INTERNATIONAL
Agrophysics
www.international-agrophysics.org

*Int. Agrophys., 2025, 39, 427-442*
doi: 10.31545/intagr/205684

# Application of soft computing techniques to estimate wind drift and evaporation loss in sprinkler irrigation**

*Ahmed A. Al-Othman*[1], *Ahmed Z. Dewidar*[2], *Mohamed A. Mattar*[2]

[1]Department of Agricultural Engineering, College of Food and Agriculture Sciences, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia
[2]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

*Received April 1, 2025; accepted May 28, 2025*

A b s t r a c t. This study investigates the predictive performance of five models: Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Multivariate Adaptive Regression Splines (MARS), Piecewise Linear Regression (PLR), and Support Vector Regression (SVR) in estimating wind drift and evaporation losses (*WDEL*) in varying environmental conditions. Using a diverse dataset encompassing key environmental variables, such as riser height, operating pressure, nozzle diameters, wind speed, air temperature, and relative humidity, the models were trained and tested to assess their ability to capture complex, nonlinear relationships affecting *WDEL*. The results reveal that ANN outperformed all the other models, achieving the highest correlation coefficient values and the lowest root mean square error and mean absolute error, highlighting its superior ability to generalize to unseen data. In contrast, ANFIS and MARS exhibited moderate success, with higher prediction errors, especially in extreme conditions. PLR and SVR, which assume linear relationships, struggled to model the nonlinear dynamics governing *WDEL*, resulting in significantly lower accuracy. These findings underscore the importance of employing nonlinear models, such as ANN, to accurately predict *WDEL* in complex environmental systems. The study concludes that ANN is the most robust and reliable model for *WDEL* prediction, offering insights into future research directions, including hybrid models and ensemble approaches to further enhance predictive accuracy.

K e y w o r d s: evaporation and drift losses, sprinkler irrigation, machine learning, regression

## 1. INTRODUCTION

Sprinkler irrigation is widely acknowledged for its ability to save up to 50% more water compared to traditional surface irrigation methods, offering such advantages as efficiency, cost-effectiveness, and ease of installation (Li *et al.*, 2015). However, despite its potential for water conservation, some of the water sprayed by the system is lost in transit before it reaches the crop canopy or soil. This loss, referred to as wind drift and evaporation losses (*WDEL*), reduces the overall effectiveness of the irrigation system, as water that could be used by crops is dissipated into the environment (Sadeghi *et al.*, 2017; Stambouli *et al.*, 2013). Accurate estimation of *WDEL* is essential because it directly affects irrigation management decisions. The challenge lies in finding the optimal balance in the system design: overdesigning can result in unnecessary costs, while underdesigning may cause insufficient water supply, leading to crop stress. Moreover, *WDEL* plays a significant role in determining water rights, which are often subject to legal and financial disputes, as well as influencing general water management strategies in agriculture (Sarwar *et al.*, 2021).

The effectiveness of sprinkler irrigation systems is subject to a variety of influencing factors. These can be broadly classified into design elements, such as the type

*Corresponding author e-mail: mmattar@ksu.edu.sa
**Ongoing Research Funding program - Research Chairs (ORF-RC-2025-5503), King Saud University, Riyadh, Saudi Arabia.

of sprinkler, nozzle diameter, and spacing, and operational considerations, including the pressure, sprinkler height, and timing of irrigation. Additionally, environmental conditions, particularly meteorological factors, also impact the system's performance (Carrión *et al.*, 2001; Playán *et al.*, 2006). In practice, not all of the water emitted by sprinklers reaches its target. *WDEL* values reported in the literature range between 2 and 50%, depending on both the type of the system used and the local climate in which it is applied (Playán *et al.*, 2005; Sarwar *et al.*, 2019). Among the design factors, the nozzle diameter has been identified as critical: larger diameters produce bigger droplets that are less prone to *WDEL* (Keller and Bliesner, 1990). Conversely, raising the height of the nozzle increases *WDEL* due to the longer drop path, which exposes the water to wind for an extended period. Additionally, increasing operational pressure reduces the size of water droplets, which exacerbates *WDEL* (Montero *et al.*, 2003). Environmental conditions, such as wind speed, further influence the trajectory of water droplets, intensifying both evaporation and *WDEL* (Dechmi *et al.*, 2003; Playán *et al.*, 2006; Zapata *et al.*, 2007; Sanchez *et al.*, 2010).

To estimate *WDEL*, a common approach involves comparing the amount of water discharged by the sprinkler system with the amount collected using various measurement devices and mathematical models (Yan *et al.*, 2010; King *et al.*, 2012; Sadeghi *et al.*, 2015). Field and laboratory studies have attempted to analyze *WDEL*, but their results often vary due to differences in measurement techniques. These disparities underscore the challenges of accurately determining *WDEL* in real-world conditions. For example, using catch cans to measure water distribution often leads to experimental errors. Consequently, mathematical models have become essential for providing more reliable *WDEL* estimates. The accurate calculation of *WDEL* is crucial not only for optimizing the sprinkler system design but also for effective water resource management (Maroufpoor *et al.*, 2018).

In recent years, advances in artificial intelligence (AI) have proven valuable in various domains of water engineering, irrigation, and agriculture (Supharatid, 2003; Kumar *et al.*, 2004; Cigizoglu and Kisi, 2006; Chang *et al.*, 2010; Ismail *et al.*, 2012; Kalra *et al.*, 2013; Petković *et al.*, 2017; Mattar and Alamoud, 2015; Mattar *et al.*, 2015, 2018; Kuzman *et al.*, 2021; Lakovic *et al.*, 2021; Roy *et al.*, 2021). Several AI techniques have gained popularity, including Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multivariate Adaptive Regression Spline (MARS), Probabilistic Linear Regression (PLR), and Support Vector Regression (SVR). These techniques are valued for their ability to handle nonlinear problems, their quick processing capabilities, parallel operation, self-learning abilities, and potential for generalization (Basheer and Hajmeer, 2000). By utilizing mathematics and statistical methods, these AI approaches can link input parameters with system outputs (Hamdia *et al.*, 2015). For instance, ANN has the ability to learn from examples, identify data patterns, and process information quickly. ANFIS combines neural networks with fuzzy logic, allowing it to handle more complex reasoning tasks. These methods provide useful alternatives to traditional empirical or physically based formulas, such as the Christiansen uniformity coefficient or the F-factor for multiple outlets. These traditional models, while widely used in irrigation engineering, often rely on simplifying assumptions that may not capture the nuanced, nonlinear interactions among environmental and design parameters influencing *WDEL* (Hamdia *et al.*, 2015). For instance, they are typically limited in representing complex feedback mechanisms involving wind dynamics and temperature fluctuations. Consequently, soft computing techniques like ANN, ANFIS, and MARS offer more robust and flexible modeling capabilities, especially in variable field conditions. MARS, a flexible modeling technique suitable for high-dimensional data, does not assume a specific relationship between input and output variables. Instead, it constructs models using spline functions, which are selected dynamically based on the data (Friedman, 1991; Friedman and Roosen, 1995). MARS serves as a bridge between parametric and nonparametric methods (Huang *et al.*, 2019). Meanwhile, PLR applies Bayesian methods for parameter estimation by combining prior, likelihood, and posterior distributions (Permai and Tanty, 2018). SVR is rooted in constrained optimization theory, focusing on minimizing structural risk to produce optimal solutions (Tang *et al.*, 2019).

Collectively, the cited studies form a strong foundation for this research, with each contributing specific insight that informs the selection and application of the soft computing techniques used herein. For example, Cigizoglu and Kisi (2006) demonstrated the improved accuracy of ANFIS over empirical methods in modeling daily evaporation rates, which supports our selection of ANFIS for modeling *WDEL*. Zarei *et al.* (2021) explored how different climatic variables influence potential evapotranspiration using machine learning methods like PLR and random forest, providing a methodological basis for integrating similar regression-based models in our study. Emamgholizadeh and Mohammadi (2021) combined SVR with hybrid optimization techniques to estimate soil properties, highlighting the model's capability in capturing nonlinear patterns in environmental data. These works validate the relevance and applicability of AI-driven techniques in hydrological and agricultural modeling.

Building on the insights and methodologies established by these previous studies, the current research addresses the critical need for efficient and accurate estimation of *WDEL*, especially given the practical limitations and variability associated with direct field measurements. In this context, mathematical modeling offers a viable alternative by enabling the analysis of *WDEL* across a broad range of operational and environmental conditions. Accordingly, the primary objectives of this study are threefold: (1) to evaluate

the applicability of three machine learning techniques (ANN, ANFIS, and SVR) and two regression techniques (MARS and PLR) for modeling *WDEL* in sprinkler irrigation systems; (2) to assess and compare the predictive performance of these models; and (3) to identify the most influential input variables driving *WDEL* through a comprehensive contribution analysis.

## 2. MATERIAL AND METHODS

### 2.1. Experimental datasets

The modeling in this study is based on experimental data compiled from several peer-reviewed sources, including Abo-Ghobar (1993), Dechmi *et al.* (2003), Sanchez *et al.* (2010, 2011). These studies investigated wind drift and evaporation losses (*WDEL*) in diverse environmental and operational conditions, providing a total of 109 data points from multiple field locations. The data were used to train and test the five predictive models developed in this study: Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Multivariate Adaptive Regression Splines (MARS), Probabilistic Linear Regression (PLR), and Support Vector Regression (SVR). Before detailing each experiment individually, it is important to describe the field measurement techniques employed in the original studies, as these methods directly influence the quality and interpretation of the dataset.

#### 2.1.1. Field measurement techniques

In all referenced experiments, *WDEL* was assessed using standard field measurement tools, primarily catch cans and pluviometers. Catch cans were strategically placed across the irrigated plots to collect water droplets during sprinkler operation, allowing the calculation of the average irrigation depth received at the ground surface. In studies like Sanchez *et al.* (2010), pluviometers were additionally used to capture water applied over vegetative surfaces such as maize and alfalfa, improving the precision of drift and evaporation loss estimates in crop canopy conditions. Despite their practicality, these techniques are subject to experimental uncertainties due to wind variability, spray overlap, and spatial distribution inconsistencies. Nonetheless, they remain widely accepted for *WDEL* assessment and form the basis of the dataset used in this research.

#### 2.1.2. Summary of experimental studies

The individual experiments contributing to the dataset are summarized below:

First Experiment (Abo-Ghobar, 1993): Conducted at the educational farm in Riyadh, Saudi Arabia, this study measured *WDEL* in a wide range of meteorological and operational conditions during the period from April to July 1991. Ten types of impact sprinklers with varying nozzle sizes were tested. Catch cans were used for water collection, and the systems operated for durations between 1 to 2 h depending on nozzle configuration. The experiment provided valuable baseline data on how nozzle characteristics and environmental variables influence *WDEL* in arid climates.

Second Experiment (Dechmi *et al.*, 2003): This experiment took place at the Zaragoza experimental farm in Spain from May to August 2000. It focused on evaluating the performance of a solid-set sprinkler system irrigating maize. The sprinklers were equipped with a 4.4 mm main nozzle and a 2.4 mm auxiliary nozzle, both mounted at a riser height of 2.3 m. Operating at a pressure of 300 kPa, the system achieved a wetted radius of 11 meters. Data on water application and losses were collected using standard field instrumentation in varying environmental conditions.

Third Experiment (Sanchez *et al.*, 2010): Conducted in 2006, this study also took place at the Zaragoza experimental farm. It focused on assessing *WDEL* in solid-set sprinkler systems irrigating alfalfa and maize. A 4 mm main nozzle with a jet-straightening vane and a 2.4 mm auxiliary nozzle were used, and pluviometers were deployed over crop canopies to capture applied water. The experiment was designed to evaluate the impact of vegetation, wind speed, and humidity on water losses, offering critical insights into how plant cover affects *WDEL* measurement.

Fourth Experiment (Sanchez *et al.*, 2011): Performed between 2003 and 2004 at the same Zaragoza facility, this study investigated the influence of different nozzle diameters, operating pressures (ranging from 180 to 420 kPa), and meteorological conditions on sprinkler performance. The system utilized a 4.4 mm main nozzle and a 2.4 mm spreader nozzle in low wind conditions. Each test was conducted over a two-hour period, and measurements were taken using catch cans to determine the spatial distribution and quantify *WDEL*.

The experiments compiled in this study utilized solid-set sprinkler systems arranged in a rectangular configuration. This layout provides statistically consistent coverage and uniform spatial sampling, which is advantageous for analyzing water distribution and *WDEL*. It contrasts with center-pivot systems, where radial variability in droplet trajectory and coverage patterns introduces additional complexity in estimating losses. The parameters presented in Table 1 encompass both system design features and environmental conditions that influence *WDEL*. In addition to the core input variables used in model development, we also considered two derived hydraulic parameters: discharge coefficient (CD) and discharge exponent (n) to provide further context. These were obtained either directly from the literature or inferred using manufacturer data and typical sprinkler operating ranges. Although CD and n were not included as direct inputs in model training, they offer valuable insight into how sprinkler design characteristics interact with operating pressure to affect discharge behavior and water distribution in field conditions.

*WDEL* was calculated by comparing the applied irrigation depth ($\overline{ID}_C$) with the average irrigation depth ($ID_D$), following Christiansen's (1942) formula. The Christiansen's

(1942) method applied in this study refers specifically to the uniformity coefficient (CU) approach, which evaluates the evenness of water distribution based on catch can data. This method is particularly relevant for systems with multiple outlets and is widely accepted in sprinkler irrigation evaluations. By using CU, the study ensures a standardized and comparable estimation of distribution uniformity and associated water losses across experimental setups.

$$\mathrm{WDEL} = \frac{ID_D - \overline{ID}_C}{ID_D}, \quad (1)$$

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

where: $C_D$ is the discharge coefficient, $A$ is the nozzle orifice area, $g$ is the acceleration due to gravity, $n$ is the discharge exponent, $t$ is the operation time, and both $S_S$ and $S_L$ are the sprinkler's allocated area.

The findings from these experiments indicate that a portion of the applied irrigation water was consistently lost due to evaporation and wind drift, with droplets often displaced beyond the intended irrigated area. Table 1 summarizes the experimental data used to develop the predictive models. Data were collected under a broad spectrum of design, operational, and meteorological conditions. The modeling variables include riser height, pressure, nozzle dimensions, discharge rates, wind speed, air temperature, and relative humidity. To estimate wind drift and evaporation losses (*WDEL*), five models ANN, ANFIS, MARS, PLR, and SVR were developed using this dataset. A total of 109 observations were compiled from four independent experimental studies. The dataset was randomly divided into 70% for training and 30% for testing, ensuring consistent model evaluation across all methods. To promote transparency and reproducibility, the complete dataset is available from the authors upon reasonable request, along with supporting metadata and documentation.

Table 2 offers a statistical overview of the input and output parameters utilized for training and testing the models. The parameters include maximum ($X_m$), mean ($X_a$), minimum ($X_n$), standard deviation ($S_x$), kurtosis coefficient (Kx), and skewness coefficient ($C_{sx}$). The maximum values for both design and operational parameters ($H_r$, P, $d_m$, $d_a$, $q_m$, $q_a$) were consistent across both training and testing datasets, as were the minimum values. For wind speed (WS) and relative humidity (RH), maximum values reached over 7 m s$^{-1}$ and 80%, respectively, in both datasets. Of all the variables, the auxiliary nozzle diameter ($d_a$) showed the highest kurtosis and skewness coefficient.

### 2.2. Artificial neural network (ANN)

Artificial Neural Networks (ANNs) inherently possess the capability to identify and model complex, nonlinear relationships by learning from data inputs (Swingler, 2001). The architecture of an ANN is comprised of three primary layers: the input layer, which holds the independent variables (i); the hidden layer (j); and the output layer, where the dependent variables (k) reside (Fig. 1). Developing an efficient ANN involves careful selection of several structural components, including the number of hidden layers, the number of neurons in each hidden layer, the learning algorithm, and the activation functions (Dogan *et al.*, 2008) (Table 3). One of the most commonly utilized algorithms for training feed-forward ANNs is the back-propagation algorithm (Thirumalaiah and Deo, 1998; Ciĝizoĝlu, 2003; Jain and Srinivasulu, 2004; Fernando and Shamseldin, 2009), which is particularly suited for modeling nonlinear and complex functions (Salehi and Razavi, 2012). In line with this, the current study implemented back-propagation to train the ANN, featuring a hidden layer and an output layer. Prior to training, the input and output data undergo automatic normalization within a range of 0.15 to 0.85. This normalization accelerates the training process and enhances the model's generalization abilities. Once normalized,

**Table 1.** 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-Ghobur (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) |

$H_r$ – riser height, P – operating pressure, $d_m$ – main nozzle diameter, $d_a$ – auxiliary nozzle diameter, $q_m$ – discharge from main nozzle, $q_a$ – discharge from auxiliary nozzle, WS – wind speed, T – air temperature, RH – relative humidity, *WDEL* – wind drift and evaporation losses.

**Table 2.** Descriptive analysis of the training and testing variables

| Variable | 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 (ºC) | 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 |

$X_m$ – maximum value, $X_a$ – mean value, $X_n$ – minimum value, $S_x$ – standard deviation, $K_x$ – kurtosis coefficient, $C_{sx}$ – skewness coefficient, $H_r$ – riser height, P – operating pressure, $d_m$ – main nozzle diameter, $d_a$ – auxiliary nozzle diameter, $q_m$ – water discharge by main nozzle, $q_a$ – water discharge by auxiliary nozzle, WS – wind speed, T – air temperature, RH – relative humidity, and *WDEL* – wind drift and evaporation losses.

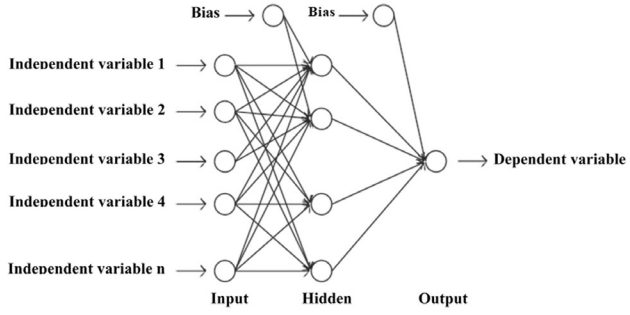


**Fig. 1.** Structural diagram of a standard artificial neural network (Li *et al.*, 2017).

data is fed unidirectionally through the network, beginning from the input layer, passing through the hidden layer, and finally reaching the output layer.

Each neuron in the ANN operates as an independent processing unit, comprising three fundamental elements: input links, a central processing unit, and output links. Connections between neurons are assigned weights (W), which quantify the strength of associations between neurons across adjacent layers. Additionally, a bias term (B), or threshold, is applied within the central processing unit to refine the processing of weight information (Izadifar, 2010). The activation functions (f), responsible for transforming input signals in both the hidden and output layers, are executed at this point. In many engineering applications, the hyperbolic tangent (tanh) function is one of the most widely employed activation functions (Dawson and Wilby, 1998; Zanetti *et al.*, 2007). In this study, the tanh function was adopted for both the hidden and output layers to optimize the network's training process. The general mathematical form of the tanh activation function is as follows:

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

The output of the neural network can be described by the following expression (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: $Y_k$ represents the network's output; $X_i$ is the normalized input parameter; $(W_1)_{ji}$ is the weight of the input layer in connection to the hidden layer; $(W_2)_{kj}$ is the weight of the hidden layer in connection to the output layer; $n_i$ and $n_j$ represent the number of neurons in the input and output layers, respectively; $(B_1)_j$ and $(B_2)_k$ denote the biases in the hidden and output layers, respectively.

After each forward pass through the network, the error, defined as the difference between the predicted output and the target values, is computed. If this error surpasses the predetermined threshold, the network adjusts the associated weights through a process called back-propagation. Conversely, if the error falls within acceptable limits, the training process is considered complete. The optimal number of neurons in the hidden layer is determined using a trial-and-error approach (Jain *et al.*, 2008). Upon the completion of training, the model is evaluated by testing it on previously unseen data to assess its predictive accuracy and performance.

**Table 3.** Tuning parameters used for each model

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

### 2.3. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The process of developing an Adaptive Neuro-Fuzzy Inference System (ANFIS) starts with creating an initial Fuzzy Inference System (FIS), whose parameters are fine-tuned through a hybrid optimization algorithm to complete the ANFIS structure. To construct the initial FIS, the fuzzy c-means (FCM) clustering algorithm (Bezdek *et al.*, 1984) was used. FCM is an effective method for organizing datasets into clusters, thereby reducing both the linear and nonlinear parameters in the FIS. The number of clusters influences the number of rules generated in the FIS (Table 3). For this study, a Sugeno-type FIS-commonly referred to as the Takagi-Sugeno-Kang model, introduced by Sugeno in 1985-was chosen as the initial FIS (Sugeno, 1985). Sugeno-type FISs are founded on the principles of fuzzy logic, which integrates fuzzy set theory, fuzzy "if-then" rules, and fuzzy reasoning. The fundamental components of an FIS consist of: (i) a rule base composed of fuzzy "if-then" rules, (ii) a database specifying the membership functions (MFs) in terms of type, number, and range, and (iii) an inference mechanism that governs the reasoning process (Jang *et al.*, 1997). These "if-then" fuzzy rules enable the FIS to model nonlinear correlations between input and output variables. In a Sugeno FIS, the input variable membership functions are typically Gaussian, while the output functions are linear. The Gaussian membership function is characterized by two key parameters and is mathematically defined as follows (Jang *et al.*, 1997):

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

where: $c$ represents the center of the membership function (MF) and $\sigma$ denotes its width.

In this study, a Sugeno-type Fuzzy Inference System (FIS) was created using the fuzzy c-means (FCM) clustering algorithm to serve as the foundation for the final Adaptive Neuro-Fuzzy Inference System (ANFIS). The optimal number of clusters was determined through multiple trials, in which different cluster configurations were tested by assessing the root mean square error (*RMSE*) of the predictive outcomes. The cluster arrangement that minimized *RMSE* and exhibited the smallest discrepancy between the *RMSE* values of the training and testing datasets was chosen as the most effective.

Subsequently, the Sugeno-type ANFIS was refined by applying a hybrid algorithm to adjust the parameters of the initial FIS. Compared to the Mamdani ANFIS structure, Sugeno-type ANFIS models are simpler in their design but offer enhanced learning capabilities (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:

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

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

These rules are illustrated in Fig. 2, which shows the architecture of the Sugeno fuzzy inference system-based ANFIS. This architecture is composed of five layers: the fuzzy layer, the product layer, the normalized layer, the defuzzification layer, and the total output layer.

### 2.4. Multivariate Adaptive Regression Spline (MARS)

Multivariate Adaptive Regression Splines (MARS) is a flexible, non-parametric approach for adaptive regression, enabling the creation of predictive models that can adapt to varying data structures (Friedman, 1991). MARS segments the solution space into several regions based on the input variables and applies individual splines or basis functions to each region (Bera *et al.*, 2006). In this study, MARS-based models were developed to forecast *WDEL* using nine predictor variables. The MARS algorithm employs a combination of forward and backward stepwise techniques to model the input-output relationships. Notably, the backward stepwise process aids in removing irrelevant

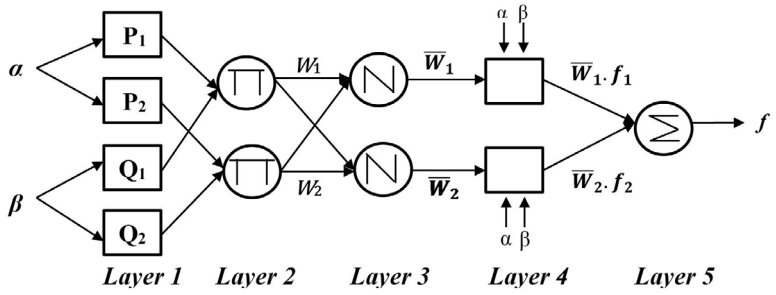


**Fig. 2.** ANFIS architecture built on a Sugeno fuzzy inference system with two inputs (Jang, 1993).

input variables, thus simplifying the model and mitigating the risk of overfitting. The mathematical formulation of the MARS model is outlined as follows (Roy and Datta, 2017):

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

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

where: *i* and *j* are the indices for Basis functions and input variables, respectively; $BF_i$ denotes *ith* Basis functions; *Xj* represents *jth* input variables; $\alpha$ is a constant known as knot; $\beta$ is a constant value; $\gamma_k$ denotes the corresponding coefficients of $BF_i(X)$.

The selection of the number of basis functions is guided by the number of significant variables required to construct a robust model (Table 3). In this study, up to 100 forward steps were permitted during the model development phase. MARS employs internal heuristics to determine the minimum number of observations between knots, taking into account both the sample size and the complexity of the model. During the forward selection process, all variables are treated equally without any penalties. However, in the backward elimination phase, only the most relevant input variables are retained, promoting model simplicity and reducing the risk of overfitting.

### 2.5. Probabilistic Linear Regression (PLR)

Probabilistic Linear Regression (PLR), also known as empirical Bayesian Linear Regression, is employed to model the relationship between input variables and a target output variable. To solve the regression problem, PLR typically utilizes either the Expectation-Maximization (EM) algorithm (Dempster *et al.*, 1977) or the Mackay fix-point iteration method (MacKay, 1992). In this study, the EM algorithm was selected for developing the PLR model (Table 3). The EM algorithm functions by maximizing the log-likelihood or log-posterior density functions, which are integrated with latent variables. Throughout the optimization process, the EM algorithm implicitly constructs a lower bound, $\mathcal{F}_{EM}$ of the objective function, $l(\alpha)$, ensuring that:

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

where: $q(.)$ refers to an arbitrary probability distribution in the space $z$. As stated by Dempster *et al.* (1977), the following are the widely recognized results in the EM literature, derived from Jensen's Inequality theorem (Jensen, 1906).

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

An alternative optimization problem is now constructed to be addressed instead of directly optimizing the original objective function $l(\alpha)$. This redefined optimization task is expressed as:

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

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:

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

$$\text{M-Step: } \propto^{(t+1)} := \text{argmax}_{\propto} \mathcal{F}_{EM}(q^{(t)}, \propto) = \\ \text{argmax}_{\propto} \mathbb{E}_{q^{(t)}}[\log\, p(z|\propto)], \quad (12)$$

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

$$\mathcal{F}_{EM}(q^{(t)}, \alpha^{(t)}) = l(\propto^{(t)}). \quad (13)$$

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 and therefore is guaranteed to converge.

### 2.6. Support Vector Regression (SVR)

Support Vector Regression (SVR) is a non-parametric modeling technique known for its effectiveness in developing robust predictive models. This method fundamentally relies on kernel functions to construct prediction models, offering flexibility and adaptability in handling various types of data. A key feature of SVR is its use of an asymmetrical loss function for supervised learning tasks. One

of the primary advantages of SVR is its ability to maintain computational efficiency, regardless of the dimensionality of the input space, which makes it particularly suitable for high-dimensional datasets (Awad and Khanna, 2015). Detailed descriptions of the SVR approach are provided by Vapnik (1995) and Gunn (1997), and this section presents a concise overview of its methodology. In a typical regression problem, the training dataset can be represented as:

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

where: $a_i$ ($i$=1,2,3, …, $N$) represents a vector consisting of real independent variables; $b_i$ ($i$=1,2,3, …, $N$) is the associated scalar real independent variable. The regression equation in the feature space for the training dataset can be expressed as:

$$z(a, w) = (w\, \emptyset(a) + c), \quad (15)$$

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

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

and,

$$L_\varepsilon\big(b, z(a, w)\big) = \begin{cases} 0 & \text{if } |b - z(a, w)| \le \varepsilon \\ |b - z(a, w)| - \varepsilon & \text{otherwise} \end{cases}. \quad (17)$$

The first term on the left side of Eq. (16) signifies the empirical error, while the parameter C governs the balance between this empirical error and the complexity of the model, which is reflected in the second term of the equation. Equation (17) defines a loss function known as the $\varepsilon$ - insensitive loss function (Vapnik *et al.*, 1996). By introducing Lagrangian multipliers $\beta$ and $\beta$*, the optimization problem expressed in Eq. (16) is reformulated into its dual form. The non-zero coefficients, together with their associated input vectors, $a_i$, are referred to as support vectors. The final expression of the equation is presented as follows:

$$z(a, \beta_i, \beta_i^*) = \sum_{i=1}^{N_{sv}}(\beta_i - \beta_i^*)\left(\emptyset(a_i)\, \emptyset(a_j)\right) + c. \quad (18)$$

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

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

The Karush-Kuhn-Tucker condition is applied to compute the term $c$ in Eqs (18) and (19). The key parameters that influence regression problems using the SVR method include the cost function $C$, the radius of the insensitive tube $\varepsilon$, and the kernel parameters $K(x_i, x_j)$, as shown in Table 3.

### 2.7. Entropy weight calculation

Machine learning-based prediction models often demonstrate variability in their predictive performance depending on the specific evaluation metrics employed. For example, a model's performance might show improvement when assessed using Wilmott's index of agreement (IA), while another model might exhibit superior performance based on the root mean square error (*RMSE*). This inconsistency suggests that multiple performance indices should be utilized when evaluating models, particularly within a decision-theoretic framework, to determine the optimal model. In this study, four benefit indices (where higher values are better) and four cost indices (where lower values are better) were applied to assess the performances of the developed prediction models, with Shannon's entropy method used for ranking.

The integration of prediction models with performance evaluation metrics leads to the formation of a decision matrix. For example, given $m$ prediction models and $l$ performance evaluation indices, the decision matrix can be mathematically represented as follows (Wu *et al.*, 2011):

$$WDEL_{ij} = \begin{bmatrix} WDEL_{11} & WDEL_{21} & \cdots & WDEL_{m1} \\ WDEL_{12} & WDEL_{22} & \cdots & WDEL_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ WDEL_{1l} & WDEL_{2l} & \cdots & WDEL_{ml} \end{bmatrix}. \quad (20)$$

This matrix was normalized to reduce the dimensionality effects of the indices, with the performance values transformed to lie between 0 and 1, denoted as $S_{ij} \in [0,1]$, $i$=1,2,…m; j= 1,2,…, $l$. The standardized matrix, $S_{ij}$ is given by Wu *et al.* (2011):

$$S_{ij} = \begin{cases} \dfrac{WDEL_{ij}}{\max(WDEL_{i1}, WDEL_{i2}, \ldots, WDEL_{il})}, & for\ benefit\ indexe. \\ \dfrac{X_{ij}}{\min(WDEL_{i1}, WDEL_{i2}, \ldots, WDEL_{il})}, & for\ cost\ indexes. \end{cases} \quad (21)$$

The entropy-based ranking method was then employed to assign weights to each prediction model. The process, as outlined by Li *et al.* (2011), involved several steps:

Step 1*:* Calculation of the entropy value of each index using Shannon's information entropy principles. The entropy value of the *jth* index was calculated as:

$$Entropy_j = -k\sum_{i=1}^{m} f_{ij} \ln f_{ij}, \quad (22)$$

where:

$$f_{ij} = S_{ij} / \sum_{i=1}^{m} S_{ij}, \quad (23)$$

$$k = 1/\ln m. \quad (24)$$

Step 2: Calculation of each index's entropy weight. The *jth* index's entropy weight was calculated by:

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

This weight reflects the importance of the index in the decision-making process, with higher weights indicating that the index contains more information and is more critical to the decision.

Step 3: The ranking weight for each model is determined by multiplying the entropy weight of each index by its corresponding standardized value. This process can be mathematically represented as:

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

Step 4: Determination of model ranking:

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

Step 5: Calculation of entropy weight for individual prediction models:

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

### 2.8. Statistical performance criteria

Five statistical metrics were utilized to assess the performance of the five proposed models: ANN, ANFIS, MARS, PLR, and SVR. The evaluation criteria included the correlation coefficient (R), Index of Agreement (IA), Kling-Gupta Efficiency (KGE), Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), and Mean Absolute Relative Error (*MARE*). These metrics are mathematically defined by the following equations:

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

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

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

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

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

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

where: $WDEL_{i,a}$ and $WDEL_{i,p}$ are *WDEL* values at the *ith* 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

### 3.1. Descriptive statistics of input and output variables

The study developed five models: Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Multivariate Adaptive Regression Splines (MARS), Partial Least Regression (PLR), and Support Vector Regression (SVR) to estimate Wind Drift and Evaporation Loss (*WDEL*). The models used several critical environmental variables, including riser height (Hr), operating pressure (P), nozzle diameters (dm and da), wind speed (WS), air temperature (T), and relative humidity (RH) to account for *WDEL*'s complex behavior. The training dataset showed an average $H_r$ of 2.11 m, P of 313.2 kPa, and WS between 0.4 and 8 m $s^{-1}$. *WDEL*, the output variable, ranged widely from 2.4 to 39.6%, underscoring the need for models that can handle the nonlinear relationships driven by environmental conditions, particularly WS and RH.

In Table 2, the comparison of training and testing datasets reveals consistent patterns in operational variables like $H_r$ and P, but significant variability in environmental conditions during the testing phase. For instance, $H_r$ remained stable, with values between 1 m and 2.3 m, though the standard deviation increased from 0.27 m in training to 0.37 m in testing. WS displayed greater variation during testing, with an average increase from 2.71 to 3.28 m $s^{-1}$ and a standard deviation rising to 1.93 m $s^{-1}$, alongside a more negatively skewed distribution (Csx = -1.03). T followed a similar trend, increasing from a maximum of 36°C in training to 42°C in testing, with greater variability (standard deviation = 7.96°C) and a positively skewed distribution (Csx = 0.85). The auxiliary nozzle diameter (da) showed extreme kurtosis (Kx = 31.53) and skewness (Csx = 5.64) in the training set, which slightly decreased in the testing phase (Kx = 21.31, Csx = 4.45), reflecting a more balanced yet still non-normal distribution. RH and *WDEL* also exhibited more variability in testing, with the latter showing a broader dispersion (Csx = 0.08). These findings suggest that more extreme environmental conditions and increased variability in testing could affect the performance and accuracy of the prediction models.

### 3.2. Model performance during the training phase

During the training phase, model performance was evaluated using several statistical metrics: the correlation coefficient (R), Root Mean Square Error (*RMSE*), Mean Absolute Error (*MAE*), Index of Agreement (IA), Kling-Gupta Efficiency (KGE), and Mean Absolute Relative Error (*MARE*). These metrics capture various aspects of prediction quality, including accuracy, agreement, and consistency. Their full definitions and mathematical formulations are provided in Section 2.8. ANN demonstrated superior performance, achieving the highest R value (0.956), indicating a strong relationship between the predicted and observed *WDEL* values (Table 4). Moreover, ANN achieved an IA of 0.978 and a KGE of 0.925, highlighting its ability to capture both the strength and direction of the data variation. ANN's lower *RMSE* (2.662) and *MAE* (2.197) further illustrate its precision during the training phase, as it consistently produced lower prediction errors compared to the other models. Its scatter plot regression equation, y = 0.8985x + 1.561, shows a near-perfect alignment with the observed *WDEL* values, confirming its superior predictive accuracy (Fig. 3). ANFIS followed closely behind ANN, with R = 0.871, IA = 0.927, and higher errors (*RMSE* = 4.462, *MAE* = 3.417, *MARE* = 32.24), with a regression equation of y = 0.758x + 3.8526 (Fig. 3), indicating some deviation from observed values. While ANFIS captured significant patterns in the data, its *RMSE* (4.462) and *MAE* (3.417)

**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 |

R – correlation coefficient, IA – index of agreement, KGE – Kling-Gupta efficiency, *RMSE* – root mean square error, *MAE* – mean absolute error, *MARE* – mean absolute relative error.

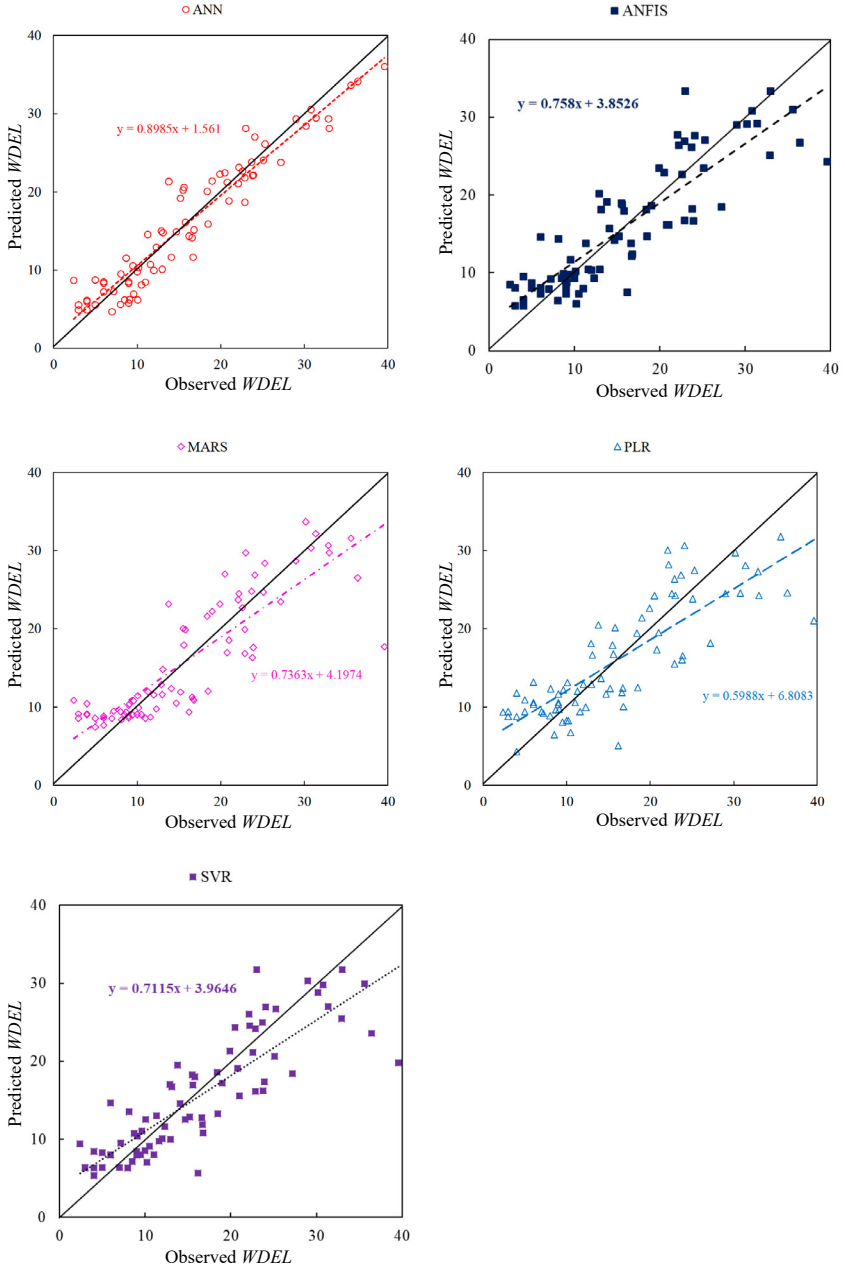


**Fig. 3.** Scatter plot of observed and predicted *WDEL* values for models built throughout the training process.

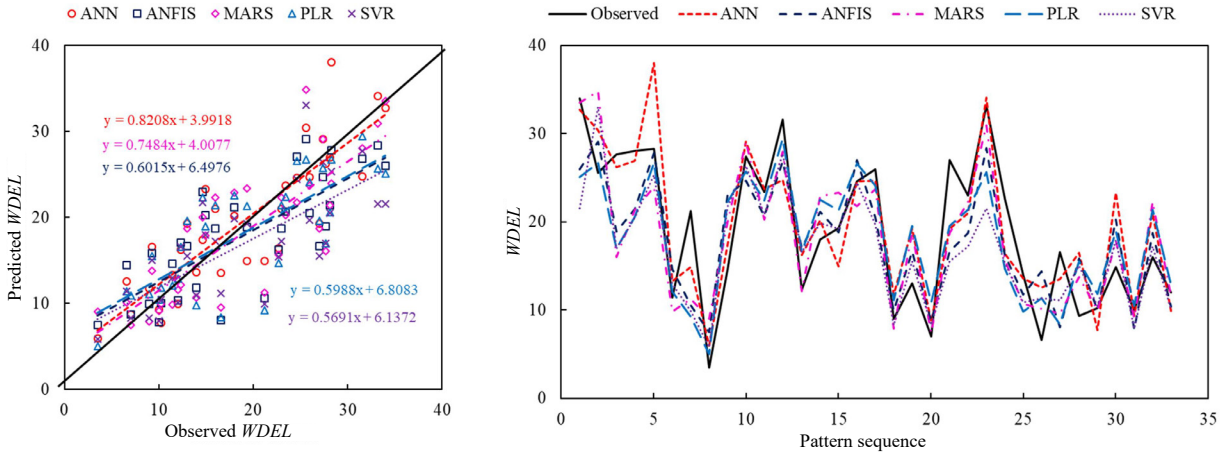


**Fig. 4.** Comparison of observed and predicted *WDEL* values for the developed models throughout the testing process.

were higher than those of ANN, indicating less precision. MARS, a flexible regression model that accommodates interactions between variables, performed reasonably well during training, achieving an R value of 0.858 and an IA of 0.919 (Table 4), with a regression line of y = 0.7363x + 4.1974 (Fig. 3). However, with a higher *RMSE* (4.658) and *MAE* (3.387), MARS exhibited limitations in its predictive precision, particularly for extreme values of *WDEL*. PLR and SVR exhibited the weakest performance during training. PLR, with its assumption of a linear relationship between the input and output variables, achieved the lowest R value of 0.822 and the highest *RMSE* (5.169) and *MAE* (4.108) (Table 4), and a poorly fitting regression equation of y = 0.5988x + 6.8083 (Fig. 3), suggesting that it was poorly suited to model the complex, nonlinear dynamics inherent in the dataset. Finally, SVR, with R = 0.857, IA = 0.914, and *RMSE* of 4.716, performs similarly to MARS but exhibits higher deviations (*MAE* = 3.505, *MARE* = 31.39) (Table 4), with a regression line of y = 0.7115x + 3.9646 (Fig. 3).

### 3.3. Model performance during the testing phase

During the testing phase (Table 4), which provided an opportunity to evaluate the models' ability to generalize to new data, ANN consistently outperforms the other models, achieving the highest R value (0.861), reflecting the strongest linear relationship between predicted and observed values. In contrast, SVR exhibits the lowest R value (0.749), indicating a weaker correlation. The IA further supports ANN's superiority, with a value of 0.925, the highest among the models, demonstrating the best agreement between predictions and observed data. By comparison, SVR again falls short with an IA of 0.831. Similarly, the KGE for ANN is 0.850, highlighting its balanced performance across correlation, bias, and variability, whereas SVR shows the lowest KGE (0.637), indicating less efficient performance. In terms of error metrics, ANN again excels with the lowest *RMSE* (4.356), *MAE* (3.490), and

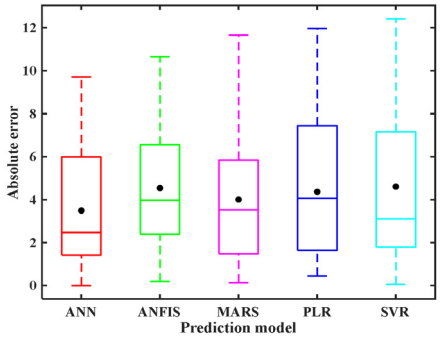


**Fig. 5.** Boxplot for absolute error in predicted *WDEL* with the prediction models during the testing process.

*MARE* (24.15), which underscore its precision in minimizing the magnitude of prediction errors. Conversely, SVR displays the highest *RMSE* (5.876) and *MAE* (4.603), indicating larger discrepancies between predicted and actual values, while ANFIS and PLR perform moderately with higher error rates than ANN but better than SVR. Figure 4 visually confirms these results by plotting the observed versus predicted *WDEL* values. ANN's trend line (y = 0.8208x + 3.9918) is the closest to the ideal line of perfect prediction (y = x), indicating strong alignment between predicted and observed values. In contrast, SVR (y = 0.5691x + 6.1372) and PLR (y = 0.5988x + 6.8083) demonstrate flatter slopes, suggesting significant underprediction. ANFIS (y = 0.6015x + 6.4976) and MARS (y = 0.7484x + 4.0077) perform better than SVR and PLR but still exhibit deviations from the ideal, especially in predicting higher *WDEL* values. Figure 5 further illustrates model performance by comparing the absolute error distributions through a boxplot. ANN consistently shows the narrowest error range and the lowest median error, reflecting its precision and stability. SVR, on the other hand, presents the widest interquartile range and the highest median error, indicating high variability in its predictions. ANFIS and PLR exhibit wider error distributions compared to ANN, but MARS performs moderately, with a relatively tighter error range.

### 3.4. Entropy-based ranking of models

Shannon's entropy-based decision theory was used to rank the prediction models based on entropy weights, which measure the consistency and reliability of each model's performance (Table 5). ANN holds the highest entropy weight (0.999), confirming its top performance. MARS ranks second with a weight of 0.896, reflecting its solid, though slightly less consistent, performance compared to ANN. ANFIS follows closely in third place with a weight of 0.865, showing moderate variability in predictions but still outpacing models like PLR and SVR. PLR is ranked fourth with an entropy weight of 0.833, signifying weaker and less stable predictions. SVR ranks last with the lowest weight (0.794), reflecting its inconsistent performance and higher error variability. Overall, the entropy weight rankings align closely with prior statistical evaluations, positioning ANN as the most reliable model, followed by MARS and ANFIS, while PLR and SVR demonstrate the least reliability and consistency.

### 3.5. Contribution ratio of the ANN model's input variables

The ANN model's superior predictive ability was used to assess the contribution of various input variables ($H_r$, P, $d_a$, $d_m$, $q_a$, $q_m$, WS, RH, and T) in modeling *WDEL*. By analyzing the connection weights from the ANN architecture, the contribution ratios for each input were calculated, revealing the most impactful variables for predicting *WDEL*. Figure 6 provides a detailed analysis of the contribution ratios of the input variables within the ANN model, expressed as percentages. In this figure, it is evident that da has the highest contribution ratio, approximately 17.5%, suggesting that this variable plays a dominant role in the model's predictions. Following closely is Hr, contributing nearly 16.3%, indicating a strong influence on the ANN's outputs. These hydraulic design parameters, i.e. riser height, pressure, and nozzle diameter, are not independent of one another. Instead, they function together to determine the sprinkler's discharge behavior and droplet dynamics. While the models in this study treated these as separate input variables, their collective hydraulic influence is inherently embedded in the dataset, allowing the models to learn and reflect these complex interactions during training. Other significant contributors include qm and WS, each contributing 12.57 and 12.19%, respectively. On the other hand, variables such as qa and dm have lower contribution ratios, closer to 5%, indicating that they have a smaller impact on the model's performance. RH and P, with contributions of around 8-9%, appear to have moderate importance in influencing the predictions.

**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 |

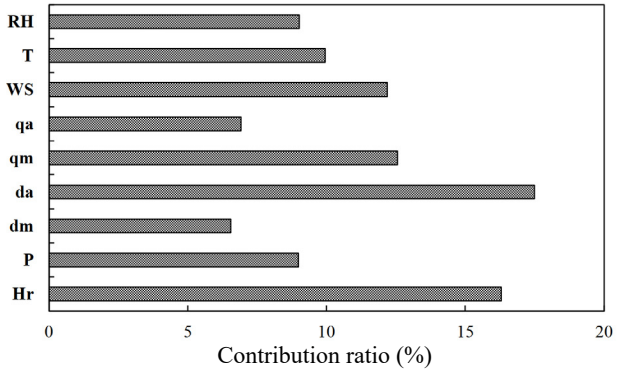


**Fig. 6.** Contribution ratio of the ANN model's input variables.

## 4. DISCUSSION

### 4.1. Model performance and implications

The findings from this study provide valuable insights into the application of advanced machine learning and regression models for predicting wind drift and evaporation losses (*WDEL*) in sprinkler irrigation systems. By comparing five computational approaches – Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Multivariate Adaptive Regression Splines (MARS), Probabilistic Linear Regression (PLR), and Support Vector Regression (SVR), ANN demonstrated a clear advantage over the others, particularly in its ability to generalize from training to testing data, making it a strong candidate for predicting *WDEL* in varying environmental conditions. The success of ANN can be attributed to its inherent capability to self-learn from large datasets and recognize intricate patterns that simpler regression models like PLR or SVR might fail to capture. The consistent underperformance of SVR, in particular, points to the limitations of kernel-based methods in dealing with highly nonlinear, multivariate systems such as *WDEL*, where a wide range of factors influence outcomes (Awad and Khanna, 2015). Similarly, PLR, which relies on linear assumptions, struggled to account for the complex interactions between the system's design and environmental variables. In comparison, MARS and ANFIS performed moderately well, though both exhibited limitations in certain scenarios. MARS, while flexible in handling high-dimensional data (Friedman, 1991), faced challenges in managing the high variability in meteorological conditions, potentially leading to overfitting when dealing with certain input parameters. ANFIS, which combines neural networks and fuzzy logic (Jang *et al.*, 1997), showed potential but

was hindered by the sensitivity of its fuzzy inference system, particularly when fewer membership functions were used, affecting its overall predictive accuracy. Despite these limitations, ANFIS still demonstrated practical utility by modeling *WDEL* more effectively than purely regression-based approaches like PLR.

### 4.2. Role of input variables

The contribution analysis of input variables offers deeper insight into the factors most significantly impacting *WDEL* predictions. Design variables, such as nozzle diameter and riser height, along with meteorological factors like wind speed and air temperature, emerged as key determinants of *WDEL*. The prominence of nozzle diameter aligns with prior studies, which have emphasized that larger droplets produced by larger nozzles are less susceptible to wind drift and evaporation losses (Keller and Bliesner, 1990; Playán *et al.*, 2006). Similarly, wind speed remains a crucial factor due to its ability to disrupt droplet trajectories, exacerbating evaporation losses before water reaches the crop canopy (Dechmi *et al.*, 2003). Interestingly, the meteorological parameters-specifically, air temperature and relative humidity-also played a significant role in determining *WDEL*. This highlights the importance of considering local climate conditions when designing and managing irrigation systems, especially in arid and semi-arid regions where temperature and humidity variations can greatly influence evaporation rates (Molle *et al.*, 2012). The interaction between temperature and humidity is particularly relevant in understanding how the vapor pressure deficit affects droplet evaporation. Under higher temperatures and lower humidity, evaporation rates increase, leading to greater water losses, while higher humidity slows this process, reducing *WDEL*. These findings suggest that a comprehensive approach to *WDEL* modeling must integrate both design and environmental variables to achieve accurate predictions (Sarwar *et al.*, 2021).

### 4.3. Practical applications and implications for irrigation management

From a practical standpoint, the findings of this study have important implications for irrigation management and water conservation. The strong predictive capability of the ANN model makes it a valuable tool for real-time decision-making in irrigation systems, particularly in regions where water scarcity necessitates the efficient use of available resources. By accurately forecasting *WDEL*, irrigation managers can optimize system operations, adjusting such parameters as nozzle diameter and riser height in response to changing weather conditions. This flexibility could significantly reduce water losses due to drift and evaporation, improving the overall efficiency of irrigation systems and maximizing water use for crop production. The ability to predict *WDEL* with high accuracy is also crucial for designing more resilient irrigation systems that can adapt to varying climatic conditions. In particular, regions experiencing high temperatures and frequent winds could benefit from ANN-based predictions that enable precise adjustments to the system's configuration, reducing the negative impacts of weather variability on irrigation efficiency (Playán *et al.*, 2005). Moreover, these predictive models could be integrated into decision support systems, providing farmers and irrigation engineers with real-time guidance on how to minimize water losses and enhance crop water use efficiency.

### 4.4. Limitations and future research directions

While the ANN model's performance in this study was outstanding, there are limitations that future research should address. First, the datasets used in this study were derived from prior experimental research, which, while diverse, may not fully represent the range of environmental conditions present across different geographic regions. Thus, expanding the model's applicability by incorporating more diverse datasets from various climatic zones would improve its generalizability. Additionally, the ANN model's computational demands, particularly during the training phase, may pose challenges in resource-limited environments. Future work should explore the potential for hybrid models that combine the strengths of ANN with other, less computationally intensive methods, providing a balance between predictive accuracy and computational feasibility (Supharatid, 2003). Moreover, while this study focused on *WDEL* in specific operational conditions, future research could expand the scope by including additional factors, such as soil characteristics, plant canopy cover, and irrigation scheduling, which may further affect water losses in sprinkler systems. Such an integrated approach would provide a more holistic understanding of *WDEL* and could lead to the development of more comprehensive models that account for the full range of variables impacting water use efficiency in agriculture.

## 5. CONCLUSIONS

This study compared five soft computing approaches: ANN, ANFIS, MARS, PLR, and SVR for predicting wind drift and evaporation losses (*WDEL*) in sprinkler irrigation systems. Among them, ANN achieved the highest accuracy during both the training and testing phases, demonstrating strong capability in capturing nonlinear interactions between system design parameters and meteorological variables. ANFIS and MARS performed moderately well, while PLR and SVR were less effective, particularly in extreme environmental conditions.

The contribution analysis identified wind speed and relative humidity as the most influential uncontrolled variables affecting *WDEL*, followed by riser height and nozzle dimensions. Based on the compiled dataset, practical operational thresholds were derived. *WDEL* remained below

10% when wind speed was less than or equal to 2 m $s^{-1}$ and relative humidity was greater than or equal to 60%. Losses exceeded 25% when wind speed was above 5 m $s^{-1}$ or relative humidity dropped below 40%. For crops shorter than 1 m, maintaining riser height at or below 2 m helped limit drift losses to under 12%. These findings offer valuable guidance for optimizing the sprinkler design and operation in variable field conditions.

The results highlight the potential of ANN-based tools to enhance irrigation system management through more precise design and scheduling. Future research should focus on validating the proposed thresholds across diverse climatic regions, exploring hybrid and ensemble modeling strategies, and incorporating additional factors such as canopy characteristics and soil moisture to improve model generalizability and practical utility.

## 6. ACKNOWLEDGMENTS

The authors extend their appreciation to Ongoing Research Funding program – Research Chairs (ORF-RC-2025-5503), King Saud University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The Authors do not declare any conflict of interest.

## 7. 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.

Awad, M., Khanna, R., 2015. Support vector regression, in Efficient learning machines: Theories, concepts, and applications for engineers and system designers, Ed. Berkeley, CA: Apress, 67-80. https://doi.org/10.1007/978-1-4302-5990-9_4

Basheer, I.A., Hajmeer, M., 2000. Artificial neural networks: fundamentals, computing, design, and application. J. Microbiol. Methods 43, 3-31. https://doi.org/10.1016/S0167-7012(00)00201-3

Bera, P., Prasher, S.O., Patel, R.M., Madani, A., Lacroix, R., Gaynor, J.D., *et al.*, 2006. Application of MARS in simulating pesticide concentrations in soil. Trans. ASABE 49(1): 297-307. https://doi.org/10.13031/2013.20228

Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. Comput. Geosci. 10, 191-203. https://doi.org/10.1016/0098-3004(84)90020-7

Carrión, P., Tarjuelo, J., Montero, J., 2001. SIRIAS: a simulation model for sprinkler irrigation. Irrigation Sci. 20(2), 73-84. https://doi.org/10.1007/s002710000031

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. https://doi.org/10.1016/j.jhydrol.2010.01.016

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. https://doi.org/10.1623/hysj.48.3.349.45288

Cigizoglu, H.K., Kisi, Ö., 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol. 317(3-4), 221-238. https://doi.org/10.1016/j.jhydrol.2005.05.019

Dawson, W.C., Wilby, R., 1998. An artificial neural networks approach to rainfall runoff modeling. Hydrol. Sci. J. 43(1), 47-66. https://doi.org/10.1080/02626669809492102

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*). Irrigation Sci. 22(2), 67-77. https://doi.org/10.1007/s00271-003-0071-9

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the em algorithm. J. Royal Statistical Society. Series B (methodological), 1-38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

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. https://doi.org/10.1002/ep.10295

Emamgholizadeh, S., Mohammadi, B., 2021. New hybrid nature-based algorithm to integration support vector machine for prediction of soil cation exchange capacity. Soft Computing 25(21), 13451-13464. https://doi.org/10.1007/s00500-021-06095-4

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. https://doi.org/10.1061/(ASCE)1084-0699(2009)14:3(286)

Friedman, J.H., 1991. Multivariate adaptive regression splines (with Discussion). Annals of Statistics 19(1), 1-67. https://doi.org/10.1214/aos/1176347963

Friedman, J.H., Roosen, C.B., 1995. An introduction to multivariate adaptive regression splines. Statistical Methods in Medical Research 4(3), 197-217. https://doi.org/10.1177/096228029500400303

Gunn, S., 1997. Support vector machines for classification and regression. ISIS Technical Report, University of Southampton, 1-42.

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. Computational Materials Sci. 102, 304-313. https://doi.org/10.1016/j.commatsci.2015.02.045

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., *et al.*, 2019. Multivariate adaptive regression splines for estimating riverine constituent concentrations. Hydrological Processes 34(5),1-15. https://doi.org/10.1002/hyp.13669

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 System Sci. 16(11), 4417-4433. https://doi.org/10.5194/hess-16-4417-2012

Izadifar, Z., 2010. Modelling and Analysis of Actual Evapotranspiration Using DataDriven and Wavelet Techniques Master Thesis. University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

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. https://doi.org/10.1029/2003WR002355

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. https://doi.org/10.1002/hyp.6819

Jang, J.-S.R., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. 23, 665-685. https://doi.org/10.1109/21.256541

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. https://doi.org/10.1109/TAC.1997.633847

Jensen, J.L.W.V., 1906. Sur les fonctions bconvexes et les ingalits entre les valeurs moyennes. Acta Mathematica 30(1), 175-193. https://doi.org/10.1007/BF02418571

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. Hydrologic Eng. 18(8), 1031-1040. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000707

Keller, J., Bliesner. R.D., 1990. Sprinkler and trickle irrigation. Van Nostrand Reinhold, New York. https://doi.org/10.1007/978-1-4757-1425-8

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 Eng. Annual International Meeting 2012, American Society of Agricultural and Biological Engineers, Dallas, Texas, USA. Paper No. 121336891.

Kumar, D.N., Raju, K.S., Sathish, T., 2004. River flow forecasting using recurrent neural networks. Water Resour. Management. 18(2), 143-161. https://doi.org/10.1023/B:WARM.0000024727.94701.12

Kuzman, B., Petković, B., Denić, N., Petković, D., Ćirković, B., Stojanović, J., *et al.*, 2021. Estimation of optimal fertilizers for optimal crop yield by adaptive neuro fuzzy logic. Rhizosphere 18, 100358. https://doi.org/10.1016/j.rhisph.2021.100358

Lakovic, N., Khan, A., Petković, B., Petkovic, D., Kuzman, B., Resic, S., *et al.*, 2021. Management of higher heating value sensitivity of biomass by hybrid learning technique. Biomass Conv. 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. https://doi.org/10.1155/2017/4194251

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. https://doi.org/10.1016/j.proeng.2011.11.2410

Li, Y., Bai, G., Yan, H., 2015. Development and validation of a modified model to simulate the sprinkler water distribution. Computers and Electronics in Agriculture 111, 38-47. https://doi.org/10.1016/j.compag.2014.12.003

MacKay, D.J.C., 1992. The evidence framework applied to classification networks. Neural Computation (4), 720-736. https://doi.org/10.1162/neco.1992.4.5.720

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. https://doi.org/10.1002/ird.2182

Mattar, M.A., Alamoud, A.I., 2015. Artificial neural networks for estimating the hydraulic performance of labyrinth-channel emitters. Comput. Electron. Agric. 114 (5), 189-201. https://doi.org/10.1016/j.compag.2015.04.007

Mattar, M.A., Alazba, A.A., Zin El-Abedin, T.K., 2015. Forecasting furrow irrigation infiltration using artificial neural networks. Agric. Water Manag. 148 (1), 63-71. https://doi.org/10.1016/j.agwat.2014.09.015

Mattar, M.A., 2018. Using gene expression programming in monthly reference evapotranspiration modeling: A case study in Egypt. Agric. Water Manag. 198, 28-38. https://doi.org/10.1016/j.agwat.2017.12.017

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. https://doi.org/10.1002/ird.648

Montero, J., Tarjuelo, J.M., Carrión, P., 2003. Sprinkler droplet size distribution measured with an optical spectropluviometer. Irrigation Sci. 22(2), 47-56. https://doi.org/10.1007/s00271-003-0069-3

Permai, S.D., Tanty, H., 2018. Linear regression model using bayesian approach for energy performance of residential building. Procedia Computer Sci. 135, 671-677. https://doi.org/10.1016/j.procs.2018.08.219

Petković, D., Gocic, M., Trajkovic, S., Milovančević, M., Šević, D., 2017. Precipitation concentration index management by adaptive neuro-fuzzy methodology. Climatic Change 141, 655-669. https://doi.org/10.1007/s10584-017-1907-2

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. Agricultural Water Manag. 76(3), 139-159. https://doi.org/10.1016/j.agwat.2005.01.015

Playán, E., Zapata, N., Faci, J.M., Tolosa, D., Lacueva, J.L., *et al.*, 2006. Assessing sprinkler irrigation uniformity using a ballistic simulation model. Agric. Water Manag. 84(1-2), 89-100. https://doi.org/10.1016/j.agwat.2006.01.006

Roy, D.K., Biswas, S.K., Mattar, M.A., El-Shafei, A.A., Murad, K.F.I., Saha, K.K., *et al.*, 2021. Groundwater level prediction using a multiple objective genetic algorithm-grey relational analysis based weighted ensemble of ANFIS models. Water 13, 3130. https://doi.org/10.3390/w13213130

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. https://doi.org/10.1080/10286608.2018.1431777

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. https://doi.org/10.1016/j.agwat.2016.12.009

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. https://doi.org/10.1016/j.biosystemseng.2015.04.011

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. https://doi.org/10.1080/19443994.2012.664683

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. https://doi.org/10.1016/j.agwat.2011.10.002

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. Agricultural Water Manag. 97(10), 1591-1601. https://doi.org/10.1016/j.agwat.2010.05.013

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. https://doi.org/10.1016/j.agwat.2019.04.028

Sarwar, A., Peters, R.T., Shafeeque, M., Mohamed, A., Arshad, A., Ullah, I., *et al.*, 2021. Accurate measurement of wind drift and evaporation losses could improve water application efficiency of sprinkler irrigation systems − A comparison of measuring techniques. Agricultural Water Manag. 258. 107209. https://doi.org/10.1016/j.agwat.2021.107209

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. https://doi.org/10.1007/s00271-012-0389-2

Sugeno, M., 1985. Industrial Applications of Fuzzy Control Elsevier Science Inc: pp 269.

Supharatid, S., 2003. Application of a neural network model in establishing a stage-discharge relationship for a tidal river. Hydrolog. Processes 17(15), 3085-3099. https://doi.org/10.1002/hyp.1278

Swingler, K., 2001. Applying Neural Networks, A Practical Guide, 3rd ed. Academic Press, San Francisco, CA, USA.

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. https://doi.org/10.1016/j.jhydrol.2019.06.058

Thirumalaiah, K., Deo, M.C., 1998. River stage forecasting using artificial neural networks. J. Hydrol. Eng. ASCE 3(1), 26-32. https://doi.org/10.1061/(ASCE)1084-0699(1998)3:1(26)

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory, Springer, New York. https://doi.org/10.1007/978-1-4757-2440-0

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.

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. https://doi.org/10.1016/j.eswa.2010.10.046

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. https://doi.org/10.1016/j.biosystemseng.2010.05.008

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. ASCE 133(2), 83-89. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:2(83)

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. Agricultural Water Manag. 87(2), 187-199. https://doi.org/10.1016/j.agwat.2006.06.018

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. https://doi.org/10.1007/s10333-020-00825-4