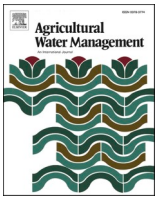
Agricultural Water Management 265 (2022) 107529

Contents lists available at ScienceDirect

Agricultural Water Management

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



Machine learning and regression-based techniques for predicting sprinkler irrigation's wind drift and evaporation losses

Mohamed A. Mattar [a,b,c], Dilip Kumar Roy [d], Hussein M. Al-Ghobari [a], Ahmed Z. Dewidar [c,*]

[a] Agricultural Engineering Department, College of Food and Agriculture Sciences, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia
[b] Agricultural Engineering Research Institute (AEnRI), Agricultural Research Centre, P.O. Box 256, Giza, Egypt
[c] Prince Sultan Bin Abdulaziz International Prize for Water Chair, Prince Sultan Institute for Environmental, Water and Desert Research, King Saud University, P. O. Box 2454, Riyadh 11451, Saudi Arabia
[d] Irrigation and Water Management Division, Bangladesh Agricultural Research Institute, Gazipur 1701, Bangladesh

ARTICLE INFO

Handling editor - Dr. B.E. Clothier

Keywords:
Sprinkler irrigation
Artificial neural network
Adaptive neuro-fuzzy Inference system
Multivariate adaptive regression spline
Probabilistic linear regression
Support vector regression
Evaporation and drift losses

ABSTRACT

Wind drift and evaporation losses (WDEL), which can occur as a result of operational and meteorological factors, are two of the most significant sprinkler-irrigation losses that can occur even in a well-managed irrigation system. A proper understanding of factors that influence WDEL in sprinkle irrigation is critical for developing water conservation strategies that significantly impact the quality and return on investment of irrigation projects. The specific objective of this research was to determine the predictive ability of five soft computing approaches (artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS), multivariate adaptive regression spline (MARS), probabilistic linear regression (PLR), and support vector regression (SVR)) for predicting WDEL on a sprinkler irrigation system under design, operational, and meteorological conditions. Datasets were collected from previously published studies conducted under a variety of conditions. The results showed that the five approaches yielded statistically different WDEL predictions. The ANN model produced the most accurate WDEL predictions compared to the other models with the training and testing dataset. The ANFIS, MARS, PLR, and SVR models' performance ranks were found to be inconsistent across a variety of statistical performance criteria. Hence, Shannon's entropy-based decision theory was used to rank these models. The MARS model was ranked second (0.896), followed by the ANFIS model (0.865), the PLR model (0.833), and the SVR model (0.794). The design variable "auxiliary nozzle diameter" and climate variable "wind speed" both had high contribution ratios (17.5% and 12.19%, respectively) in WDEL modeling to produce a robust predictive model. In general, the developed models, particularly the ANN model, demonstrated a high degree of accuracy in estimating the WDEL of sprinkler irrigation systems.

1. Introduction

Sprinkler irrigation is a method of applying water that saves up to 50% more water than surface irrigation while maintaining a high level of quality, affordability, and simplicity of installation (Li et al., 2015). Water flowing through the air towards the crop canopy or soil surface is a critical component of sprinkler irrigation efficiency, as some water does not reach the irrigated surface. This water is ineffectively utilized by the crop and is referred to as wind drift and evaporation losses (WDEL) (Sadeghi et al., 2017; Stambouli et al., 2013). Estimating WDEL from sprinklers is extremely important. Irrigation system design (too high is costly, too low causes crop water stress), water rights allocation (very contentious and involves enormous sums of money), and the overall requirements for irrigation water management are all influenced by irrigation efficiency, and as a result, the WDEL (Sarwar et al., 2021).

Sprinkler irrigation efficiency depends on various parameters such as the design variables (e.g., sprinkler type, sprinkler nozzle diameter, and sprinkler spacing) and operating variables (e.g., operating pressure, sprinkler height, and irrigation time), as well as weather conditions (Carrión et al., 2001; Playán et al., 2006). These parameters are responsible for a portion of the discharge water from the sprinkler nozzle that is wasted and does not reach the crop canopy. The amount of WDEL published in the literature have been reported to range from 2% and 50% with different operating and sprinkler system types under a wide

* Corresponding author.
E-mail addresses: mmattar@ksu.edu.sa (M.A. Mattar), dilip.roy@my.jcu.edu.au (D.K. Roy), hghobari@ksu.edu.sa (H.M. Al-Ghobari), adewidar@ksu.edu.sa (A.Z. Dewidar).

https://doi.org/10.1016/j.agwat.2022.107529
Received 17 November 2021; Received in revised form 2 February 2022; Accepted 3 February 2022
Available online 12 February 2022
0378-3774/Published by Elsevier B.V.

variety of climates (Playán et al., 2005, 2017; Sarwar et al., 2019). Nozzle diameter, one of the design aspects, plays a crucial role in the water drop diameter. In other words, a large nozzle diameter leads to large drop diameters that are less affected by WDEL (Keller and Bliesner, 1990). On the other hand, raising the nozzle height above the soil surface produces a higher WDEL due to the long drop trajectory and increased exposure to wind. In operating factors, a rise in operating pressure causes a decrease in drop diameters and increased WDEL (Montero et al., 2003). Wind speed is another essential meteorological factor that breaks droplet trajectory, causing a rise in both the droplet evaporation and WDEL (Dechmi et al., 2003; Playán et al., 2006; Zapata et al., 2007; Sanchez et al., 2010).

WDEL is estimated as the difference between the amount of water discharged by sprinklers and the water gathered using measurement techniques and mathematical methods based on physical principles (Yan et al., 2010; King et al., 2012; Sadeghi et al., 2015). Many researchers have documented the use of sprinkle irrigation to analyze the WDEL during the water application process (e.g., laboratory and field tests) (Uddin, 2012; Mohamed et al., 2021; Molaei et al., 2021). Due to the fact that various studies have varying degrees of accuracy, the results obtained vary significantly. As a result, estimation of the WDEL using these devices during field distribution experiments is imprecise and challenging for testing conducted with water-collecting devices. The issues connected with water-collection equipment (e.g., catch cans) and common experimental mistakes in the measuring procedure have encouraged the notion that mathematical models are critical in WDEL calculation. Precise estimation of the WDEL is important for water management and the optimal design of sprinkler systems (Maroufpoor et al., 2018).

In recent decades, several studies, such as Supharatid (2003), Kumar et al. (2004), Cigizoglu and Kisi (2006), Chang et al. (2010), Ismail et al. (2012), Kalra et al. (2013), Yassin et al. (2016), Petković et al. (2017), Mattar (2018), Mattar et al. (2020), Kuzman et al. (2021), Lakovic et al. (2021), and Roy et al. (2021) reported on successful applications of artificial intelligence techniques in a variety field of water engineering, irrigation, and agricultural topics. Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multivariate Adaptive Regression Spline (MARS), Probabilistic Linear Regression (PLR), and Support Vector Regression (SVR) are in use because of their ability to handle nonlinear problems, high processing speed, advanced parallelism, self-learning, and generalization capabilities (Basheer and Hajmeer, 2000). These artificial techniques are centered on computational intelligence and machine learning technologies that utilize mathematics and statistics to correlate the system's input parameters to the system's outputs (Hamdia et al., 2015). For instance, ANN is able to learn from examples, recognize data patterns, and process information quickly, whereas ANFIS presents a neural networking combination and a fuzzy system dealing with higher-level thinking. Consequently, these techniques may provide a feasible alternative to empirical or physically based analytical formulas with a high number of variables (Hamdia et al., 2015). MARS is a method for modeling high-dimensional data in a flexible way. MARS does not prescribe a certain type of link between the response and predictor variables; instead, it expands in product spline functions, with the number of spline functions and interactions chosen dynamically by the data (Friedman, 1991; Friedman and Roosen, 1995). MARS is a semiparametric technique that frequently bridges the gap between parametric and nonparametric approaches (Huang et al., 2019). PLR is a type of regression modeling that employs a Bayesian technique for parameter estimation. There are three distributions in the Bayesian approach: a prior, a likelihood distribution, and a posterior distribution. Bayesian parameter estimation is accomplished by multiplying the prior distribution by the likelihood (Permai and Tanty, 2018). The SVR is a high-performance learning system based on constrained optimization theory that uses the notion of structural error minimization induction to provide an optimal solution (Tang et al., 2019).

For accuracy, more than one model is often utilized in which the outputs of different models are compared. For example, Tabari et al. (2010) used ANN and multivariate non-linear regression approaches to estimate daily pan evaporation in a semi-arid region of Iran. Similarly, Kişi (2006) compared the ANFIS technique to the ANN and Stephens–-Stewart methods for estimating daily evaporation. Several prior studies have also proved the MARS approach's potential applicability in other problem issues with reasonable prediction accuracy. These include scour-depth estimates (Samadi et al., 2015), runoff predictions (Adamowski et al., 2012; Sharda et al., 2008), downscaling daily precipitations (Nasseri et al., 2013), and potential groundwater mapping (Zabihi et al., 2016). Application of PLR, random forest, and cross-correlation function to investigate the effectiveness rate of the climatic parameters on the amount of potential evapotranspiration of barley during the growth period was studied by Zarei et al. (2021). The performance of SVR with hybrid nature-based algorithms in order to model soil cation exchange capacity was investigated by Emamgholizadeh and Mohammadi (2021).

However, due to the time-consuming and challenging nature of field measurements, mathematical algorithms have been developed to rapidly and accurately estimate the WDEL of sprinkler irrigation systems operating under a variety of operating situations. To the authors' best knowledge, no studies have compared the capability of the ANN, ANFIS, MARS, PLR, and SVR methods for modeling WDEL. Therefore, the aims of the study are to (1) investigate the applicability of three machine learning techniques (i.e., ANN, ANFIS, and SVR) and two regression techniques (i.e., MARS and PLR) for modeling WDEL from sprinkler irrigation systems, (2) assess the effectiveness of the ANN, ANFIS, MARS, PLR, and SVR's predictive capabilities to select the best prediction technique, and (3) determine which relevant factors influencing WDEL using the models' contribution analysis.

2. Material and methods

2.1. Experimental datasets

Table 1 shows the experimental data (Abo-Ghobar, 1993; Dechmi et al., 2003; Sanchez et al., 2010, 2011) used to develop the ANN, ANFIS, MARS, PLR, and SVR models. In total, 109 data points were gathered in various locations. To determine WDEL, experiments were conducted utilizing sprinklers with double nozzles. In detail, these experiments are described as follows:

The WDEL from sprinklers was tested in the first experiment (Abo-Ghobar, 1993) under various meteorological and operating circumstances combinations. During the months of April to July 1991, the experiment was carried out at the educational farm, Riyadh, Saudi Arabia. A total of ten types of impact sprinklers with various nozzle sizes were chosen for field irrigation. Water was collected in the catch cans after 1–2 h of operation, depending on the size of the sprayer's nozzle. In the second experiment (Dechmi et al., 2003), the effect of a solid set sprinkler system on maize (Zea mays L.) yield was investigated. The experiment was carried out from May to August 2000 at the Zaragoza experimental farm, Spain. The sizes of the nozzles were 4.4 mm for the main and 2.4 mm for the auxiliary, and they were positioned 2.30 m above the soil surface. The working pressure of the nozzle was maintained at 300 kPa throughout the season, resulting in an 11-meter wetted radius. In the third experiment (Sanchez et al., 2010), WDEL was determined using data gathered from pluviometers above alfalfa and maize that had been irrigated with sprinkler solid-set system under similar operational and technical conditions. The experiment was conducted at Zaragoza's experimental farm, Spain, in 2006. The main nozzle had a jet-straightening vane with a diameter of 4 mm, while the auxiliary nozzle had a diameter of 2.4 mm. In the fourth experiment (Sanchez et al., 2011), the impact sprinklers' performance was investigated with regard to the nozzle diameter, pressure, and meteorological conditions. This study was conducted in the Zaragoza experimental farm over the course of two years, 2003 and 2004. The main nozzle's

Table 1
A summary of the experimental data used to construct the models.

Input variables									Output variable	Location	References
H_r (m)	P (kPa)	d_m (mm)	d_a (mm)	q_m (L s^{-1})	q_a (L s^{-1})	WS (m s^{-1})	T (°C)	RH (%)	WDEL (%)		
1	200–300	5.5–6.1	3.0–4.2	0.41–0.50	0.10–0.12	2.33–3.65	34–42	42.0–56.7	22.6–34.0	Saudi Arabia	Abo-Ghobar (1993)
2.3	300	4.4	2.4	0.26	0.08	0.6–6.5	12–31	31–64	6–39.6	Spain	Dechmi et al. (2003)
2.3	290–355	4	2.4	0.10–0.33	0.04–0.12	0.8–5.6	19–27	44–68	3–33	Spain	Sanchez et al. (2010)
2	230–420	4.0–4.8	2.4	0.19–0.37	0.07–0.09	0.4–8.0	5.0–27.1	40–86	2.4–35.6	Spain	Sanchez et al. (2011)

Hr = riser height, P = operating pressure, dm = main nozzle diameter, da = auxiliary nozzle diameter, qm = water discharge by main nozzle, qa = water discharge by auxiliary nozzle, WS = wind speed, T = air temperature, RH = relative humidity, and WDEL = wind drift and evaporation losses.

diameter was 4.4 mm, with a spreader nozzle of 2.4 mm. All tests were conducted for two hours with a pressure range of 180–420 kPa and low wind conditions. In each experiment with the solid-set, the amount of water collected in catch-cans and pluviometers following each irrigation event is recorded and divided by the number of catch-cans to obtain the average irrigation depth ($\overline{ID}_C$). The WDEL is calculated using the sprinkler system's irrigation applied depth (ID_D) and the $\overline{ID}_C$ (Christiansen, 1942) as:

$$\text{WDEL} = \frac{ID_D - \overline{ID}}{ID_D} \quad (1)$$

$$ID_D = \frac{C_D \times A \times (2gp)^n \times t}{S_S \times S_L} \quad (2)$$

where C_D is a coefficient of discharge, A is the orifice area of the nozzle, g is acceleration of gravity, n is the exponent of discharge, t is the running time, and both S_S and S_L are the sprinkler's allocated area.

According to the findings of these studies, a portion of the applied water is lost to evaporation, most notably by a drift away from the irrigated region. Table 1 summarizes the experimental data used to create the models, collected at several sites and under various design, operational, and weather conditions. The data sets which are used as input variables to create the suggested models included design, operational, and meteorological parameters. The design and operating parameters are riser height (H_r), operating pressure (P), main nozzle diameter (d_m), auxiliary nozzle diameter (d_a), and the water discharged by both the main (q_m) and auxiliary (q_a) nozzles. Wind speed (WS), air temperature (T), and relative humidity (RH) are all examples of meteorological parameters. To estimate WDEL, the ANN, ANFIS, MARS, PLR, and SVR models were developed. All data (109 data points) were randomly divided into 70% for training and 30% for testing. It is worth noting that all models use the same training and test data sets to allow for fair performance comparison. The training data set must have many values to include all of the problem's characters and ensure accuracy. It is important also that the training data set contains all of the problem's maximal and lowest values.

The statistics parameters for the inputs and outputs used to train and test the proposed models are summarized in Table 2. Maximum, mean, minimum, standard deviation, kurtosis coefficient, and skewness coefficient are denoted as X_m, X_a, X_n, S_x, K_x, and C_{sx}, respectively. Maximum values for design and operating parameters (H_r, P, d_m, d_a, q_m, q_a) were highly similar in both training and testing sets, as were minimum values. For both sets, the maximum values of WS and RH were greater than 7 m s^{-1} and 80%, respectively. The d_a had the largest K_x and C_{sx} magnitudes compared to the other variables.

2.2. Artificial neural network (ANN)

ANN has a fundamental ability to identify and comprehend complex and nonlinear relationships through experience. (Swingler, 2001). Three layers of neurons comprise the ANN: an input layer (independent variables, *i*), a hidden layer (*j*), and an output layer (dependent variables, *k*) (Fig. 1). To design the optimal ANN structure, the number of hidden layers, the neurons within each hidden layer, the learning algorithm, and the activation functions must all be carefully chosen (Dogan et al., 2008) (Table 3). The most widely used algorithm for training feed-forward ANNs is back-propagation (Thirumalaiah and Deo, 1998; Ciğizoğlu, 2003; Jain and Srinivasulu, 2004; Fernando and Shamseldin, 2009). Nonlinear and complex functions are more accurately modeled by having hidden layers in feed-forward networks (Salehi and Razavi, 2012). Accordingly, this study used the back-propagation learning algorithm to train the ANN with a hidden and an output layer. The network input and output are automatically standardized to [0.15, 0.85] prior to transferring the data to the ANN for training. This normalization speeds up the training process and significantly enhances the network's generalization capabilities. After normalization, the input data are processed unidirectionally from the input layer to the hidden layer and finally to the output layer. Each neuron (information processing unit, *n*) in the ANN is composed of three components: input connection links, a central processing unit, and output connection linkages. Weights (W) are assigned to the connections to store information (the strength of the link between two neurons in neighboring layers) that is then processed

Table 2
A descriptive analysis of the training and testing variables.

Variables	Training set						Testing set					
	X_m	X_a	X_n	S_x	K_x	C_{sx}	X_m	X_a	X_n	S_x	K_x	C_{sx}
H_r (m)	2.3	2.11	1	0.27	9.26	-2.67	2.3	2.05	1	0.37	4.38	-2.16
P (kPa)	420	313.2	200	59.34	-0.53	0.29	420	311.91	200	59.69	-0.17	0.34
d_m (mm)	6.1	4.35	4	0.4	4.6	1.74	6.1	4.42	4	0.56	3.42	1.86
d_a (mm)	4.2	2.46	2.4	0.3	31.53	5.64	4.2	2.49	2.4	0.34	21.31	4.45
q_m (L s^{-1})	0.5	0.28	0.1	0.06	2.7	0.3	0.5	0.3	0.19	0.07	2.27	1.28
q_a (L s^{-1})	0.12	0.09	0.04	0.02	-0.52	0.12	0.12	0.09	0.07	0.02	-1.51	0.28
WS (m s^{-1})	8	2.71	0.4	2.14	-0.4	0.92	7.1	3.28	0.8	1.93	-1.03	0.2
T (°C)	36	20.73	5	6.41	0.31	-0.3	42	21.17	7	7.96	1.38	0.85
RH (%)	86	55.42	31	10.82	0.32	0.64	85	54.02	42	9.03	2.97	1.31
WDEL (%)	39.6	15.92	2.4	9.13	-0.38	0.6	34	18.83	3.5	8.47	-1.14	0.08

Xm = maximum value, Xa = mean value, Xn = minimum value, Sx = standard deviation, Kx = kurtosis coefficient, Csx = skewness coefficient, Hr = riser height, P = operating pressure, dm = main nozzle diameter, da = auxiliary nozzle diameter, qm = water discharge by main nozzle, qa = water discharge by auxiliary nozzle, WS = wind speed, T = air temperature, RH = relative humidity, and WDEL = wind drift and evaporation losses.

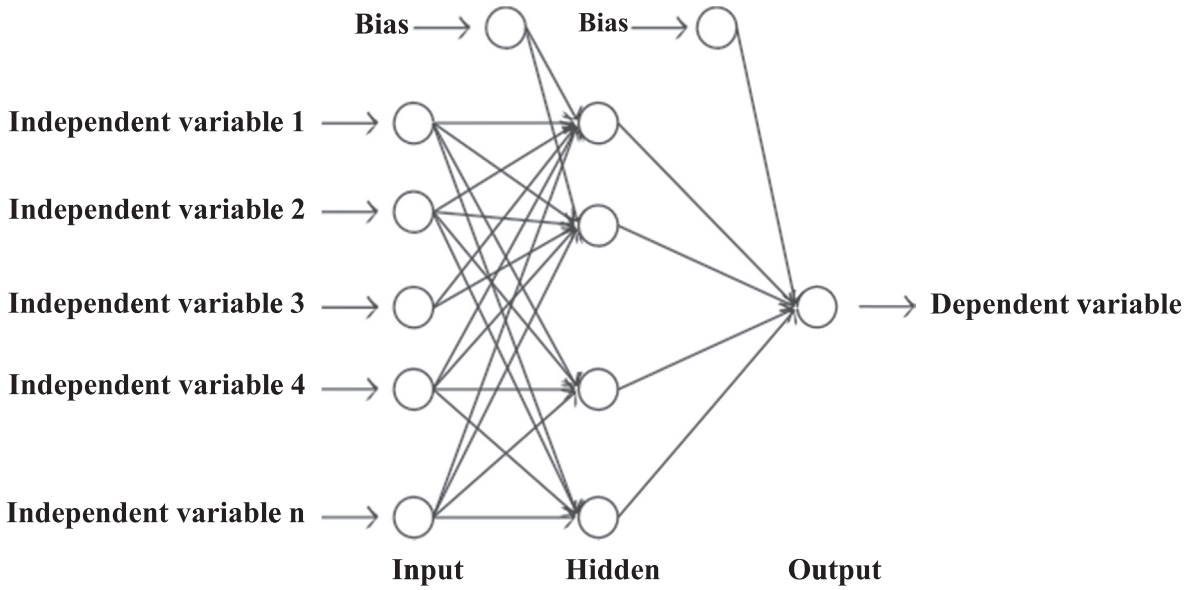


Fig. 1. A typical artificial neural network's structural diagram (Li et al., 2017).

Table 3
Tuning parameters used for each model.

Model	Tuning Parameters
ANN	Maximum number of iterations = 30,000 Learning rate = 0.01 Momentum = 0.8 Activation function = Tanh Number of hidden layers = 1 Number of neurons of the hidden layer = 6
ANFIS	Number of clusters: 5 Initial FIS: Fuzzy partition matrix exponent = 2 Maximum number of iterations = 200 Minimum improvement = $1e^{-5}$ ANFIS: Maximum number of epochs: 200 Error goal = 0 Initial step size = 0.01 Step size decrease rate = 0.9 Step size increase rate = 1.1
MARS	Number of basis functions at the forward pass = 100 Number of basis functions at the backward pass = 50 Minimum number of observations between the knots = 3 No penalty is added to the variables to give equal priority to all input variables
PLR	Prior parameters: $\alpha = 0.02$, $\beta = 0.5$ Maximum number of iterations = 1000 Tolerance = $1e^{-4}$
SVR	Kernel function = linear Box constraint = 1 Kernel scale = 1 Epsilon = 0.2

by adding bias (threshold value, B) in the central processing unit (Izadifar, 2010). The activation (transfer) functions (f) in the hidden and output neurons' processing units are then implemented. The hyperbolic tangent (tanh) function is the most often used transfer function in engineering applications (Dawson and Wilby, 1998; Zanetti et al., 2007). The neural network's tanh transfer functions in the hidden and output layers were applied in the study to train the network. The tanh transfer function's general functional form is:

$$f(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \tag{3}$$

The network's output can be expressed as follows (Haykin, 1999):

$$Y_k = f\left(\sum_{j=1}^{n_j} (W_2)_{kj} f\left(\sum_{i=1}^{n_i} (W_1)_{ji} X_i + (B_1)_j\right) + (B_2)_k\right) \tag{4}$$

where Y_k denotes the output; X_i is the input parameter with normalized value; $(W_1)_{ji}$ is the input layer's weight in relation to the hidden layer; $(W_2)_{kj}$ is the hidden layer's weight in relation to the output layer; n_i and n_j denote the input and output neuron numbers, respectively.; $(B_1)_j$ and $(B_2)_k$ denote the biases in the hidden and output layers, respectively.

At the end of each forward pass in the forward-propagation step, the differences (error) between the output and target values are determined. If the error exceeds the acceptable threshold limit, the weights are modified layer by layer (back-propagation step). Otherwise, the training is terminated. The optimal number of hidden neurons is determined using a trial-and-error approach (Jain et al., 2008). Following that, the trained ANN is tested on the unseen data set.

2.3. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The first step of developing an ANFIS based model involves constructing an initial Fuzzy Inference System (FIS) whose parameters are adjusted using a hybrid algorithm to develop the final ANFIS. A fuzzy c-mean clustering algorithm (FCM) (Bezdek et al., 1984) was used to develop the initial FIS. FCM is a valuable tool in compressing the dataset by dividing them into identical clusters that significantly reduce the number of linear and non-linear parameters of FIS. The number of clusters also determines the number of rules in the developed FIS (Table 3). A Sugeno-type FIS, also known as the Takagi-Sugeno-Kang model introduced in 1985 (Sugeno, 1985), is ideal for creating an initial FIS for ANFIS training. Sugeno FIS's computational framework is based on fuzzy logic theory, which integrates fuzzy set theory, fuzzy if-then rules, and fuzzy reasoning. The FIS's fundamental structure is built of three components: (i) a rule base composed of fuzzy if-then rules, (ii) a database that specifies the size, type, and membership functions (MF) number utilized in fuzzy rules, and (iii) a reasoning mechanism that performs inference (Jang et al., 1997). Using several of these fuzzy "if-then" rules, FISs can be employed to map nonlinear input and output spaces. A Sugeno type FIS's input and output membership functions are Gaussian and linear, respectively. A Gaussian MF defined by two parameters can be mathematically described as follows (Jang et al., 1997):

$$gaussian(x,c,\sigma) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \quad (5)$$

where c represents the center of MF and is the width of MF.

The present study utilizes a Sugeno-type FIS developed using the FCM algorithm as an initial FIS to develop the final ANFIS. The optimal number of clusters was determined by performing multiple trials with various clusters and analyzing the root mean square error (RMSE) of the prediction that resulted. The number of clusters that produce minimum RMSE and create the least difference in RMSE values between the training and testing datasets are chosen as adequate.

Sugeno-type ANFIS was produced by utilizing a hybrid algorithm to tune the parameters of the initial Sugeno FIS. Sugeno ANFIS structures are simpler in architecture than Mamdani ANFIS structures. However, they have superior learning capabilities compared to other ANFIS structures (Jang et al., 1997).

The fuzzy "if-then" rule set for a Sugeno first-order FIS with two inputs (α and β), one output (γ), and two fuzzy "if-then" rules are as follows:

$$Rule\ \ 1: If\ \ \alpha\ \ is\ \ P_1\ \ and\ \ \beta\ \ is\ \ Q_1\ \ then\ \ f_1 = p_1\alpha + q_1\beta + r_1 \quad (6)$$

$$Rule\ \ 2: If\ \ \alpha\ \ is\ \ P_2\ \ and\ \ \beta\ \ is\ \ Q_2\ \ then\ \ f_2 = p_2\alpha + q_2\beta + r_2 \quad (7)$$

This is illustrated graphically in Fig. 2, which depicts a Sugeno fuzzy inference system-based ANFIS architecture. Five layers comprise the proposed Sugeno ANFIS: "a fuzzy layer, a product layer, a normalized layer, a defuzzification layer, and a total output layer".

2.4. Multivariate Adaptive Regression Spline (MARS)

MARS is a method for non-parametric adaptive regression that enables the construction of flexible regression models (Friedman, 1991). MARS divides the total solution space into numerous input variable intervals and constructs a regression model by fitting individual splines or basis functions to each interval (Bera et al., 2006). MARS-based prediction models were developed to predict WDEL with nine predictors. The MARS uses both the forward and backward stepwise approaches to input-output mapping in order to forecast WDEL values. MARS utilizes the backward stepwise technique, which eliminates irrelevant input factors in generating the output variable to avoid unnecessary complexity and model overfitting. The proposed MARS model's input-output mapping is as follows (Roy and Datta, 2017):

$$BF_i(X) = max(0, X_j - \alpha) or$$

$$= max(0, \alpha - X_j) \quad (8)$$

$$Y = f(X) = \beta \pm \gamma_k \times BF_i(X) \quad (9)$$

where i and j are the indices for Basis functions and input variables, respectively; BF_i denotes i^{th} Basis functions; X_j represents j^{th}input variables; α is a constant known as knot; β is a constant value; γ_k denotes the corresponding coefficients of $BF_i(X)$.

The number of basic functions is chosen in accordance with the number of significant variables required to build a robust MARS model (Table 3). In this study, 100 forward steps were permitted during MARS model construction. MARS uses internal heuristics to select the minimal number of observations between knots based on sample size and model complexity. In the MARS forward-stepping methodology, no penalty is given to the variables, ensuring MARS can assign the same importance to all of the variables supplied into the model generation process. However, throughout the backward stepping process, MARS selects only the most pertinent input variables required to forecast the output variables. Model overfitting is minimized with this backward step, which makes the constructed model as simple as possible.

2.5. Probabilistic Linear Regression (PLR)

The PLR, also known as empirical Bayesian Linear regression, captures the input-output mapping for the predicted target variable. In PLR, the regression task is executed via the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) or the Mackay fixpoint iteration method (MacKay, 1992). For developing the PLR model, the present study adopts the EM algorithm (Table 3). The EM algorithm is relied on maximizing either the log-likelihood or log-posterior density functions. These functions are integrated with certain latent variables. During the optimization process, a lower bound, $\mathscr{F}_{EM}$ of the objective function, $l(\alpha)$ is implicitly constructed by the EM algorithm, such that:

$$\mathscr{F}_{EM}(q,\alpha) := \mathbb{E}_q\left[\log\frac{p(D|z)p(z|\alpha)}{q(z)}\right] = \int q(z)\log\frac{p(D|z)p(z|\alpha)}{q(z)}dz \quad (10)$$

where, $q()$ Refers to an arbitrary probability distribution in the space z. According to Dempster et al. (1977), the following are the well-known results in the literature of EM based on Jensen's Inequality theorem (Jensen, 1906).

Lemma 1. $\mathscr{F}_{EM}(q,\alpha) \leq l(\alpha)$, in which the equality is attained if and only if $q(z)$ is the posterior distribution $p(z|D,\propto)$.

Now an alternative optimization problem is defined to be optimized instead of optimizing the original objective function $l(\alpha)$. This alternate optimization problem is formulated as:

AltOptEM : Find $\alpha$and$a$distribution$q$that maximize $\mathscr{F}_{EM}(q,\alpha)$

The EM algorithm will then act as the coordinate ascent solver for the *AltOptEM*. More specifically, the updating of the rule at the *ith* iteration of the coordinate ascent can be represented as follows:

$$\text{E} - \text{Step}: q^{(t)} := \text{argmax}_q \mathscr{F}_{EM}\left(q, \alpha^{(t)}\right) = p\left(z|D, \propto^{(t)}\right) \quad (11)$$

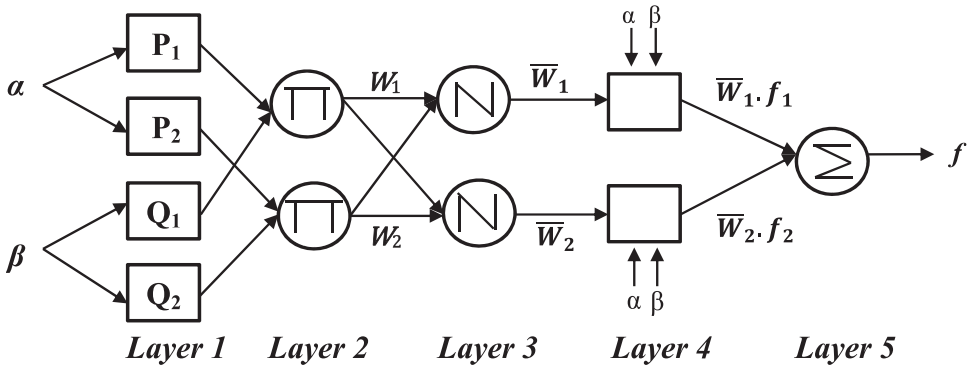


Fig. 2. ANFIS architecture based on a Sugeno FIS with two inputs (Jang, 1993).

$$\mathrm{M-Step}: \propto^{(t+1)} := \mathrm{argmax}_{\propto} \mathscr{T}_{EM}\left(q^{(t)}, \propto\right) = \mathrm{argmax}_{\propto} \mathbb{E}_{q^{(t)}}[\log \quad p(z|\propto)] \tag{12}$$

It is noted that by Lemma 1, at the end of E-step,

$$\mathscr{T}_{EM}\left(q^{(t)}, \alpha^{(t)}\right) = l\left(\propto^{(t)}\right) \tag{13}$$

This provides rise to the following Lemma (Dempster et al., 1977):

Lemma 2. *The iteration of the EM algorithm continuously increases the log-evidence function* $l(\propto)$ and therefore is guaranteed to converge.

2.6. Support Vector Regression (SVR)

The SVR is a nonparametric modeling approach that is capable of building robust prediction tools. This approach relies primarily on kernel functions in developing the prediction models. For supervised learning, the SVR approach uses an asymmetrical loss function. The main advantage of employing the SVR approach for prediction purposes is that the computational efficiency is not affected by the dimensionality of the input space (Awad and Khanna, 2015). The detailed description of SVR was presented in Vapnik (1995) and Gunn (1997). A brief outline of SVR is presented in this effort. For a regression problem, a training dataset can be given as:

$$P = \{(a_1, b_1), (a_2, b_2), (a_3, b_3), \ldots., (a_N, b_N)\} \tag{14}$$

where, $a_i (i = 1, 2, 3, \ldots., N)$ represents a vector consisting of real independent variables; $b_i (i = 1, 2, 3, \ldots., N)$ is the associated scalar real independent variable. For the training dataset, the regression equation in the feature space can be represented by:

$$z(a, w) = (w \bullet \varnothing(\mathrm{a}) + \mathrm{c}) \tag{15}$$

where, w states the weight vector; c indicates a constant; $\varnothing(a)$ represents the feature function; and $w \bullet \varnothing(\mathrm{a})$ is the dot product. SVR works by minimizing the following function:

$$\mathrm{Minimize}: Q(f) = C\frac{1}{N}L_{\varepsilon}(b, z(a, w)) + \frac{1}{2}\left\|W^2\right\| \tag{16}$$

And,

$$L_{\varepsilon}(b, z(a, w)) = \left\{ \begin{array}{cc} 0 & \mathrm{if}|b - z(a, w)| \le \varepsilon \\ |b - z(a, w)| - \varepsilon & \mathrm{otherwise} \end{array} \right\} \tag{17}$$

The left-hand side term of Eq. (16) denotes the empirical error. In contrast, the term C indicates the degree to which the empirical error is optimized in relation to the model complexity given by the second component of Eq. (16). Eq. (17) describes a loss function names ε-insensitive loss function (Vapnik et al., 1996). Through the incorporation of Lagrangian multipliers β and β^*, the optimization problem defined in Eq. (16) is converted into the dual problem. Only the non-zero coefficients, in combination with their input vectors, a_i are termed as the support vectors. The final form of the equation comes out as one presented below:

$$z\left(a, \beta_i, \beta_i^*\right) = \sum_{i=1}^{N_{sv}}\left(\beta_i - \beta_i^*\right)\left(\varnothing(a_i) \bullet \varnothing\left(\mathrm{a_j}\right)\right) + \mathrm{c} \tag{18}$$

Using the kernel function $K\left(x_i, x_j\right)$, the SVR function can be written as:

$$z\left(a, \beta_i, \beta_i^*\right) = \sum_{i=1}^{N_{sv}}\left(\beta_i - \beta_i^*\right)K(a, a_i) + c \tag{19}$$

The Karush-Kuhn-Tucker condition is applied to compute the term c in Eqs. (18) and (19). The cost function C, the radius of the insensitive tube ε and the kernel parameters $K\left(x_i, x_j\right)$ are considered as the most important parameters that control regression problems utilizing the SVR approach (Table 3).

2.7. Entropy weight calculation

Machine learning-based prediction models provide inconsistent prediction performances time and again when different performance evaluation indices are employed. In other words, for instance, the performance of the model improves when Wilmott's index of agreement (IA) is taken into consideration. In comparison, other models might provide higher performance with respect to the RMSE. As a result, several performance indices should be generated on the test dataset to be used within a decision theory framework to determine the best-performing model. Four benefit indices (the higher, the better) and four cost indices (the lower, the better) were utilized in this study to rank the built prediction models' performances. Shannon's entropy was used to determine the ranking.

The prediction models and performance evaluation indices were combined to generate a decision matrix. For instance, if there are m prediction models and l performance evaluation indices, the decision matrix can be stated mathematically as (Wu et al., 2011):

$$WDEL_{ij} = \left[\begin{array}{cccc} WDEL_{11} & WDEL_{21} & \cdots & WDEL_{m1} \\ WDEL_{12} & WDEL_{22} & \cdots & WDEL_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ WDEL_{1l} & WDEL_{2l} & \cdots & WDEL_{ml} \end{array} \right] \tag{20}$$

The decision matrix was then normalized to minimize the index dimensionality effect. The performance indices were normalized to values between 0 and 1such that $S_{ij} \in [0, 1], i = 1, 2, \ldots, m; j = 1, 2, \ldots, l$. The standardized matrix, S_{ij} is expressed as (Wu et al., 2011):

$$S_{ij} = \left\{ \begin{array}{c} \dfrac{WDEL_{ij}}{\max(WDEL_{i1}, WDEL_{i2}, \ldots, WDEL_{il})}, \quad for \quad benefit \quad indexes \\ \dfrac{X_{ij}}{\min(WDEL_{i1}, WDEL_{i2}, \ldots, WDEL_{il})}, \quad for \quad cost \quad indexes \end{array} \right\} \tag{21}$$

The entropy-based ranking was used to allocate entropy weights to individual prediction models. The ranking procedure was as follows (Li et al., 2011):

Step1: Calculation of the entropy value of each index using Shannon's information entropy principles. Entropy value of the j^{th} index was calculated as:

$$Entropy_j = -k\sum_{i=1}^{m} f_{ij} \ln f_{ij} \tag{22}$$

where,

$$f_{ij} = S_{ij} \Big/ \sum_{i=1}^{m} S_{ij} \tag{23}$$

$$k = 1/\ln m \tag{24}$$

Step2: Calculation of each index's entropy weight. j^{th} index's entropy weight was calculated by:

$$w(entropy)_j = \frac{1 - Entropy_j}{l - \sum_{j=1}^{l} Entropy_j} \tag{25}$$

This entropy weight reflects the index's importance in the decision-making process. The greater the entropy-based weight, the more information is contained in the index, and the more essential the index is in decision-making.

Step3: The ranking weight for each model is calculated by multiplying the entropy weight of each index by its standardized value. A mathematical representation of this step is:

$$w(entropy)_i = \sum_{j=1}^{l} S_{ij} \times w(entropy)_j \tag{26}$$

Step4: Determination of model ranking

$$\max \quad [w(entropy)_i], \ldots, \min[w(entropy)_i]; \text{for} i = 1, 2, \ldots, \quad m \tag{27}$$

Step5: Calculation of entropy weight for individual prediction models

$$W(entropy)_i = w(entropy)_i \Big/ \sum_{i=1}^{m} w(entropy)_i \tag{28}$$

2.8. Statistical performance criteria

Five statistical criteria were used to evaluate the accuracy of the five proposed models (ANN, ANFIS, MARS, PLR, and SVR). These statistical criteria are correlation coefficient (R), IA, Kling-Gupta efficiency (KGE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute relative error (MARE). Following are the equations that determine the statistical criteria:

$$R = \frac{\sum_{i=1}^{n}\left(WDEL_{i,a} - \overline{WDEL_a}\right)\left(WDEL_{i,a} - \overline{WDEL_p}\right)}{\sqrt{\sum_{i=1}^{n}\left(WDEL_{i,a} - \overline{WDEL_a}\right)^2}\sqrt{\sum_{i=1}^{n}\left(WDEL_{i,p} - \overline{WDEL_p}\right)^2}} \tag{29}$$

$$IA = 1 - \frac{\sum_{i=1}^{n}\left(WDEL_{i,a} - WDEL_{i,p}\right)^2}{\sum_{i=1}^{n}\left(\left|WDEL_{i,p} - \overline{WDEL_a}\right| + \left|WDEL_{i,a} - \overline{WDEL_a}\right|\right)^2} \tag{30}$$

$$KGE = 1 - \sqrt{(R-1)^2 + \left(\frac{pd}{ad} - 1\right)^2 + \left(\frac{pm}{am} - 1\right)^2} \tag{31}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(WDEL_{i,a} - WDEL_{i,p}\right)^2} \tag{32}$$

$$MAE = \frac{1}{n} \quad \left[\left|WDEL_{i,a} - WDEL_{i,p}\right|\right] \tag{33}$$

$$MARE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{WDEL_{i,a} - WDEL_{i,p}}{WDEL_{i,a}}\right| \tag{34}$$

where $WDEL_{i,a}$ and $WDEL_{i,p}$ are *WDEL* values at the i^{th} step obtained through experiments (actual values) and prediction models, respectively. *n* is the number of the time steps. $\overline{WDEL_a}$ is the mean value of the actual *WDEL*, *n* is the number of data points, *pd* is the standard deviation of predicted *WDEL* values, *ad* is the standard deviation of actual *WDEL* values, *pm* is average of predicted *WDEL* values, and *am* is average of predicted *WDEL* values.

3. Results and discussion

3.1. Development of predication models

To determine the optimal ANN topology, statistical parameters were used to analyze multiple networks with varying numbers of neurons in the hidden layer. With the addition of more hidden neurons for the training process, the ANN architectures significantly improved. Initially, the R, RMSE, and MAE values for five neurons in the hidden layer were 0.945, 2.962, and 2.457, respectively. By increasing the number of hidden neurons from five to six, the network performance was improved slightly. The R-value was enhanced by 1.16%, while the RMSE and MAE values were reduced by 10.13% and 10.58%, respectively. Following that, raising the number to six resulted in a degraded performance of the ANN. Thus, the optimal ANN model included an input layer with nine nodes (nine inputs), a hidden layer with six neurons, and an output layer with a single node (WDEL). The optimal network's R, IA, KGE, RMSE, MAE, and MARE were computed as 0.956%, 0.978%, 0.925%, 2.662%, 2.197%, and 22.29%, respectively (Table 4). Fig. 3 clearly illustrates that the WDEL values followed a straight line during the training process, demonstrating that the model's predicted and experimental data were in extremely close agreement. The fitted line equation for the ANN model indicated that the slope (α_o) and intercept (α_1) approached one and zero, respectively, (assuming that the equation is $y = \alpha_o x + \alpha_1$).

To determine the optimal ANFIS network, a hybrid-learning approach was used to characterize the link between the input variables and output. This strategy combines both the backpropagation and least-squares methods to quickly train and adapt the FIS to determine the optimal MF distribution. To design the optimal ANFIS network, the Gaussian MF was employed and assessed. Experimentally, multiple models were developed, and their associated parameters were studied to calculate the number of clusters. Two to five clusters were the most effective range for training and testing. Once the differences in absolute errors between training and testing processes were minimized, the ANFIS model was chosen. The ANFIS's optimal results are summarized in Table 4. For the training process, R, IA, KGE, RMSE, MAE, and MARE were calculated as 0.871%, 0.927%, 0.817%, 4.462%, 3.417%, and 32.24%, respectively. According to Fig. 3, the values of α_o and α_1 for the fitted line equation were 0.758 and 3.853, respectively.

The MARS model was developed using the most significant inputs in the WDEL. As a result, several combinations of related input variables were concurrently assessed to determine the WDEL values. The predicted WDEL was straightforward in which based on the MARS model's regression equations. MARS learns and extracts knowledge from the dataset while generating the regression equations and then presents an explicit relationship between the input and output variables. The accuracy of the MARS model was determined by its performance with the training and testing datasets during the training phase. As shown in Table 4, the model's performance was found to be satisfactory with the statistical indices. For the training dataset, MARS yielded R = 0.858, IA = 0.919, KGE = 0.799, RMSE = 4.658, MAE = 3.387, and MARE = 34.75%. Results presented in Table 4 show that the training and testing statistics indices were not significantly different. This demonstrates that the MARS model was not considerably overfitting during training development. The scatter plots in Fig. 3 indicated that the values of α_o and α_1 for the fitted line equation are 0.736 and 4.197, respectively.

Table 4 summarizes the PLR model's statistical parameters. Additionally, Fig. 3 depicts the PLR model's predicted and observed values. The PLR model had the lowest R (0.822), IA (0.891), and KGE (0.729) and the highest RMSE (5.169), MAE (4.108), and MARE (40.17%) during the training process. As a result, the PLR model was the worst for the study's WDEL estimation. In addition, the fit line equation yielded the lowest value of αo (0.599).

To improve the accuracy of the WDEL prediction of the SVR model, the parameters of the SVR were optimized in the training mode. The training R, RMSE, and MAE were used to set the ε and the C for the selected kernel function and kernel parameter. The SVR model's predicted WDEL values were found to differ from their observed values. The R, IA, KGE, RMSE, MAE, and MARE of the SVM results were 0.857%, 0.914%, 0.775%, 4.716%, 3.505%, and 31.39% during the training process (Table 4). In Fig. 3, the line's equation represents the observed and expected WDEL values. The αo and α1 are line's equation parameters, with αo being 0.712 and α1 being 3.965.

3.2. Comparison between prediction models

The comparison of the predicted WDEL values by the developed ANN, ANFIS, MARS, PLR, and SVR models to the observed values using the training and testing datasets is shown in Figs. 3 and 4. In addition, Table 4 summarizes the statistical criteria used to evaluate each model

Table 4
Statistical performance of the prediction models during the training and testing processes.

Model	Training process						Testing process					
	R	IA	KGE	RMSE	MAE	MARE	R	IA	KGE	RMSE	MAE	MARE
ANN	0.956	0.978	0.925	2.662	2.197	22.29	0.861	0.925	0.850	4.356	3.490	24.15
ANFIS	0.871	0.927	0.817	4.462	3.417	32.24	0.776	0.860	0.678	5.361	4.535	29.68
MARS	0.858	0.919	0.799	4.658	3.387	34.75	0.813	0.897	0.793	5.000	4.008	25.76
PLR	0.822	0.891	0.729	5.169	4.108	40.17	0.769	0.859	0.678	5.381	4.368	27.00
SVR	0.857	0.914	0.775	4.716	3.505	31.39	0.749	0.831	0.637	5.876	4.603	26.00

during training and testing datasets. As shown in Fig. 3, the α_o and α_1 values of the fitted line and the scatter plots' visual inspection indicated that the ANN model's WDEL estimations were closer to the observed values than the other models' estimations. The α_o values for the ANFIS, MARS, PLR, and SVR models were 15.64%, 18.05%, 33.36%, and 20.81%, respectively, lower than those for the ANN model. The α_1, on the other hand, increased by factors of 1.47, 1.69, 3.36, and 1.54, respectively. Additionally, R, IA, and KGE had the highest ANN statistical parameters values, whilst RMSE, MAE, and MARE had the lowest values (Table 3). Modeling the WDEL using the developed ANFIS, MARS, PLR, and SVR resulted in lower R, IA, and KGE values and higher RMSE, MAE, and MARE values than the ANN model during the training dataset. These four models, ANFIS, MARS, PLR, and SVR were each able to reduce R, IA, and KGE values by (8.97%, 10.25%, 14%, and 10.32%), (5.2%, 6.04%, 8.89%, and 6.51%), and (11.68%, 13.62%, 21.19%, and 16.22%), respectively, compared to the results of the ANN model. Additionally, the RMSE was increased to 67.6%, 74.94%, 94.14%, and 77.13, respectively, using the developed ANFIS, MARS, PLR, and SVR models. Similarly, the MAE was enhanced to 55.55%, 54.16%, 86.97%, and 59.54%, respectively. MARE was also raised to 44.63%, 55.91%, 80.19% and 40.82%, respectively. In terms of prediction accuracy, the ANFIS model came in second, followed by the MARS, SVR, and PLR models, using R, IA, KGE, and RMSE. The R-values for the ANFIS model were around 1.46%, 5.89%, and 1.54% higher than those for the MARS, PLR, and SVR models, respectively. Other statistical parameters in the ANFIS model outperformed their values in the MARS, PLR, and SVR models, except for MAE in the MARS model and MARE in the SVR model that had the lowest values. The PLR model had the lowest α_o (0.599) and the highest α_1 (6.808) values, showing that there was significantly less scatter and more clustering around the 1:1 line during the training process (Fig. 3). Table 4 demonstrates that the PLR model had the lowest R, IA, and KGE and highest RMSE, MAE, and MARE values, showing that the PLR model provided slightly inaccurate estimates.

During the testing phase (Fig. 4), the ANN model predicted WDEL values that were more tightly clustered around the 1:1 line than the other models. Mainly due to the ANN model's α_o (0.821) and α_1 (3.992) values which were closer to one and zero, respectively, compared to the other prediction models' values. For ANFIS, MARS, PLR, and SVR models, the α_o values varied between 0.599 and 0.748, whereas the α_1 values ranged between 4.008 and 6.808. The ANN model's results for performance statistics provided the best overall results (Table 4), having the highest R (0.861), IA (0.925), and KGE (0.85) values. In contrast, the RMSE (4.356%), MAE (3.49%), and MARE (24.15%) were the lowest. Clearly, the R, IA, and KGE values for the MARS model decreased by 5.63%, 3.05%, and 6.69%, respectively, compared to the ANN model. RMSE, MAE, and MARE values, on the other hand, increased by 14.79%, 14.82%, and 6.69%, respectively. The ANFIS model came in the third rank with slightly higher R, IA, and KGE values of 0.81%, 0.05%, and 0.07%, respectively, compared to the PLR model. On the other hand, the ANFIS model had a lower RMSE value of 0.38% than the PLR models. In accordance with MAE and MARE, the PLR model had better performance than the ANFIS model. The R, AI, and KGE values for the SVM model were the lowest, while the RMSE and MAE values were the highest.

The Whisker boxplots in Fig. 5 show the absolute error distribution associated with the projected values for the developed models. The range of predicted error values was demarcated by quartiles of the absolute errors of prediction. The lower quartile (25th percentile) represents the boxplot's lower bound, while the upper quartile represents the boxplot's top-bound (75th percentile). The horizontal line within each box shows the median of the prediction model's absolute errors. The '×' symbols indicate the average value of the absolute errors. The lower Whisker extends from the box bottom to the lowest non-outlier in the absolute error dataset, while the upper Whisker extends from the box top to the largest non-outlier. From Fig. 5, it appears that the ANN model had the smallest extent of errors, while the SVR model had the largest ones. As a result, the ANN model performed better than the other models, with the fewest mean and median absolute errors. Additionally, the ANN model had the smallest lower quartile value, followed by the MARS model. ANFIS's upper quartile value was ranked third, following the ANN and MARS models. The SVR model generated the maximum mean value of absolute errors, whereas the PLR produced the maximum median value of absolute errors.

The models' performances were found to be inconsistent across a variety of performance metrics. As a result, selecting the optimal model was difficult to use these criteria. Therefore, the Shannon's entropy-based decision theory method, which is widely utilized (Zou et al., 2006; Liu et al., 2010), was used to rank the models for WDEL predicting (Table 5). The major advantage of this method is the avoidance of human influences interfering with indicator weighting, which improves the objectivity of the outcomes of the overall assessment (Taheriyoun et al., 2010; Ding et al., 2017). As a result, the method has become increasingly popular in recent years as a decision-making tool (Wu et al., 2015; Yan et al., 2016; Feng et al., 2019). The predicted outcomes were prioritized according to the weight assignments that were applied to the processes to manage the contradictory performances across the performance criteria. This method revealed that the ANN model with the highest weight (0.999) was the best performing, followed by MARS, ANFIS, and so on. The SVR, on the other hand, was found to be the least ranked model.

3.3. Contribution analysis

Due to the superior predictive ability of the ANN model, ANN was utilized to study the contribution of the inputs (i.e., H_r, P, d_a, d_m, q_a, q_m, WS, RH, and T) that were utilized to model the WDEL. The connection weights acquired from the ANN architecture were used to generate the contribution ratios for each input (training). As a result of this technique, it was possible to determine which inputs were most effective at generating the model's output predictions. Fig. 6 illustrates WDEL prediction's contribution ratio values for each input variable. As illustrated in this Figure, d_a (17.5%) and H_r (16.3%) were the most influential factors, while q_m and WS came in third and fourth, with a contribution ratio of 12.57% and 12.19%, respectively. The T, RH, and P had a moderate effect, with their contribution ratio values of 9.96%, 9.02%, and 8.99% respectively. According to Playán et al. (2005), Al-Ghobari et al. (2018) and Sarwar et al. (2019), T and RH are contributing variables in WDEL under a variety of technical and operational conditions. Molle et al. (2012) demonstrate that droplet evaporation is T and RH dependent, with evaporation occurring as a result of the difference

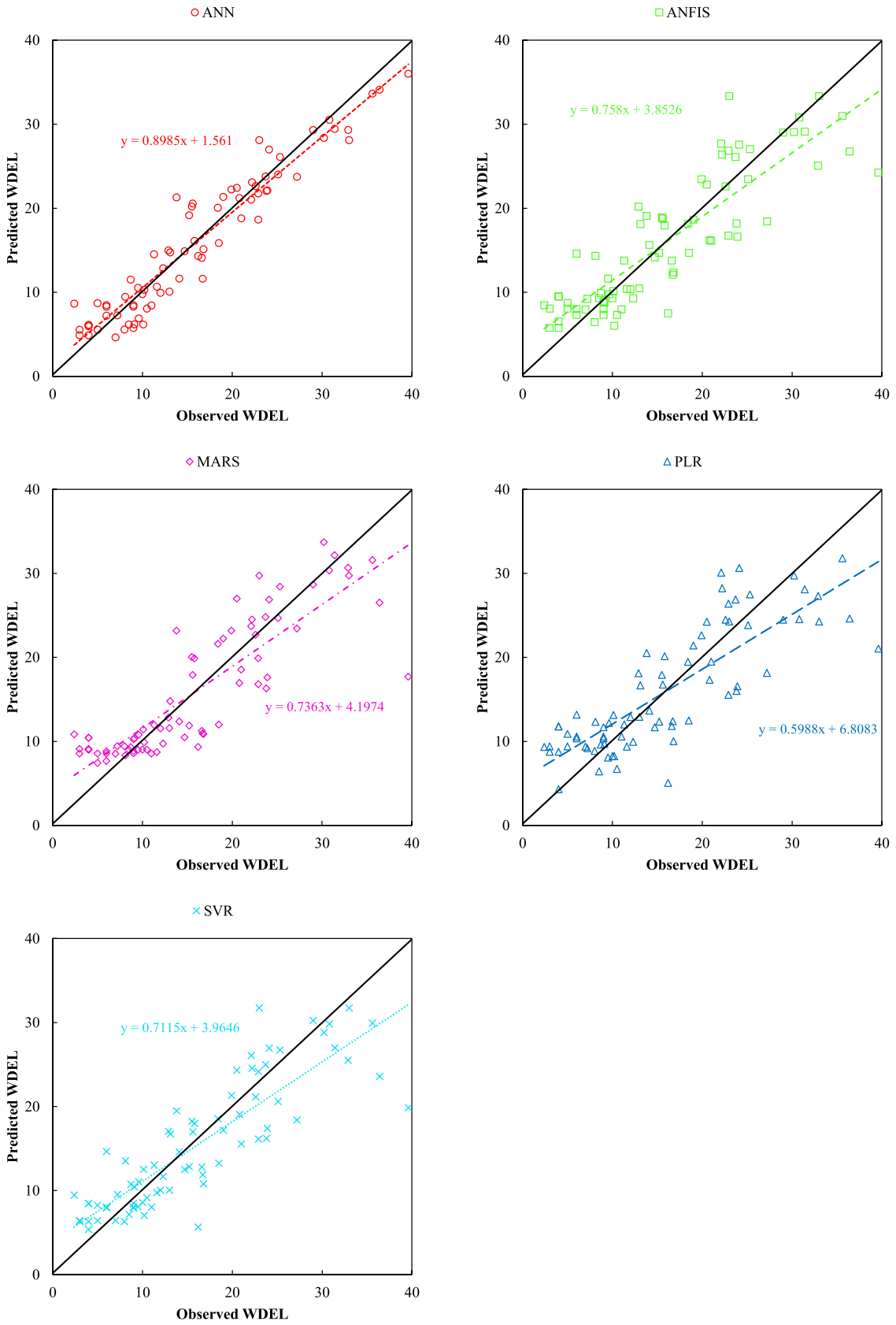


Fig. 3. Scatter plot of observed and predicted WDEL values for models built throughout the training process.

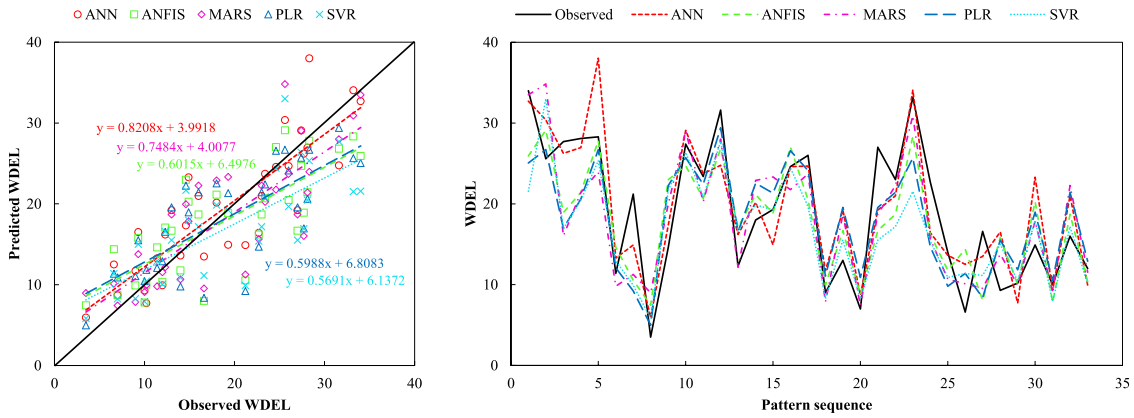


Fig. 4. A comparison of the observed and predicted WDEL values for the developed models throughout the testing process.

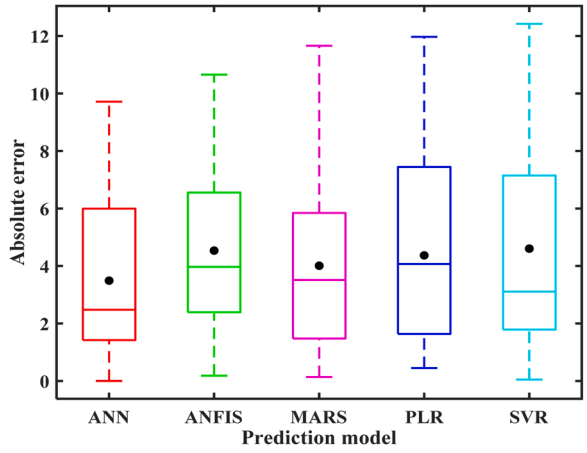


Fig. 5. Boxplot for absolute error in predicted WDEL with the predication models during the testing process.

Table 5
Entropy weights and ranking of prediction models.

Model	Weights	Ranking
ANN	0.999	1
ANFIS	0.865	3
MARS	0.896	2
PLR	0.833	4
SVR	0.794	5

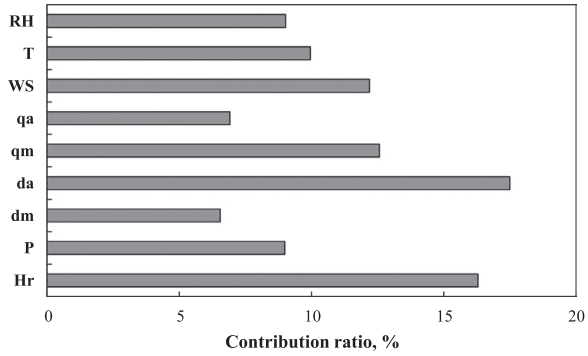


Fig. 6. Contribution ratio of the ANN model's input variables.

between the actual and saturation water vapor pressures at a certain ambient air temperature (i.e., vapor pressure deficit). When the surrounding air has a high RH, the droplets evaporate more slowly due to the low vapor pressure deficit, resulting in inadequate energy to convert the droplet's sensible heat to latent heat.

4. Conclusion

Wind drift and evaporation losses (WDEL) are the wide range of sprinkler water application losses that are the most important factors required for irrigation system management and design. The artificial neural network (ANN), the adaptive neuro-fuzzy inference system (ANFIS), the multivariate adaptive regression spline (MARS), the probabilistic linear regression (PLR), and the support vector regression (SVR) were used to develop models that could predict of WDEL using design, operational, and meteorological variables in sprinkler irrigation. These models were trained and tested using datasets from climates and sprinklers published by other researchers. According to the statistical evaluation criteria and Shannon's entropy-based decision theory, the ANN model outperformed the other models in terms of accuracy. The SVR model produced high error, implying that the prediction values were less accurate. In addition, the MARS model ranked second, followed by ANFIS and PLR models. The auxiliary nozzle diameter of sprinkler and wind speed were dominating design and climate variables, producing a robust predictive model. Incorporating the air temperature and relative humidity into the modeling process via the vapor pressure deficit is desirable. In brief, ANN was a reliable and powerful prediction tool for modeling WDEL in the sprinkler irrigation system. Using the developed ANN model, irrigation system managers could adjust various field operations to maximize water efficiency during the growth season, lowering WDEL in various meteorological situations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was financially supported by the Vice Deanship of Research Chairs at King Saud University.

References

Abo-Ghobar, H.M., 1993. Evaporation and drift losses from sprinkler irrigation systems under hot and dry conditions. J. King. Saud. Univ. Agric. Sci. 5 (2), 153–164.
Adamowski, J., Chan, H.F., Prasher, S.O., Sharda, V.N., 2012. Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial

neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. J. Hydroinform. 14 (3), 731–744.
Al-Ghobari, H.M., El-Marazky, M.S., Dewidar, A.Z., Mattar, M.A., 2018. Prediction of wind drift and evaporation losses from sprinkler irrigation using neural network and multiple regression techniques. Agric. Water Manag. 195, 211–221.
Awad, M., Khanna, R., 2015. Support vector regression. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers. Apress, Berkeley, CA, pp. 67–80 ed.
Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. J. Microbiol. Methods 43, 3–31.
Bera, P., Prasher, S.O., Patel, R.M., Madani, A., Lacroix, R., Gaynor, J.D., Tan, C.S., Kim, S.H., 2006. Application of MARS in simulating pesticide concentrations in soil. Trans. ASABE 49 (1), 297–307.
Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: the fuzzy c-means clustering algorithm. Comput. Geosci. 10, 191–203.
Carrión, P., Tarjuelo, J., Montero, J., 2001. SIRIAS: a simulation model for sprinkler irrigation. Irrig. Sci. 20 (2), 73–84.
Chang, F.J., Chang, L.C., Kao, H.S., Wu, G.R., 2010. Assessing the effort of meteorological variables for evaporation estimation by self-organizing map neural network. J. Hydrol. 384 (1–2), 118–129.
Christiansen, J.E., 1942. Irrigation by sprinkling. Univ Calif Agric Exp Stn Bull 670.
Ciğizoğlu, H.K., 2003. Estimation, forecasting and extrapolation of river flows by artificial neural networks. Hydrol. Sci. J. 48 (3), 349–361.
Cigizoglu, H.K., Kisi, Ö., 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol. 317 (3–4), 221–238.
Dawson, W.C., Wilby, R., 1998. An artificial neural networks approach to rainfall runoff modeling. Hydrol. Sci. J. 43 (1), 47–66.
Dechmi, F., Playán, E., Cavero, J., Faci, J.M., Martínez-Cob, A., 2003. Wind effects on solid set sprinkler irrigation depth and yield of maize (Zea mays). Irrig. Sci. 22 (2), 67–77.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B (Methodol.) 39, 1–38.
Jensen, J.L.W.V., 1906. Sur les fonctions bconvexes et les ingalits entre les valeurs moyennes. Acta Math. 30 (1), 175–193.
Ding, X., Chong, X., Bao, Z., Xue, Y., Zhang, S., 2017. Fuzzy comprehensive assessment method based on the entropy weight method and its application in the water environmental safety evaluation of the Heshangshan drinking water source area, Three Gorges Reservoir area, China. Water 9 (5), 329.
Dogan, E., Ates, A., Yilmaz, E.C., Eren, B., 2008. Application of artificial neural networks to estimate wastewater treatment plant inlet biochemical oxygen demand. Environ. Prog. 27, 439–446.
Emamgholizadeh, S., Mohammadi, B., 2021. New hybrid nature-based algorithm to integration support vector machine for prediction of soil cation exchange capacity. Soft Comput. 25 (21), 13451–13464.
Feng, Y., Bao, Q., Xiao, X., Lin, M., 2019. Geo-accumulation vector miodel for evaluating the heavy metal pollution in the sediments of Western Dongting Lake. J. Hydrol. 573, 40–48.
Fernando, A., Shamseldin, A.Y., 2009. Investigation of the internal functioning of the radial basis function neural network river flow forecasting models. J. Hydrol. Eng. 14 (3), 1–7.
Friedman, J.H., 1991. Multivariate adaptive regression splines (with Discussion). Ann. Stat. 19 (1), 1–67.
Friedman, J.H., Roosen, C.B., 1995. An introduction to multivariate adaptive regression splines. Stat. Methods Med. Res. 4 (3), 197–217.
Gunn, S., 1997. Support Vector Machines for Classification and Regression. University of Southampton, pp. 1–42. ISIS Technical Report.
Hamdia, K.M., Lahmer, T., Nguyen-Thoi, T., Rabczuk, T., 2015. Predicting the fracture toughness of PNCs: a stochastic approach based on ANN and ANFIS. Comput. Mater. Sci. 102, 304–313.
Haykin, S., 1999. Neural Networks. A Comprehensive Foundation. Prentice Hall International Inc., New Jersey.
Huang, H., Ji, X., Xia, F., Huang, S., Shang, X., Chen, H., Zhang, M., Dahlgren, R.A., Mei, K., 2019. Multivariate adaptive regression splines for estimating riverine constituent concentrations. Hydrol. Process. 34 (5), 1–15.
Ismail, S., Shabri, A., Samsudin, R., 2012. A hybrid model of self organizing maps and least square support vector machine for river flow forecasting. Hydrol. Earth Syst. Sci. 16 (11), 4417–4433.
Izadifar, Z., 2010. Modelling and Analysis of Actual Evapotranspiration Using DataDriven and Wavelet Techniques. University of Saskatchewan, Saskatoon, Saskatchewan, Canada. Master Thesis.
Jain, A., Srinivasulu, S., 2004. Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms, and artificial neural network techniques. Water Resour. Res. 40 (4), W04302.
Jain, S.K., Nayak, P.C., Sudhir, K.P., 2008. Models for estimating evapotranspiration using artificial neural networks, and their physical interpretation. Hydrol. Process. 22 (13), 2225–2234.
Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. 23, 665–685.
Jang, J.-S.R., Sun, C.T., Mizutani, E., 1997. Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. Prentice Hall, Upper Saddle River, New Jersey.
Kalra, A., Li, L., Li, X., Ahmad, S., 2013. Improving streamflow forecast lead time using oceanic-atmospheric oscillations for Kaidu river basin, Xinjiang, china. J. Hydrol. Eng. 18 (8), 1031–1040.
Keller, J., Bliesner, R.D., 1990. Sprinkler and Trickle Irrigation. Van Nostrand Reinhold, New York.
King, B.A., Dungan, R.S., Bjorneberg, D.L., 2012, Evaluation of center pivot sprinkler wind drift and evaporation loss. Presented at the American Society of Agricultural and Biological Engineers Annual International Meeting 2012, American Society of Agricultural and Biological Engineers, Dallas, Texas, USA. Paper No. 121336891.
Kişi, Ö., 2006. Daily pan evaporation modelling using a neuro-fuzzy computing technique. J. Hydrol. 329 (3–4), 636–646.
Kumar, D.N., Raju, K.S., Sathish, T., 2004. River flow forecasting using recurrent neural networks. Water Resour. Manag. 18 (2), 143–161.
Kuzman, B., Petković, B., Denić, N., Petković, D., Ćirković, B., Stojanović, J., Milić, M., 2021. Estimation of optimal fertilizers for optimal crop yield by adaptive neuro fuzzy logic. Rhizosphere 18, 100358.
Lakovic, N., Khan, A., Petković, B., Petkovic, D., Kuzman, B., Resic, S., Jermsittiparsert, K., Azam, S., 2021. Management of higher heating value sensitivity of biomass by hybrid learning technique. Biomass Convers. Bioref. https://doi.org/10.1007/s13399-020-01223-w.
Li, H., Liu, Z., Liu, K., Zhang, Z., 2017. Predictive power of machine learning for optimizing solar water heater performance: the potential application of high-throughput screening. Int. J. Photoenergy, 4194251.
Li, Y., Bai, G., Yan, H., 2015. Development and validation of a modified model to simulate the sprinkler water distribution. Comput. Electron. Agric. 111, 38–47.
Li, X., Wang, K., Liu, L., Xin, J., Yang, H., Gao, C., 2011. Application of the entropy weight and TOPSIS method in safety evaluation of coal mines. Procedia Eng. 26, 2085–2091.
Liu, L., Zhou, J., An, X., Zhang, Y., Yang, L., 2010. Using fuzzy theory and information entropy for water quality assessment in Three Gorges region, China. Expert Syst. Appl. 37 (3), 2517–2521.
MacKay, D.J.C., 1992. The evidence framework applied to classification networks. Neural Comput. 4, 720–736.
Maroufpoor, E., Sanikhani, H., Emamgholizadeh, S., Kişi, Ö., 2018. Estimation of wind drift and evaporation losses from sprinkler irrigation systems by different data-driven methods. Irrig. Drain. 67, 222–232.
Mattar, M.A., 2018. Using gene expression programming in monthly reference evapotranspiration modeling: a case study in Egypt. Agric. Water Manage. 198, 28–38.
Mattar, M.A., Alamoud, A.I., Al-Othman, A.A., Elansary, H.O., Farah, A.H., 2020. Hydraulic performance of labyrinth-channel emitters: experimental study, ANN, and GEP modeling. Irrig. Sci. 38, 1–16.
Mohamed, A.Z., Peters, R.T., Sarwar, A., Molaei, B., McMoran, D., 2021. Impact of the intermittency movement of center pivots on irrigation uniformity. Water 13, 1167.
Montero, J., Tarjuelo, J.M., Carrión, P., 2003. Sprinkler droplet size distribution measured with an optical spectropluviometer. Irrig. Sci. 22 (2), 47–56.
Molaei, B., Peters, R.T., Mohamed, A.Z., Sarwar, A., 2021. Large scale evaluation of a LEPA/LESA system compared with MESA on spearmint and peppermint. Ind. Crops Prod. 159, 113048.
Molle, B., Tomas, S., Hendawi, M., Granier, J., 2012. Evaporation and wind drift losses during sprinkler irrigation influenced by droplet size distribution. Irrig. Drain. 61, 240–250.
Nasseri, M., Tavakol-Davani, H., Zahraie, B., 2013. Performance assessment of different data mining methods in statistical downscaling of daily precipitation. J. Hydrol. 492, 1–14.
Permai, S.D., Tanty, H., 2018. Linear regression model using bayesian approach for energy performance of residential building. Procedia Comput. Sci. 135, 671–677.
Petković, D., Gocic, M., Trajkovic, S., Milovančević, M., Šević, D., 2017. Precipitation concentration index management by adaptive neuro-fuzzy methodology. Clim. Change 141, 655–669.
Playán, E., Salvador, R., Faci, J.M., Zapata, N., Martínez-Cob, A., Sánchez, I., 2005. Day and night wind drift and evaporation losses in sprinkler solid-sets and moving laterals. Agric. Water Manag. 76 (3), 139–159.
Playán, E., Zapata, N., Faci, J.M., Tolosa, D., Lacueva, J.L., Pelegrín, J., Salvador, R., Sánchez, I., Lafita, A., 2006. Assessing sprinkler irrigation uniformity using a ballistic simulation model. Agric. Water Manag. 84 (1–2), 89–100.
Roy, D.K., Biswas, S.K., Mattar, M.A., El-Shafei, A.A., Murad, K.F.I., Saha, K.K., Datta, B., Dewidar, A.Z., 2021. Groundwater level prediction using a multiple objective genetic algorithm-grey relational analysis based weighted ensemble of ANFIS models. Water 13, 3130.
Roy, D.K., Datta, B., 2017. A surrogate based multi-objective management model to control saltwater intrusion in multi-layered coastal aquifer systems. Civ. Eng. Environ. Syst. 34, 238–263.
Sadeghi, S.-H., Peters, T.R., Amini, M.Z., Malone, S.L., Loescher, H.W., 2015. Novel approach to evaluate the dynamic variation of wind drift and evaporation losses under moving irrigation systems. Biosyst. Eng. 135, 44–53.
Sadeghi, S.H., Peters, T., Shafii, B., Amini, M.Z., Stöckle, C., 2017. Continuous variation of wind drift and evaporation losses under a linear move irrigation system. Agric. Water Manag. 182, 39–54.
Sarwar, A., Peters, R.T., Mehanna, H., Amini, M.Z., Mohamed, A.Z., 2019. Evaluating water application efficiency of low and mid elevation spray application under changing weather conditions. Agric. Water Manag. 221, 84–91.
Salehi, F., Razavi, S.M.A., 2012. Dynamic modeling of flux and total hydraulic resistance in nanofiltration treatment of regeneration waste brine using artificial neural networks. Desalin. Water Treat. 41, 95–104.
Samadi, M., Jabbari, E., Azamathulla, H.M., Mojallal, M., 2015. Estimation of scour depth below free overfall spillways using multivariate adaptive regression splines and artificial neural networks. Eng. Appl. Comput. Fluid Mech. 9 (1), 291–300.
Sanchez, I., Faci, J.M., Zapata, N., 2011. The effects of pressure, nozzle diameter and meteorological conditions on the performance of agricultural impact sprinklers. Agric. Water Manag. 102, 13–24.

Sanchez, I., Zapata, N., Faci, J.M., 2010. Combined effect of technical, meteorological and agronomical factors on solid-set sprinkler irrigation: II. Modifications of the wind velocity and of the water interception plane by the crop canopy. Agric. Water Manag. 97 (10), 1591–1601.
Sarwar, A., Peters, R.T., Shafeeque, M., Mohamed, A., Arshad, A., Ullah, I., Saddique, N., Muzammil, M., Aslam, R.A., 2021. Accurate measurement of wind drift and evaporation losses could improve water application efficiency of sprinkler irrigation systems – A comparison of measuring techniques. Agric. Water Manag. 258, 107209.
Stambouli, T., Martínez-Cob, A., Faci, J.M., Howell, T., Zapata, N., 2013. Sprinkler evaporation losses in alfalfa during solid-set sprinkler irrigation in semiarid areas. Irrig. Sci. 31, 1075–1089.
Sharda, V.N., Prasher, S.O., Patel, R.M., Ojasvi, P.R., Prakash, C., 2008. Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data. Hydrol. Sci. J. 53 (6), 1165–1175.
Sugeno, M., 1985. Industrial Applications of Fuzzy Control. Elsevier Science Inc, p. 269.
Supharatid, S., 2003. Application of a neural network model in establishing a stage–discharge relationship for a tidal river. Hydrol. Process. 17 (15), 3085–3099.
Swingler, K., 2001. Applying Neural Networks, A Practical Guide, third ed. Academic Press, San Francisco, CA.
Tabari, H., Marofi, S., Sabziparvar, A.A., 2010. Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. Irrig. Sci. 28 (5), 399–406.
Taheriyoun, M., Karamouz, M., Baghvand, A., 2010. Development of an entropy-based fuzzy eutrophication index for reservoir water quality evaluation. J. Environ. Health Sci. Eng. 7 (1), 1–14.
Tang, X., Hong, H., Shu, Y., Tang, H., Li, J., Liu, W., 2019. Urban waterlogging susceptibility assessment based on a PSO-SVM method using a novel repeatedly random sampling idea to select negative samples. J. Hydrol. 576, 583–595.
Thirumalaiah, K., Deo, M.C., 1998. River stage forecasting using artificial neural networks. J. Hydrol. Eng. 26–32.
Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer, New York.
Vapnik, V.N., Golowich, S., Smola, A.J., 1996. Support vector method for function approximation, regression estimation and signal processing. Adv. Neural Inform. Process. Syst. 9, 281–287.
Uddin, J.M., 2012. Measurements of Evaporation During Sprinkler Irrigation. University of Southern Queensland (Doctor of Philosophy (PhD) thesis).
Wu, J., Li, P., Qian, H., Chen, J., 2015. On the sensitivity of entropy weight to sample statistics in assessing water quality: statistical analysis based on large stochastic samples. Environ. Earth Sci. 74 (3), 2185–2195.
Wu, J., Sun, J., Liang, L., Zha, Y., 2011. Determination of weights for ultimate cross efficiency using Shannon entropy. Expert Syst. Appl. 38, 5162–5165.
Yan, F., Qiao, D., Qian, B., Ma, L., Xing, X., Zhang, Y., Wang, X., 2016. Improvement of CCME WQI using grey relational method. J. Hydrol. 543, 316–323.
Yan, H.J., Bai, G., He, J.Q., Li, Y.J., 2010. Model of droplet dynamics and evaporation for sprinkler irrigation. Biosyst. Eng. 106, 440–447.
Yassin, M.A., Alazba, A.A., Mattar, M.A., 2016. Artificial neural networks versus gene expression programming for estimating reference evapotranspiration in arid climate. Agric. Water Manag. 163, 110–124.
Zabihi, M., Pourghasemi, H.R., Pourtaghi, Z.S., Behzadfar, M., 2016. GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. Environ. Earth Sci. 75 (8), 665.
Zanetti, S.S., Sousa, E.F., Oliveira, V.P.S., 2007. Estimating evapotranspiration using artificial neural network and minimum climatological data. J. Irrig. Drain Eng 83–89.
Zapata, N., Playán, E., Martinez-Cob, A., Sánchez, I., Faci, J.M., Lecina, S., 2007. From on-farm solid-set sprinkler irrigation design to collective irrigation network design in windy areas. Agric. Water Manag. 87 (2), 187–199.
Zarei, A.R., Mahmoudi, M.R., Shabani, A., 2021. Investigating of the climatic parameters effectiveness rate on barley water requirement using the random forest algorithm, Bayesian multiple linear regression and cross-correlation function. Paddy Water Environ. 19, 137–148.
Zou, Z.H., Yi, Y., Sun, J.N., 2006. Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. J. Environ. Sci. 18 (5), 1020–1023.