Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2023.29531

293 (2023) 67–81
May

Linear and non-linear ensemble concepts for pan evaporation modeling

Jazuli Abdullahi[a,*], Gözen Elkiran[b], Fidan Aslanova[c], Derin Orhon[d]

[a]*Center for Clean Energy and Climate Change, Baze University, Plot 686 Cadastral Zone COO, Abuja, Nigeria, email: jazuli.abdullahi@bazeuniversity.edu.ng*
[b]*Department of Civil Engineering, Faculty of Civil and Environmental Engineering, Near East University, Nicosia, Mersin 10, Turkey, email: gozen.elkiran@neu.edu.tr*
[c]*Department of Environmental Sciences and Engineering, Faculty of Civil and Environmental Engineering, Near East University, Nicosia, Mersin 10, Turkey, email: fidan.aslanova@neu.edu.tr*
[d]*Department of Environmental Engineering, Faculty of Civil Engineering, Istanbul Technical University, The Science Academy, Istanbul, Turkey, email: orhon@itu.edu.tr*

Received 1 December 2022; Accepted 3 April 2023

ABSTRACT

Modeling of pan evaporation (Ep) is of paramount importance in the evaluation of drinking water supplies, planning of regional water resources and reservoir management. The main aim of this study is to investigate the accuracy of linear and non-linear ensemble approaches for monthly Ep modeling in Erbil and Salahaddin meteorological stations of Iraq. For this purpose, sensitivity analysis was performed to determine the dominant input parameters. The results showed that T_{mean}, T_{max} and T_{min} are the most effective parameters. Thereafter, two scenarios were involved for the Ep modeling. In scenario 1, the ability of artificial neural network, least-squares support-vector machine and multiple linear regression models was examined for the estimation of Ep. The results demonstrated that different input combinations led to different performance, model 3 (which has T_{mean}, T_{max}, T_{min}, R_H) for Erbil station and model 2 (which has T_{mean}, T_{max}, T_{min}) for Salahaddin station provided the best performance among several models developed. In scenario 2, linear and non-linear ensemble approaches were employed as simple linear average, weighted linear average and non-linear ensemble (NLE) models to improve predictions of the single models. The results reported that ensemble modeling could improve performance of single models and NLE model provided the best results due to its non-linear nature. The general results demonstrated that the proposed ensemble models could improve predictions of single models up to 5% and 16% for Erbil and Salahaddin stations, respectively.

Keywords: Pan evaporation; Artificial neural network; Ensemble modeling; Erbil; station

1. Introduction

Estimation of evaporation with reliable accuracy is very crucial for reservoir control, regional water resources planning, drought management and domestic water supplies [1]. For irrigation systems and various water resources planning, water loss due to evaporation should be a well-thought-out issue. In scarce rainfall areas, the water loss by evaporation for a reservoir or lake can constitute huge amount of water budget, and tremendously contributes to dropping of surface water level [2].

For Erbil and Salahaddin stations in the Kurdistan region, the climate is semiarid that characterized by high temperatures and decline in precipitation amount with visible negative effects that include vegetation cover desiccation and reduced surface water amongst others [3].

* Corresponding author.

Consequently, proper estimation of evaporation loss from the water body in such climate regions has essential importance for water resources allocation and monitoring at regional as well as at farm scales [4].

Direct and indirect approaches are generally the two methods used for calculating or predicting evaporation. The direct method employs the use of instruments for measuring the evaporation (such as Ep); however, practical issues such as maintenance and measurement errors as well as instrumental limits may deter the efficiency of the evaporation measurements [5]. Hence, for the evaporation prediction, several methods using observed meteorological parameters have been proposed, by modeling the relationship linearly between Ep and meteorological data (including solar radiation, sunshine hours, air pressure, relative humidity, air temperature, etc.) [5].

The evaporation process has an intricate stochastic characteristic which cannot be simulated sufficiently by the linear or empirical modeling techniques and thus, can substantially amalgamate the prediction errors [6]. Moreover, the coefficients of the empirical methods must be calibrated before their application to various agroclimatic zones as under different conditions, they possess different behavior [7]. The evaporation process is yet non-linear, unsteady, incidental and complex [2]. Therefore, driven accurate relationship that will represent the physical processes involved between climatic parameters and Ep is difficult to be achieved [8]. Consequently, the use of non-linear data driven methods for hydrological modeling studies on Ep have been emphasized by many researchers [9].

In the last decades, artificial intelligence (AI) methods such as artificial neural network (ANN) and least-squares support-vector machine (LS-SVM) have been successfully applied for Ep modeling [2,10]. For instance, Rahimikhoob [11] estimated Ep on daily basis in a semiarid environment using ANN as a function of air temperature in the southwest of Iran. Shirsath and Singh [12] applied ANN, multiple linear regression (MLR) and climate-based models for daily Ep estimation. Kisi [2] applied LS-SVM, multivariate adaptive regression splines and M5 model tree for Ep modeling in Antalya and Mersin stations of Turkey. Wang et al. [5] used four heuristic approaches including least squares support vector regression (LS-SVR) and MLR for daily Ep estimation in Dongting Lake Basin, China. Qasem et al. [13] modeled monthly Ep using ANN, support vector regression (SVR) and their hybrid forms. Chen et al. [14] investigated the performance of support vector machine (SVM) in modeling monthly Ep in Three Gorges Reservoir Area, China. However, predictions by AI techniques are affected by the quality standard of the used data, implying that flawed dataset could lead to unreliable predictions by AI models. According to Zhang et al. [15], for a successful application of soft computing methods, high quality datasets with well extracted features that are closely related with the dependent responses are critical. Estévez et al. [16] applied range test and other four quality control procedures to ascertain the quality and validity of meteorological datasets.

Though, quite reasonable and reliable results could be achieved by the mentioned black box models (ANN, LS-SVM and MLR) using dataset of a standard quality, it is apparent that different outcomes could be resulted from different models for a particular problem. Thus, by assembling different techniques, more accuracy with less error would be accomplished than application of sole method [17]. Study by Makridakis et al. [18] also revealed that enhancement of forecasting accuracy through combination of numerous single models has become a common practice. According to Kiran and Ravi [19], the overall idea of ensemble modeling is the presentation of dataset in different pattern through combination of outputs from different models in a unique framework.

Ensemble modeling has been applied recently in different fields of hydrology, hydro-environmental and hydro-climatological studies. For instance, Sharghi et al. [20] performed earth fill dam seepage analysis by employing ensemble approaches to improve performance of AI models. Nourani et al. [21] modeled reference evapotranspiration (ET_0) at several meteorological stations in Turkey, Cyprus, Iraq, Iran and Libya using ensemble-based modeling approaches. Nourani et al. [22] investigated the capability of ensemble models in improving the prediction accuracy of AI based models for daily global solar radiation modeling across four stations in Iraq. Nourani et al. [23] examined the potential of ensemble learning to improve AI based predictions of single and multi-step ahead ET_0 process in different climatic regions. Up-to-date scrutiny of the current literature indicates that no ensemble modeling study was performed for evaporation process.

The primary aim of this study was to apply linear and non-linear ensemble concepts to enhance prediction of AI based models for Ep modeling in Iraq. The data obtained were validated using range (fixed) test method of quality control procedures. Due to the importance of appropriate input selection for AI based modeling, firstly, sensitivity analysis was performed to determine the dominant variables. Then, the Ep modeling was performed in scenarios 1 and 2. In scenario 1, ANN, LS-SVM and MLR were trained and validated separately for monthly Ep modeling in Erbil and Salahaddin stations. Scenario 2 involved the application of simple linear average (SLA), weighted linear average (WLA) and non-linear ensemble (NLE) methods to improve performance of the single models.

2. Materials and method

2.1. Study locations and data

Erbil is the largest and the capital city of Kurdistan region in northern Iraq. Its location is within a continental semiarid climate. Erbil experiences cool and rainy winters with warm and dry summers [21]. Erbil governorate estimated population in 2010 was 1,820,000 whereas the city population was 852,000. The Erbil district population density in terms of persons/km^2 was 472.9 [24]. Salahaddin city is also located in Kurdistan region in further north of Iraq. The climate of Salahaddin is considered semiarid according to Şarlak and Agha [25] study. Fig. 1 shows map of Iraq and the respective study stations.

The data used in this study were of 20 years duration (1992–2011), which were measurements of daily values averaged over the month including pan evaporation (Ep) (mm/month), maximum air temperature (T_{max}) (°C), minimum

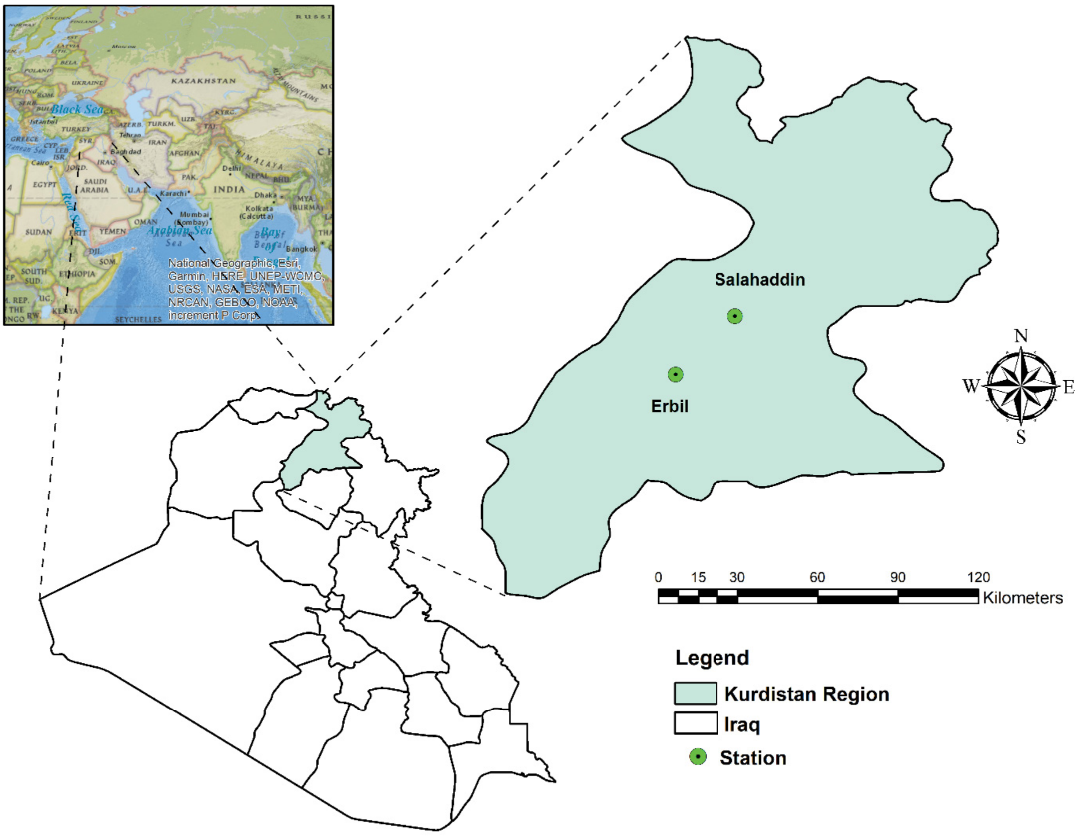

Fig. 1. Study country and location of the study stations.

air temperature (T_{min}) (°C), mean air temperature (T_{mean}) (°C), relative humidity (R_H) (%), vapor pressure (V_p) (mbar) and wind speed (U_2) (m/s) obtained for each study station from general directorate of dams and reservoirs, Kurdistan. Table 1 gives statistical description of the used data.

As seen in Table 1, Erbil station is more semiarid which has T_{max} as high as 45°C and R_H as low as 5% whereas Salahaddin is more humid with T_{max} as 39.9°C and maximum R_H of 92%. Evaporation is less in Salahaddin station as minimum Ep value of 0.1 mm/month could be seen. This is because of the dryness of the Erbil land coupled with high temperature which increases the rate of Ep for Erbil station. To determine the effect and correlation of each variable on the target, Pearson correlation matrix was developed. Table 2 provides the results of the used correlation matrix.

The results shown in Table 2 demonstrated that the correlation between Ep and temperatures is directly proportional, implying that as the temperature increases the rate of Ep increases and vice versa. This is why the temperatures (T_{mean}, T_{max} and T_{min}) in Erbil station have higher correlation than in Salahaddin station. Among all the variables, R_H was found to be less correlated with Ep compared to the rest.

2.2. Data normalization and performance criteria

At initial stage, the data were normalized to eliminate the dimensions of inputs and output and to ensure equal attention is given to all variables. The data were normalized between 0 and 1 according to Elkiran et al. [26] as;

$$\mathrm{Ep}_{\mathrm{norm}} = \frac{\mathrm{Ep}_i - \mathrm{Ep}_{\min}}{\mathrm{Ep}_{\max} - \mathrm{Ep}_{\min}} \tag{1}$$

where $\mathrm{Ep}_{\mathrm{norm}}$, Ep_i, $\mathrm{Ep}_{\max}$ and $\mathrm{Ep}_{\min}$, respectively are the normalized, observed, maximum and minimum values of Ep.

To determine the accuracy of the applied models, this study endorsed Legates and McCabe [27] study which suggested that for any hydroclimatic model, determination Coefficient or Nash–Sutcliffe efficiency criterion (NSE) and root-mean-square error can be sufficient for performance evaluation [21], given by:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}\left(\mathrm{Ep}_i - \hat{\mathrm{Ep}}_i\right)^2}{\sum_{i=1}^{N}\left(\mathrm{Ep}_i - \overline{\mathrm{Ep}}_i\right)^2} \tag{2}$$

$$\mathrm{RMSE} = \frac{\sum_{i=1}^{N}\left(\mathrm{Ep}_i - \hat{\mathrm{Ep}}_i\right)^2}{N} \tag{3}$$

where Ep_i has been defined, N, $\hat{\mathrm{Ep}}_i$, and $\overline{\mathrm{Ep}}_i$ are the number of observations, predicted values and mean of the observed values, respectively. The NSE ranges between –∞ to 1 and root-mean-square error (RMSE) between 0 to ∞ with NSE towards 1 and RMSE close to 0 imply high efficiency.

Table 1
Data descriptive statistics of the used variables including pan evaporation (Ep) (mm/month), maximum air temperature (T_{max}) (°C), minimum air temperature (T_{min}) (°C), mean air temperature (T_{mean}) (°C), relative humidity (R_H) (%), vapor pressure (V_p) (mbar), Ep and wind speed (U_2) (m/s) for 1992–2011 study period

Station	Variable	Minimum	Maximum	Mean	Std. deviation
	T_{mean}, °C	6	37.3	21.28	9.37
	T_{max}, °C	9.5	45	27.5	10.63
	T_{min}, °C	0.6	30	15.07	8.26
Erbil	R_H, %	5	88	46.73	18.78
	V_p, mbar	3.5	18.3	11.12	2.57
	U_2, m/s	1	7	2.5	0.8
	Ep, mm/month	1	16	6.84	4.39
	T_{mean}, °C	0	34.6	18.02	9.27
	T_{max}, °C	0	39.9	22.26	10.29
	T_{min}, °C	–1.6	29.2	13.35	8.57
Salahaddin	R_H, %	24	92	52.27	16
	V_p, mbar	4.7	20.1	10.46	3.58
	U_2, m/s	1	4	2.36	0.64
	Ep, mm/month	0.1	15.5	4.99	3.48

Table 2
Results of the applied correlation matrix

Station	Variable	Ep (mm/month)	T_{mean}, °C	T_{max}, °C	T_{min}, °C	R_H, %	V_p, mbar	U_2, m/s
	Ep (mm/month)	1						
	T_{mean}, °C	0.95187	1					
	T_{max}, °C	0.947304	0.99118	1				
Erbil	T_{min}, °C	0.948491	0.989551	0.971483	1			
	R_H, %	–0.86684	–0.88408	–0.89572	–0.86305	1		
	V_p, mbar	0.725395	0.778229	0.762167	0.792531	–0.54576	1	
	U_2, m/s	0.000203	–0.03324	–0.02856	–0.04811	0.027797	0.001558	1
	Ep (mm/month)	1						
	T_{mean}, °C	0.886293	1					
	T_{max}, °C	0.887104	0.981548	1				
Salahaddin	T_{min}, °C	0.903824	0.982538	0.990335	1			
	R_H, %	–0.88738	–0.87041	–0.88985	–0.88183	1		
	V_p, mbar	0.786193	0.871481	0.873193	0.889176	–0.69521	1	
	U_2, m/s	0.240498	0.098463	0.074705	0.106517	–0.16363	0.042088	1

2.3. Quality control tests

To determine the validity of the meteorological data used in this study, quality control measures were utilized to ascertain the erroneous and suspect data from observations. Initially, to ensure that all possible data have been collected with correct and complete record structure, verification is necessary. Gaps detected in the data files would be flagged as erroneous, and should not be used as input variable in the estimations [21]. Several methods can be found in the literature for quality assurance of meteorological parameters. These include step test, range (fixed or dynamic) test, persistence test and internal consistency test [16,28]. Range (fixed) test method was selected and applied in this study, owing to its ability to detect erroneous data (data outside acceptable fixed range). Table 3 shows the applied range test procedures for data quality control of the variables used.

2.4. Proposed methodology

The Ep modeling in this study was conducted in two scenarios;

- Scenario 1

In the first scenario, AI based and MLR models were applied for modeling Ep in two meteorological stations in Iraq. The dependent variable (Ep) was used as a function of the independent variables as follows:

Table 3
Data validation procedures for meteorological variables before their use as input data for Ep estimations

Variable	Unit	Applied data validation procedure
Relative humidity (R_H)	%	$0.8 < R_H < 103$ [16]
Surface pressure (S_p)	Kpa	$80 < S_p < 105$ [29]
Maximum air temperature (T_{max}) Minimum air temperature (T_{min}) Mean air temperature (T_{mean})	(°C)	$-20 < T_{max}$, T_{min}, $T_{mean} < 50$ [28]
Mean wind speed (U_{mean}) Maximum wind speed (U_{max}) Minimum wind speed (U_{min})	m/s	$0 < U_{max} < 100$ $0 < U_{mean}$ [16]
Pan evaporation (Ep)	mm	$0 \leq \text{Ep} < 500$ [30]

$$\text{Ep}^s = f\left(T^s_{min}, T^s_{mean}, T^s_{max}, V^s_p, U^s_2, R^s_H\right) \tag{4}$$

where the superscript s (such as in Ep^s) represents the station (e.g., Erbil or Salahaddin), T_{min}, T_{mean}, T_{max}, V_p, U_2 and R_H were previously defined.

- Scenario 2

The concepts of ensemble modeling were applied for accuracy improvement of the single models in the second scenario, where the single models output were used as inputs to the ensemble models as:

$$\text{Ep}^s = f\left(\text{ANN}^s_{\text{Ep}}, \text{LLSVM}^s_{\text{Ep}}, \text{MLR}^s_{\text{Ep}}\right) \tag{5}$$

where ANN^s_{Ep}, $\text{LLSVM}^s_{\text{Ep}}$, MLR^s_{Ep} are the Ep outputs produced by ANN, LS-SVM and MLR models.

The general methodology employed in this study is given in Fig. 2. For proper comparison, same methodology is applied to the data from both Erbil and Salahaddin stations.

2.5. Model validation

A stratified k-fold cross validation was applied in this study. The main advantages of using this validation approach over hold-out validation approach (which uses single test set per station) are that both training and validation are done by all observations and each observation is used exactly once for the model validation [21]. The data were randomly divided in to 4-fold of equal subsamples. The model was trained using ¾ of the subsamples while the remaining ¼ was used for testing the model. The procedure was repeated 4 times (the number of subsamples), in each case, different k–1 (4–1) subsamples were used for training and the remaining subsample for testing the model.

2.6. Artificial neural network

ANN provides a determined approach in dealing with non-linear, noisy, and dynamic data, more specifically when the physical fundamental relationship are not completely known [10].

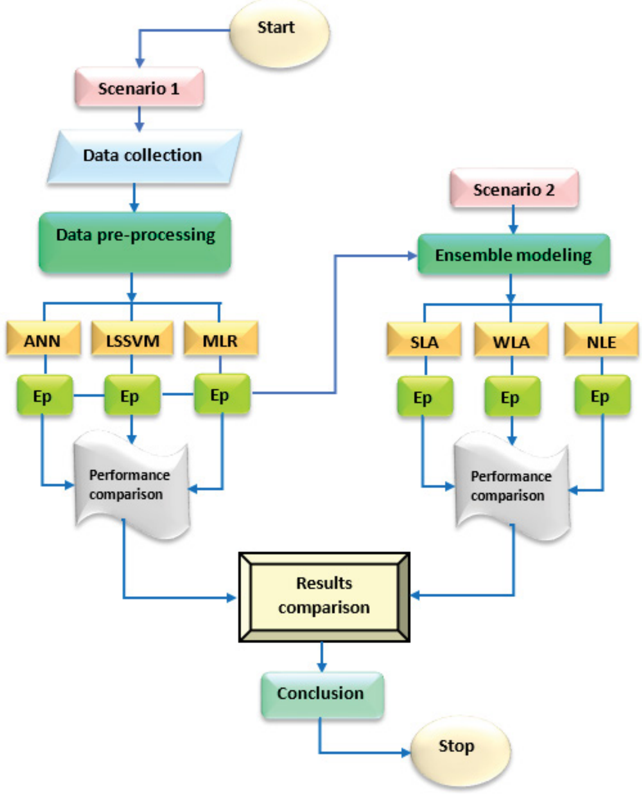

Fig. 2. The overall methodology of the study for data between 1992–2011 in 2 scenarios: (1) artificial neural network (ANN), least-squares support-vector machine (LS-SVM) and multiple linear regression (MLR) were used to predict monthly pan evaporation (Ep). (2) Simple linear average (SLA), weighted linear average (WLA) and non-linear ensemble (NLE) were used to enhance performance.

ANN constitutes a number of simple processing elements that are interconnected by nodes or neurons with fascinating characteristics of information processing including parallelism, non-linearity, generalization, capability, learning and noise tolerance. For solving many engineering problems, a feed forward neural network trained with back propagation (FFBP) algorithm is the most applied ANN method [31,32]. The FFBP method is comprised of layers of parallel processing elements known as neurons, with every successive layer neuron completely connected to its predecessor layer by weight [33]. BP algorithm generally accomplished this ANN learning [34]. Fig. 3 shows the FFBP structure (Ep).

2.7. Least-squares support-vector machine

The LS-SVM emerged from the learning context of SVM is a robust approach used for function estimation, classification and for solving non-linear problems [2]. The LS-SVM procedure was first proposed by Suykens and Vandewalle [35]. By considering the time series of x_i and y_i as inputs (meteorological data) and output (Ep values), the LS-SVM function as a non-linear function is expressed as;

$$f(X) = w^{T}\varphi(X) + b \tag{6}$$

where b represents bias term, ϕ is the mapping function and w is the m-dimensional weight vector [36]. Regarding structural minimization principle, the regression problem using the function estimation error can be expressed as;

$$\min J(w,e) = \frac{1}{2}w^{T}w + \frac{\gamma}{2}\sum_{i=1}^{m} e_i^2 \tag{7}$$

Which is dependent on the following constraints;

$$y_i = w^{T}\varphi(x_i) + b + e_i \qquad (i = 1,2,\ldots,m) \tag{8}$$

where e_i refers to x_i training error and γ denotes to regularization constant.

Lagrange multiplier optimal programming mechanism is applied for solving Eq. (7) to determine the solutions of w and e. By forming unconstraint problem through the modification of the constraint problem, the objective function can be achieved. The L Lagrange function is given by;

$$L(w,b,e,\alpha) = J(w,e) - \sum_{i=1}^{m} \alpha_i \left\{ w^{T}\varphi(x_i) + b + e_i - y_i \right\} \tag{9}$$

where α_i denotes to Lagrange multipliers.

Considering Karush–Kuhn–Tucker [37], by applying partial derivatives to Eq. (9), the optimal condition with respect to w, b, e, α can be obtained as;

$$\begin{cases} w = \sum_{i=1}^{m} \alpha_i \varphi(x_i) \\ \sum_{i=1}^{m} \alpha_i = 0 \\ \alpha_i = w e_i \\ w^{T}\varphi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{10}$$

After elimination of w and e_i, the linear equations can be derived as;

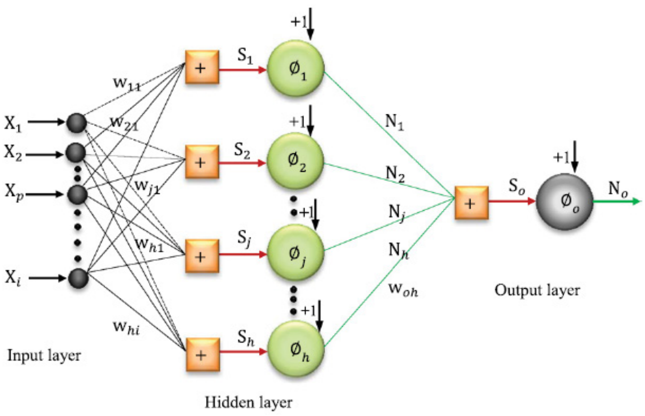

Fig. 3. A three layered FFBP structure.

$$\begin{bmatrix} 0 & -Y^{T} \\ Y & ZZ^{T} + 1/^{3} \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{11}$$

where $Y = y_1, \ldots, y_m$, $Z = \phi(x_1)^{T}y_1, \ldots, \phi(x_m)^{T}y_m$, $I = [1, \ldots, 1]$, $\alpha = [\alpha_1, \ldots, \alpha_1]$.

Kernel function can be expressed as $K(x, x_i) = \phi(x)^{T}\phi(x_i)$, $i = 1, \ldots, m$, which is satisfied with Mercer's condition. Owing to that, the LS-SVM regression becomes;

$$f(x) = \sum_{i=1}^{m} \alpha_i K(x, x_i) + b \tag{12}$$

This study employed radial basis function (RBF) kernel which is a commonly used kernel function, given as;

$$K(x, x_i) = \exp\left(-\|x - x_i\|^{2} / 2\sigma^{2}\right) \tag{13}$$

Fig. 4 shows the LS-SVM structure for pan evaporation modeling.

2.8. Multiple linear regression

Multiple linear regression (MLR) is a famous method of modeling mathematically, the linear relationship between one or more independent variables and dependent variable. In general, the dependent variable y, and n regressor variables may be related by [26]:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_ix_i + \xi \tag{14}$$

where x_i is the value of the ith predictor, b_0 is the regression constant, and b_i is the coefficient of the ith predictor and ξ is the error term.

2.9. Ensemble modeling concepts

Evaporation process, like any other natural process, may exhibit both linear and non-linear behaviors. As such, neither linear nor non-linear models could be sufficient for accurate modeling of evaporation process, because the AI models cannot deal with both linear and non-linear aspect of the system while MLR could not cope with non-linearity of the data. Therefore, application of ensemble models, which combined all models, will take care of the deficiencies

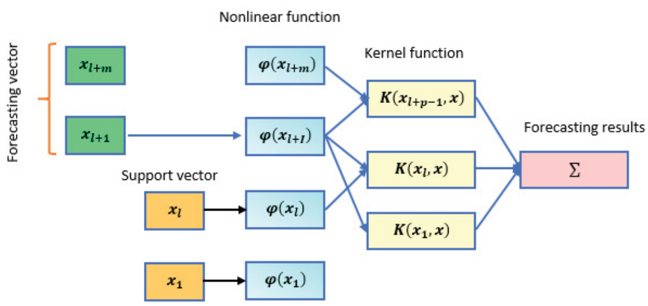

Fig. 4. Structure of LS-SVM model for Ep modeling [2].

of the single models. The ensemble approaches are categorized into two based on Kiran and Ravi [19] study, as; (i) linear ensemble approach; which includes weighted median, weighted average and linear average. (ii) non-linear ensemble approach; which involves application of non-linear model, for example, ANN.

The proposed ensemble modeling in this study includes two linear (SLA, WLA) and non-linear ensemble (NLE) approaches.

Simple linear average (SLA) is conducted as;

$$\overline{\overline{\mathrm{Ep}}} = \frac{1}{N}\sum_{i=1}^{n}\mathrm{Ep}_i \tag{15}$$

where $\overline{\overline{\mathrm{Ep}}}$ is the obtained ensemble output, Ep_i is the output produced by the *i*th model (FFBP, LS-SVM, MLR) and N is the applied number of single models.

The weighted linear average (WLA) is given as;

$$\overline{\overline{\mathrm{Ep}}} = \sum_{i=1}^{N} w_i \mathrm{Ep}_i \tag{16}$$

where w_i is the weight generated based on model performance in terms of NSE, given as;

$$w_i = \frac{\mathrm{NSE}_i}{\sum_{i=1}^{N}\mathrm{NSE}_i} \tag{17}$$

For the non-linear ensemble method, outputs of the FFBP, LS-SVM and MLR models are combined together as inputs to a new ANN model to obtain the final Ep output.

3. Results and discussion

As the general methodology of this study involves validation of the used meteorological variables, sensitivity analysis to determine dominant input variables, application of FFBP, LS-SVM (as AI based) models and conventional MLR model for single modeling in the first scenario, and utilization of SLA, WLA and NLE models in the second scenario to enhance the Ep prediction, hence, the results obtained are provided accordingly.

3.1. Range (fixed) test results

To properly validate and ensure quality standard of the variables to be used in this study, quality assurance procedures were utilized to identify erroneous values. The applied fixed range test results showed no identification of erroneous or flagged data. Indicating that all variables are within the accepted limit described in Table 3. This can also be supported by the descriptive statistics of the data given in Table 1.

3.2. Sensitivity analysis of input variables

One of the most important aspects of any AI based modeling is the appropriate selection of input variables, as failure to do so may lead to inefficient modeling. Thus, in this study, single-input single-output ANN based sensitivity analysis was conducted to determine the most dominant input variables. The results of sensitivity analysis are given in Fig. 5.

As shown in Fig. 5, the three categories of temperature (T_{mean}, T_{max}, T_{min}) are the most dominant variables in the prediction of Ep. This could be because temperature has a direct effect on evaporation process, implying that increase or decrease in temperature will lead to increase or decrease in evaporation. As such, temperature could be the main indicator for the evaporation process especially in this study stations, which have semiarid climate that constitutes higher temperature.

Visual inspection of the results depicted by Fig. 5 also shows that U_2 has the least sensitivity to Ep prediction in both stations. In other words, U_2 has the minimum effect on Ep process among the variables considered in this study. Although the rate of evaporation may increase with increase in U_2, the ineffectiveness of U_2 in Erbil and Salahaddin stations demonstrated that evaporation process is not heavily dependent on U_2. In another perspective, U_2 as a sole input variable may not have much effect on evaporation process in the two study stations. According to Nourani et al. [21] study, U_2 as a standalone variable may not have much

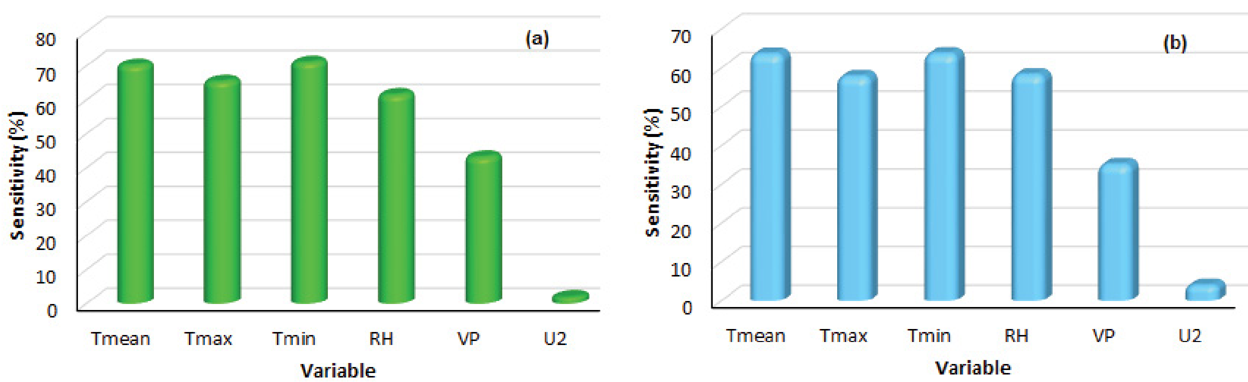

Fig. 5. Sensitivity analysis results for the study period from 1992–2011 based on maximum air temperature (T_{max}) (°C), minimum air temperature (T_{min}) (°C), mean air temperature (T_{mean}) (°C), relative humidity (R_H) (%), vapor pressure (V_p) (mbar) and wind speed (U_2) (m/s) for (a) Erbil station (b) Salahaddin station.

effect on evaporation and transpiration, but significantly increases prediction accuracy when combined with other variables. This is supported by Wang et al. [5] study which shows that addition of U_2 into model inputs significantly improve model accuracy in most cases despite having low correlation with Ep.

As can be seen also in Fig. 5, the first 4 variables have much impact on Ep and hence used in 3 different input combinations for the Ep modeling in both stations.

3.3. Scenario 1 results

For the ANN model in both stations, FFBP was used for the model training using Levenberg–Marquardt (LM) algorithm. Single hidden layer was used and via trial and error, the best number of hidden layer neurons was selected. For LS-SVM modeling, RBF is the most utilized kernel function [21] and hence used in this study. Lastly, based on input-output linear relationship, MLR modeling of Ep was performed. Table 4 shows the results of applied models for scenario 1.

For the structure of FFBP model in Table 4, x–y–z represent the number of inputs, number of hidden layer nodes and output. For LS-SVM, RBF is the kernel function used in LS-SVM construction and *x–y* for MLR model signify inputs and output number of variables, respectively.

In view of the obtained results in Table 4, it is apparent that all the applied models could lead to acceptable Ep modeling.

For Erbil station, the performance of the models in terms of NSE and RMSE are up to 0.9268 and 0.0820 for FFBP model, 0.8920 and 0.0899 for LS-SVM model and 0.8836 and 0.0941 for MLR model in the validation phase for the best performance models. Also from Table 4, it can be seen that the AI models have superior performance over MLR model both in training and validation phases. This could be attributed to the ability of AI techniques to deal with complex and non-linear Ep process.

Comparing the results in Table 4 and results obtained by Goyal et al. [38] study, it can be seen that despite using less number of input variables, this study led to better prediction performance. This implies that the higher performance of the AI based prediction of Ep lies on the proper method employed for the selection of input variables, less complexity, reduced uncertainty and a lot of time could be saved with better performance when the required inputs are used.

Similarly, the results of this study outperformed the empirical method of Priestley–Taylor, MLR and ANN based simulation developed by Bruton et al. [39] for the prediction of daily Ep at Watkinsville, Georgia. Bruton et al. [39] study obtained the highest performance by ANN with R^2 value of 0.717 and RMSE value of 1.11 mm whereas, the best performance in this study was achieved by FFBP with R^2 value of 0.9268 and RMSE of 0.0820 (normalized) in the validation phase. In addition to robust inputs selection technique applied that enhanced performance in this study, the time scale of the data may have influence on the results of the Ep prediction. With longer study period (1992–2011),

Table 4
Results of pan evaporation (Ep) modeling for the study period 1992–2011 in the first scenario based on model 1 (M1), model 2 (M2), model 3 (M3) using feed-forward back propagation (FFBP), least-squares support-vector machine (LS-SVM) and multiple linear regression (MLR) models as analyzed by Nash–Sutcliffe efficiency criterion (NSE) and root-mean-square error (RMSE) performance evaluation criteria

Station	Model	Model no.	Input	Structure	Training		Validation	
					NSE	RMSE	NSE	RMSE
		M1	T_{min}, T_{mean}	2-7-1	0.9215	0.0803	0.9042	0.0938
	FFBP	M2	T_{max}, T_{min}, T_{mean}	3-6-1	***0.9444***	***0.0675***	0.8979	0.0969
		M3	T_{max}, T_{min}, T_{mean}, R_H	4-10-1	0.9385	0.0711	***0.9268***	***0.0820***
		M1	T_{min}, T_{mean}	RBF	0.9236	0.0792	0.8988	0.0964
Erbil	LS-SVR	M2	T_{max}, T_{min}, T_{mean}	RBF	0.9282	0.0768	0.9098	0.0910
		M3	T_{max}, T_{min}, T_{mean}, R_H	RBF	0.9290	0.0763	0.9120	0.0899
		M1	T_{min}, T_{mean}	2-1	0.9104	0.0858	0.8915	0.0998
	MLR	M2	T_{max}, T_{min}, T_{mean}	3-1	0.9110	0.0855	0.8906	0.1003
		M3	T_{max}, T_{min}, T_{mean}, R_H	4-1	0.9122	0.0849	0.9036	0.0941
		M1	T_{min}, T_{mean}	2-5-1	0.8519	0.0648	0.7356	0.1236
	FFBP	M2	T_{max}, T_{min}, T_{mean}	3-10-1	***0.8612***	***0.0627***	***0.8004***	0.1074
		M3	T_{max}, T_{min}, T_{mean}, R_H	4-8-1	0.8192	0.1022	0.7699	***0.0808***
		M1	T_{min}, T_{mean}	RBF	0.8111	0.0732	0.7775	0.1134
Salahaddin	LS-SVR	M2	T_{max}, T_{min}, T_{mean}	RBF	0.8210	0.0713	0.7788	0.1130
		M3	T_{max}, T_{min}, T_{mean}, R_H	RBF	0.8356	0.0683	0.7776	0.1134
		M1	T_{min}, T_{mean}	2-1	0.7880	0.1107	0.7235	0.0886
	MLR	M2	T_{max}, T_{min}, T_{mean}	3-1	0.7888	0.1104	0.7249	0.0883
		M3	T_{max}, T_{min}, T_{mean}, R_H	4-1	0.7894	0.1103	0.6943	0.0931

Data have been normalized, so RMSE has no unit.

in this study, better simulation of the complex and uncertain phenomenon surrounding Ep might be achieved than shorter period (1992–1996) by Bruton et al. [39] study.

However, a close performance can be seen when this study results are compared with Simon-Gáspár et al. [40] study. Their predicted Ep by MLR method has maximum R^2 value of 0.62 in comparison with 0.9036 of this study. In contrast, the results obtained using Kohonen self-organizing map (K-SOM) has superior performance with maximum R^2 value of 0.98 in comparison to R^2 value of 0.9268 of this study. The similarity as well as the close performance of the two studies could be attributed to the selection of the best input variables where the sensitivity analysis of both the studies showed better correlation of temperature variables and poor correlation to R_H to Ep (Table 2).

In terms of the performance comparison between AI based and MLR models, this study is in agreement with that carried out by Wang et al. [9], where they found that across all stations, MLR model has the least Ep prediction accuracy. The reason could be due to the fact that MLR is a linear model whereas Ep may contain both linear and non-linear behaviors, hence, the MLR might generate errors from the non-linear aspect of Ep and thus, less efficiency is achieved. However, Wang et al. [9] study showed a relatively superior results to this study. Many reasons could lead to the difference, some of them include; (1) It has been proven by previous studies such as Nourani et al. [21] that there is no specific model to be employed which can perfectly simulate the underline behavior of a real world problem. (2) Wang et al. [9] included sunshine durations (HS) variable in their models development which was not available for application in this study. This may play a significant role in the difference in results of the two separate studies as HS is obviously amongst the most significant factors influencing evaporation process. (3) Another factor that is of great significance is aridity index of the study stations. In this study, both Erbil and Salahaddin are a semiarid climate stations that are characterized by high temperature and low precipitation amount. As demonstrated by many studies [5], factors affecting the rate of evaporation are the climatic variables. With the complex nature of these factors in semiarid regions, the evaporation process would be uncertain and complex. Hence, predicting and investigating the phenomenon surrounding the evaporation process in these stations would be tedious and highly competitive. This led to less predictive efficiency of the results of this study compared to Wang et al. [9] study.

The results of this study can also be compared with the results of Wang et al. [41] study. It can be vividly seen that the performances of the models are comparable for the 2 studies despite fluctuations of inferiority and superiority of one model over another and vice versa. The results similarity can be connected to the number of inputs as both studies have a maximum of 4-input variables. The little disparity observed could be due to the fact that different methods were involved to simulate the Ep process, as every technique has its unique step to follow in model development and every technique has different generalization capability.

Among the AI models in Table 4, FFBP is found to have better prediction accuracy, though fluctuations could be observed such as in M2 where better performance is achieved by LS-SVM model using both NSE and RMSE performance indicators in the validation phase. Many reasons could be associated to this development some of which include;

- Time series prediction involved complex and uncertain behavior of a system due to the huge amount of data used for a long period of time, which could be affected by seasonality, non-stationarity and missing data. This could result in increase and decrease (or rise and fall) of the observations, which in turn might lead to failure of a particular model to capture all the aspects of the data efficiently. As such, one model may perform better at certain stage and inferior at another stage of the modeling.
- Though the applied AI models are both non-linear in nature, but their methodologies of application as well as the training parameters are quite different, thus an adjustment of a particular parameter may increase accuracy of one model and could be deterrent to another.

Also, by visual inspection of the results for Erbil station, it can be seen that for all the applied models, the performance of the developed models increases as the number of input increases. This shows that evaporation process has a complex stochastic nature which its accurate prediction requires several climatological variables and depends on many factors. Despite the existence of strong correlation between Ep and temperatures, inclusion of R_H increased efficiency of the Ep modeling. For example, comparing M1 (which has only $T_{\min}$ and T_{mean}) with M3 (which has $T_{\max}$, $T_{\min}$, T_{mean} and R_H) a difference in performance in terms of NSE up to 3% could be achieved for FFBP models in the validation phase. Fig. 6 shows the time series plot of the best models for Erbil station in the validation phase.

For Salahaddin station, being both the stations have semiarid climate, the results for Salahaddin station show similar characteristics to the results obtained for Erbil station. In contrast, M2 provided the best performance and inclusion of R_H for M3 reduced the modeling performance.

Comparing the results for Erbil and Salahaddin stations in Table 4, it can be deduced that, the applied models provided better performances in Erbil stations than in Salahaddin despite having same semiarid climate. This is because evaporation has a direct relationship with temperature. As shown in Table 1, the $T_{\max}$, $T_{\min}$, and T_{mean} are all higher in Erbil station than in Salahaddin station, hence as the temperature increases the rate of evaporation increases, hence higher Ep prediction by the models. However, behavior of the climate between the stations may lead to higher results in Erbil than Salahaddin. For instance, Şarlak and Agha [25] study shows that different aridity index and period of investigation give varied climate for Salahaddin station. Using UNEP [42] aridity index, Salahaddin station was found to be semiarid between 1998–2011, subhumid between 1980–1997 and subhumid between 1980–2011. The unrealistic nature of the climate in the station leads to inefficiency of models to give comparable performance with the results of Erbil stations. Fig. 7 shows observed vs. predicted plot of the best models in validation phase for Salahaddin station.

As can be seen in Figs. 6 and 7, two points are randomly selected in each figure. At point 1 in Erbil station,

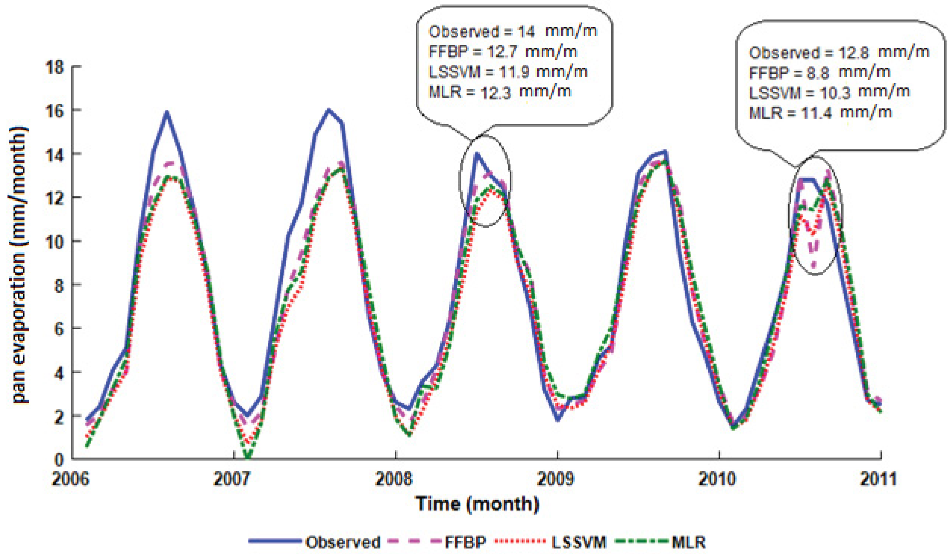

Fig. 6. Time series of the best performing models in the validation phase for Erbil station.

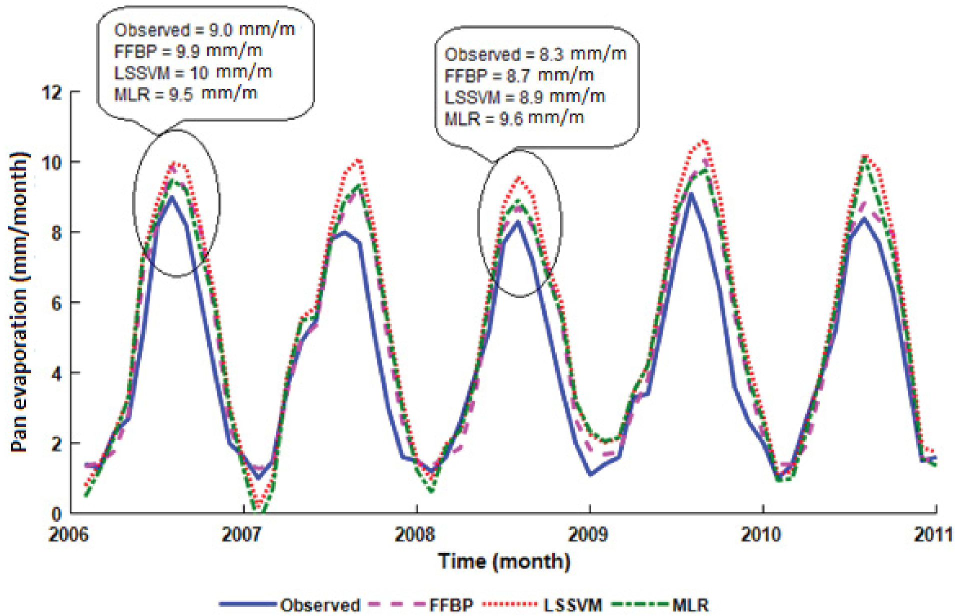

Fig. 7. Time series of the best performing models in the validation phase for Erbil station.

the observed Ep value is 14 mm/month, FFBP value is more closer to the observed value which implies better agreement between observed and predicted values and hence, better accuracy. At point 2, MLR model produced the best performance. For Salahaddin station at the first point, MLR value is more close to the observed value than the rest of the models despite having less performance when the whole data is considered (as seen in Table 3), while at point 2, FFBP has better accuracy. Based on the aforementioned development it can be deduced that, different performances could be achieved by different models at different point in time, suggesting that best performing model could be weak at certain period of a time series while the weak could produce strong performance at a given point. Thus, by assembling of the models, the gap created by the weakness of each model could be filled up. Therefore, in the next section (scenario 2) SLA, WLA and NLE approaches were applied to increase the performance of the single models.

FFBP was used as the kernel for NLE modeling. The choice of the ANN model was made based on its compatibility, accuracy and higher reported performance by recently applied ensemble studies [20–22]. Fig. 8 shows the schematic illustration of the proposed NLE concept with 3 inputs, 5 hidden neurons and an output.

3.4. Scenario 2 results

In this section, SLA, WLA and NLE approaches were applied to enhance the performance of the single models. As 3 models (M1, M2, M3) were developed by each applied technique for the single modeling, the ensemble model were produced accordingly. Table 5 shows the results of the ensemble models for Erbil station. The procedure followed for FFBP non-linear ensemble modeling is same as that of single model and the description of the model structure is same. The x–y numbers are representations of the number of inputs and output for SLA while a,b,c are the weights generated for WLA ensemble.

As demonstrated in Table 5, amalgamation of different models in form of ensemble modeling has a significance effect in Ep modeling. The applied ensemble concepts in this study have improved the performance of single models.

For Erbil station, the improvements in performance of NLE models over single models are achieved up to 2%, 4% and 1% for FFBP models, 2%, 3% and 3% for LS-SVM models and 3%, 5% and 4% for MLR models with respect to M1, M2 and M3, respectively. It is obvious from the results that SLA and WLA have comparable performance. This could be because; both the two models are derived linearly, which makes them possess similar behavior in terms of their performances. The little difference between the performances of the models is due to difference in their methodology.

For the analysis of uncertainties of ensemble predictions due to uncertainties in input data, the average performance of the ensemble concepts for each model (models 1, 2 and 3, for different set of inputs) was compared to the average performance of single models. Table 6 presents the results of the average models performance for each station.

From the results displayed by Table 6, it is obvious that the choice of input data plays an important task in Ep ensemble predictions. For instance, considering Erbil station, the average results for single models show that M2 exhibited the most reliable performance (as boldly shown) in both training and validation phases. Similarly, the ensemble models average results also show that M2 provided the most satisfactory results. Nevertheless, the results in terms of RMSE indicator show a slight under performance of M2 compared to M3 in the validation phase, which may be due to linear effect of single models on SLA and WLA models. Based on this results in Table 6, it is worthy to mention that the selection of the best input data (dominant variables) for Ep prediction is not limited to providing better model output for single modeling but rather, it also affects the performance of

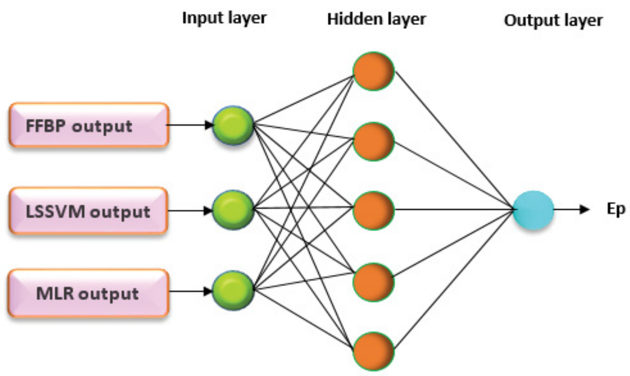

Fig. 8. Schematic illustration of the proposed non-linear ensemble concept.

Table 5
Results of the ensemble models

Station	Model	Model no.	Model structure	Training		Validation	
				NSE	RMSE	NSE	RMSE
Erbil	SLA	M1	3-1	0.9220	0.0800	0.9017	0.0950
		M2	3-1	0.9351	0.0730	0.9218	0.0848
		M3	3-1	0.9316	0.0749	0.9208	0.0853
	WLA	M1	0.3356, 0.3336, 0.3309	0.9220	0.0800	0.9018	0.0950
		M2	0.3301, 0.3372, 0.3327	0.9351	0.0730	0.9217	0.0848
		M3	0.3376, 0.3326, 0.3295	0.9317	0.0749	0.9207	0.0853
	NLE	M1	3-5-1	0.9391	0.0707	0.9163	0.0877
		M2	3-9-1	***0.9445***	***0.0675***	0.9367	0.0763
		M3	3-8-1	0.9424	0.0687	***0.9385***	***0.0752***
Salahaddin	SLA	M1	3-1	0.8175	0.0934	0.8131	0.0728
		M2	3-1	0.8172	0.0982	0.8140	0.0726
		M3	3-1	0.8075	0.0955	0.7861	0.0779
	WLA	M1	0.3289, 0.3476, 0.3235	0.8178	0.0933	0.8117	0.0731
		M2	0.3474, 0.3380, 0.3146	0.8176	0.0981	0.8143	0.0726
		M3	0.3434, 0.3469, 0.3097	0.8082	0.0953	0.7834	0.0784
	NLE	M1	3-12-1	***0.8697***	***0.0862***	***0.8672***	***0.0614***
		M2	3-9-1	0.8237	0.0909	0.8156	0.0723
		M3	3-10-1	0.8540	0.0879	0.8473	0.0658

Data have been normalized, so RMSE has no unit.

ensemble output. Fig. 9 shows the scatter plots of the ensemble models in the verification phase for Erbil station for M1, M2 and M3, respectively.

Considering the performance of ensemble models for Salahaddin station shown in Table 5, it can be deduced that the performances of the single models are improved by NLE models in the validation phase up to 13%, 2% and 8% for FFBP models, 9%, 4% and 7% for LS-SVM models and 15%, 10% and 16% for MLR models with respect to M1, M2 and M3, respectively.

Comparing the performances of the 3 ensemble models applied it can be seen that for all models, NLE provided the

Table 6
Average results for single and ensemble models

Station	Model type	Model no.	Training		Validation	
			NSE	RMSE	NSE	RMSE
Erbil	Single model	M1	0.9185	0.0818	0.8982	0.0967
		M2	***0.9279***	***0.0766***	***0.8994***	***0.0961***
		M3	0.9266	0.0774	0.9141	0.0887
	Ensemble model	M1	0.9277	0.0769	0.9066	0.0926
		M2	***0.9382***	***0.0711***	***0.9267***	0.0820
		M3	0.9352	0.0728	0.9266	***0.0819***
Salahaddin	Single model	M1	0.8170	0.0829	0.7455	0.1085
		M2	***0.8237***	***0.0815***	***0.7680***	0.1029
		M3	0.8147	0.0936	0.7472	***0.0958***
	Ensemble model	M1	***0.8350***	***0.0909***	***0.8307***	***0.0691***
		M2	0.8195	0.0957	0.8146	0.0725
		M3	0.8232	0.0929	0.8056	0.0740

Data have been normalized, so RMSE has no unit.

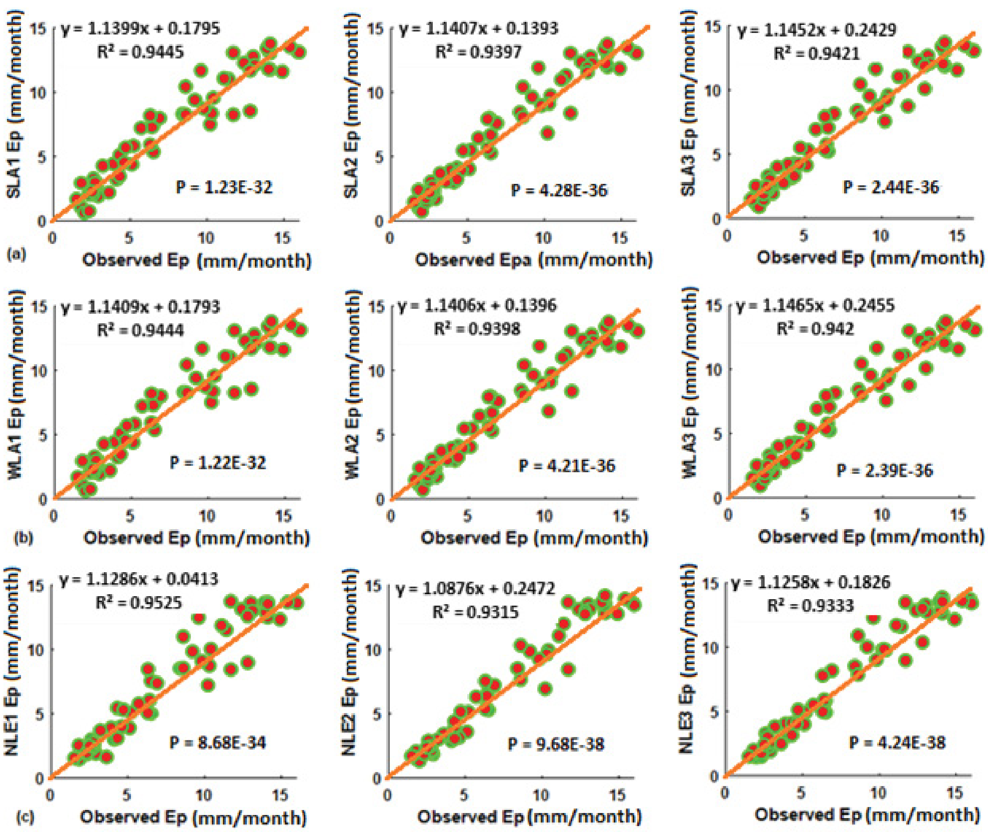

Fig. 9. Observed vs. predicted plots for Erbil station for (a) M1, (b) M2 and (c) M3.

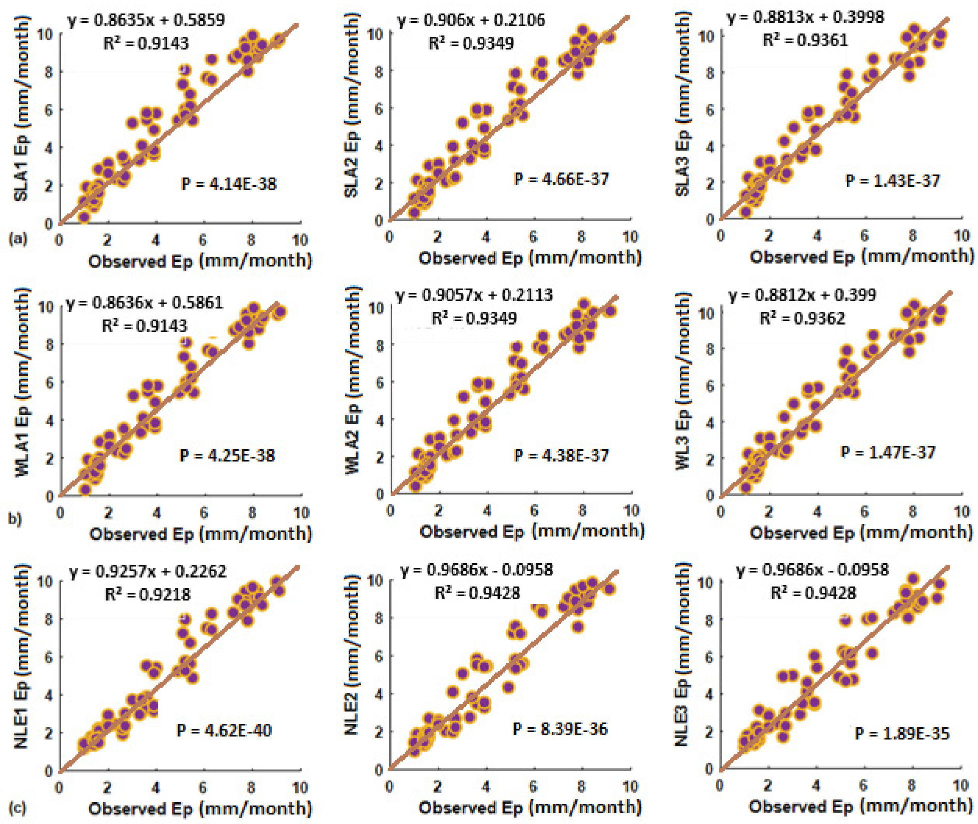

Fig. 10. Observed vs. predicted plots for Salahaddin station for (a) M1, (b) M2 and (c) M3.

best performance. This might be because; (i) non-linear kernel is used for the development of NLE model, which has the capability of dealing with non-linear aspect of the evaporation process as such, produced highest performance. (ii) Both SLA and WLA simulate the behavior of the system linearly, hence errors developed from the single models could be generated by the linear models which could reduce their performances. Fig. 10 shows scatter plots for Salahaddin station in the verification phase of Salahaddin station.

By careful observation of the obtained results in Table 4, it can be realized that similar to single models, the applied ensemble models has better results in Erbil station than Salahaddin station. However, higher percentage of ensemble performances are achieved in Salahaddin than Erbil station. These distinct characteristics implied that, ensemble models emulate the performance of single models, meaning more efficient single models will lead to more accurate ensemble models and vice versa. On the other hand, less performance single models have more space for accuracy improvement, hence, higher increment in percentage is achieved by less efficient single models.

4. Conclusion

In this study, novel artificial intelligence (AI) based ensemble techniques including simple linear average (SLA), weighted linear average (WLA) and non-linear ensemble (NLE) were applied for monthly pan evaporation (Ep) modeling across Erbil and Salahaddin stations in Iraq. The advantage of this proposed methods over others is that both linear and non-linear aspects of Ep are taken into cognizance, thereby resulting in more robust, improved and accurate predictions. For this purpose, two AI based models including feed-forward back propagation (FFBP) neural network and LS-SVM were employed initially as single models after sensitivity analysis was performed that determined the best input combinations. Additionally, a conventional multiple linear regression model was also applied for comparison. Thereafter, the ensemble techniques were applied to improve the performance of the single models.

The simulation results and the comparative analysis performed indicated that the proposed ensemble techniques (SLA, WLA and NLE) can be useful tools for performance improvement of Ep time series prediction and they have outperformed all the single models tested using the same datasets. The results also showed that ensemble models could improve predictions of single models up to 5% for Erbil station and 16% for Salahaddin station. The overall results showed that, ensemble modeling could be applied for both least performance and high-performance single models, but for better accuracy, high performance single models should be used for ensemble modeling of Ep in Erbil and Salahaddin stations.

This study has two main contributions: (i) The proposed ensemble techniques have improved the predicting performance of AI based single models. Despite the uncertainty and difficulty surrounding the Ep prediction, they have produced a promising improvement over single models. These can serve as an alternative methods for other time series and hydro-climatological studies including evapotranspiration, precipitation to mention a few. (ii) The applied ensemble methods also implied that their successful application is possible in all climate regions. Being semiarid climate stations (Erbil and Salahaddin) that characterized by scarce water resources, Ep prediction in those regions is difficult and challenging task. As such, the required performance could be achieved if the methods are applied in other water scarce regions such as arid and hyper arid climate stations. Further studies should include the application of other heuristic computing approaches and incorporation of more stations from distinct climates to investigate the behavior of ensemble models with respect to climate and different stations.

References

[1] H. Abghari, H. Ahmadi, S. Besharat, V. Rezaverdinejad, Prediction of daily pan evaporation using wavelet neural networks, Water Resour. Manage., 26 (2012) 3639–3652.

[2] O. Kisi, Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree, J. Hydrol., 528 (2015) 312–320.

[3] H.A.K.A. Khayyat, A.J.M. Sharif, M. Crespi, Assessing the Impacts of Climate Change on Natural Resources in Erbil Area, the Iraqi Kurdistan Using Geo-Information and Landsat Data, A. Al-Quraishi, A. Negm, Eds., Environmental Remote Sensing and GIS in Iraq, Springer Water, Springer, Cham. Available at: https://doi.org/10.1007/978-3-030-21344-2_19

[4] J. Piri, S. Amin, A. Moghaddamnia, A. Keshavarz, D. Han, R. Remesan, Daily pan evaporation modeling in a hot and dry climate, J. Hydrol. Eng., 14 (2009) 803–811.

[5] L. Wang, Z. Niu, O. Kisi, C.A. Li, D. Yu, Pan evaporation modeling using four different heuristic approaches, Comput. Electron. Agric., 140 (2017a) 203–213.

[6] M. Abed, M.A. Imteaz, A.N. Ahmed, Y.F. Huang, Application of long short-term memory neural network technique for predicting monthly pan evaporation, Sci. Rep., 11 (2021) 20742, doi: 10.1038/s41598-021-99999-y.

[7] M. Abed, M.A. Imteaz, A.N. Ahmed, Y.F. Huang, A novel application of transformer neural network (TNN) for estimating pan evaporation rate, Appl. Water Sci., 13 (2023) 31, doi: 10.1007/s13201-022-01834-w.

[8] S. Kim, J. Shiri, V.P. Singh, O. Kisi, G. Landeras, Predicting daily pan evaporation by soft computing models with limited climatic data, Hydrol. Sci. J., 60 (2015) 1120–1136.

[9] L. Wang, O. Kisi, M. Zounemat-Kermani, H. Li, Pan evaporation modeling using six different heuristic computing methods in different climates of China, J. Hydrol., 544 (2017b) 407–427.

[10] V. Nourani, M.S. Fard, Sensitivity analysis of the artificial neural network outputs in simulation of the evaporation process at different climatologic regimes, Adv. Eng. Software, 47 (2012) 127–146.

[11] A. Rahimikhoob, Estimating daily pan evaporation using artificial neural network in a semi-arid environment, Theor. Appl. Climatol., 98 (2009) 101–105.

[12] P.B. Shirsath, A.K. Singh, A comparative study of daily pan evaporation estimation using ANN, regression and climate based models, Water Resour. Manage., 24 (2010) 1571–1581.

[13] S.N. Qasem, S. Samadianfard, S. Kheshtgar, S. Jarhan, O. Kisi, S. Shamshirband, K.W. Chau, Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates, Eng. Appl. Comput. Fluid Mech., 13 (2019) 177–187.

[14] J.-L. Chen, H. Yang, M.-Q. Lv, Z.-L. Xiao, S.J. Wu, Estimation of monthly pan evaporation using support vector machine in Three Gorges Reservoir Area, China, Theor. Appl. Climatol., 138 (2019) 1095–1107.

[15] W.G. Zhang, H.R. Li, C.Z. Wu, Y.Q. Li, Z.Q. Liu, H.L. Liu, Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling, Underground Space, 4 (2021) 353–363.

[16] J. Estévez, P. Gavilán, J.V. Giráldez, Guidelines on validation procedures for meteorological data from automatic weather stations, J. Hydrol., 402 (2011) 144–154.

[17] J. Bates, C.W.J. Granger, The combination of forecasts, J. Oper. Res. Soc., 20 (1969) 451–468.

[18] S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, R. Winkler, The accuracy of extrapolation (time series) methods: results of a forecasting competition, Int. J. Forecast., 1 (1982) 111–153.

[19] N.R. Kiran, V. Ravi, Software reliability prediction by soft computing techniques, J. Syst. Software, 81 (2008) 576–583.

[20] E. Sharghi, V. Nourani, N. Behfar, Earthfill dam seepage analysis using ensemble artificial intelligence-based modeling, J. Hydroinf., 20 (2018) 1071–1084.

[21] V. Nourani, G. Elkiran, J. Abdullahi, Multi-station artificial intelligence-based ensemble modeling of reference evapotranspiration using pan evaporation measurements, J. Hydrol., 577 (2019a) 123958, doi: 10.1016/j.jhydrol.2019.123958.

[22] V. Nourani, G. Elkiran, J. Abdullahi, A. Tahsin, Multi-region modeling of daily global solar radiation with artificial intelligence ensemble, Nat. Resour. Res., 28 (2019b) 1217–1238.

[23] V. Nourani, G. Elkiran, J. Abdullahi, Multi-step ahead modeling of reference evapotranspiration using a multi-model approach. J. Hydrol., 581 (2020) 124434, doi: 10.1016/j.jhydrol.2019.124434.

[24] A. Rasul, H. Balzter, C. Smith, Spatial variation of the daytime Surface Urban Cool Island during the dry season in Erbil, Iraqi Kurdistan, from Landsat 8, Urban Clim., 14 (2015) 176–186.

[25] N. Şarlak, O.M.M. Agha, Spatial and temporal variations of aridity indices in Iraq, Theor. Appl. Climatol., 133 (2018) 89–99.

[26] G. Elkiran, V. Nourani, S.I. Abba, J. Abdullahi, Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river, Global J. Environ. Sci. Manage., 4 (2018) 439–450.

[27] D.R. Legates, G.J. McCabe Jr., Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35 (1999) 233–241.

[28] J. Estévez, A.P. García-Marín, J.A. Morábito, M. Cavagnaro, Quality assurance procedures for validating meteorological input variables of reference evapotranspiration in mendoza province (Argentina), Agric. Water Manage., 172 (2016) 96–109.

[29] M.A. Shafer, C.A. Fiebrich, D.S. Arndt, S.E. Fredrickson, T.W. Hughes, Quality assurance procedures in the Oklahoma Mesonetwork, J. Atmos. Oceanic Technol., 17 (2000) 474–494.

[30] S. Feng, Q. Hu, W. Qian, Quality control of daily meteorological data in China, 1951–2000: a new dataset, Int. J. Climatol.: J.R. Meteorolog. Soc., 24 (2004) 853–870.

[31] V. Nourani, M.T. Alami, M.H. Aminfar, A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation, Eng. Appl. Artif. Intell., 22 (2009) 466–472.

[32] J. Abdullahi, G. Elkiran, Prediction of the future impact of climate change on reference evapotranspiration in Cyprus using artificial neural network, Procedia Comput. Sci., 120 (2017) 276–283.

[33] J. Abdullahi, G. Elkiran, V. Nourani, Application of Artificial Neural Network to Predict Reference Evaporation in Famagusta, North Cyprus, 11th International Scientific Conference on Production Engineering Development and Modernization of Production, Bihac, Bosnia, 2017, pp. 549–554.

[34] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks, 2 (1989) 359–366.

[35] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural Process. Lett., 9 (1999) 293–300.

[36] S.G. Cao, Y.B. Liu, Y.P. Wang, A forecasting and forewarning model for methane hazard in working face of coal mine based on LS-SVM, J. China Univ. Min. Technol., 18 (2008) 172–176.
[37] R. Fletcher, Practical Methods of Optimization, John Wiley and Sons, New York, 1987.
[38] M.K. Goyal, B. Bharti, J. Quilty, J. Adamowski, A. Pandey, Modeling of daily pan evaporation in sub-tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS, Expert Syst. Appl., 41 (2014) 5267–5276.
[39] J.M. Bruton, R.W. McClendon, G. Hoogenboom, Estimating daily pan evaporation with artificial neural networks, Trans. ASAE, 43 (2000) 491–496.
[40] B. Simon-Gáspár, G. Soós, A. Anda, Pan evaporation is increased by submerged macrophytes, Hydrol. Earth Syst. Sci., 26 (2022) 4741–4756.
[41] L. Wang, O. Kisi, B. Hu, M. Bilal, M. Zounemat-Kermani, H. Li, Evaporation modelling using different machine learning techniques, Int. J. Climatol., 37 (2017c) 1076–1092.
[42] N.M. UNEP, D. Thomas, World Atlas of Desertification, Edward Arnold, London, 1992.