, T.S., Clemmens, A.J., Bautista, E., 2009. Estimation of soil and crop hydraulic properties. J. Irrig. Drain. Eng., ASCE 135 (5), 537–555.
Valiantzas, J.D., 1993. Border advance using improved volume balance model. J. Irrig. Drain. Div. ASCE 119 (6), 1006–1013.
Valiantzas, J.D., 1997. Surface irrigation advance equation: variation of subsurface shape factor. J. Irrig. Drain. Div., ASCE 123 (4), 300–306.
Valiantzas, J.D., Aggelides, S., Sassalu, A., 2001. Furrow infiltration estimation from time to a single advance point. Agric. Water Manage. 52, 17–32.
Valipour, M., 2012. Comparison of surface irrigation simulation models: full hydrodynamic, zero inertia, kinematic wave. J. Agric. Sci. 4, 68–74.
Valipour, M., Gholami Sefidkouhi, M.A., Eslamian, S., 2015. Surface irrigation simulation models: a review. Int. J. Hydrol. Sci. Technol. 5 (1), 51–70.
Valipour, M., Montazar, A.A., 2012a. Optimize of all effective infiltration parameters in furrow irrigation using visual basic and genetic algorithm programming. Aust. J. Basic Appl. Sci. 6, 132–137.
Valipour, M., Montazar, A.A., 2012b. Sensitive analysis of optimized infiltration parameters in SWDC model. Adv. Environ. Biol. 6, 2574–2581.
Valipour, M., Montazar, A.A., 2012c. An evaluation of SWDC and WINSRFR models to optimize of infiltration parameters in furrow irrigation. Am. J. Sci. Res. 69, 128–142.
Walker, W.R., Skogerboe, G.V., 1987. Surface Irrigation: Theory and Practice. Prentice-Hall Inc, Englewood Cliffs, N.J.
Walker, W.R., 1989. Guidelines for designing and evaluating surface irrigation system. In: FAO Irrigation and Drainage Paper No. 45. FAO, Rome, pp. 137.
Walker, W.R., Busman, J., 1990. Real time estimation of furrow irrigation. J. Irrig. Drain. Eng. ASCE 116, 299–317.
Walker, W.R., Prestwich, C., Spofford, T., 2006. Development of the revised USDA–NRCS intake families for surface irrigation. Agric. Water Manage. 85, 157–164.
Wu, L.-P., 1971. Overland flow hydrograph analysis to determine infiltration function. Trans. ASAE 14 (2), 295–300.
Yang, I.H., Yeo, M.S., Kim, K.W., 2003. Application of artificial neural network to predict the optimal start time for heating system in building. Energy Convers. Manage. 44, 2791–2809.
Yassin, M.A., Alazba, A.A., Mattar, M.A., 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agr. Water Manage. 163, 110–124.