

Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates

Sultan Noman Qasem, Saeed Samadianfard, Salar Kheshtgar, Salar Jarhan, Ozgur Kisi, Shahaboddin Shamshirband & Kwok-Wing Chau

To cite this article: Sultan Noman Qasem, Saeed Samadianfard, Salar Kheshtgar, Salar Jarhan, Ozgur Kisi, Shahaboddin Shamshirband & Kwok-Wing Chau (2019) Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates, Engineering Applications of Computational Fluid Mechanics, 13:1, 177-187, DOI: [10.1080/19942060.2018.1564702](https://doi.org/10.1080/19942060.2018.1564702)

To link to this article: <https://doi.org/10.1080/19942060.2018.1564702>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 11 Jan 2019.



Submit your article to this journal [↗](#)



Article views: 140



View Crossmark data [↗](#)

Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates

Sultan Noman Qasem^{a,b}, Saeed Samadianfard^c, Salar Kheshtgar^d, Salar Jarhan^e, Ozgur Kisi^f, Shahaboddin Shamshirband^{g,h} and Kwok-Wing Chauⁱ

^aComputer Science Department, College of Computer and information Sciences, Al Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia; ^bComputer Science Department, Faculty of Applied Sciences, Taiz University, Taiz, Yemen; ^cDepartment of Water Engineering, University of Tabriz, Tabriz, Iran; ^dDepartment of Building, Civil and Environmental Engineering, Faculty of Engineering and Computer Science, Concordia University, Montreal, Canada; ^eDepartment of Hydrosocieties, Technische Universität Dresden, Dresden, Germany; ^fFaculty of Natural Sciences and Engineering, Ilia State University, Tbilisi, Georgia; ^gDepartment for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam; ^hFaculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam; ⁱDepartment of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

ABSTRACT

Evaporation rate is one of the key parameters in determining the ecological conditions and it has an irrefutable role in the proper management of water resources. In this paper, the efficiency of some data-driven techniques including support vector regression (SVR) and artificial neural networks (ANN) and combination of them with wavelet transforms (WSVR and WANN) were investigated for predicting evaporation rates at Tabriz (Iran) and Antalya (Turkey) stations. For evaluating the performances of studied techniques, four different statistical indicators were utilized namely the root mean square error (RMSE), the mean absolute error (MAE), the correlation coefficient (R), and Nash–Sutcliffe efficiency (NSE). Additionally, Taylor diagrams were implemented to test the similarity among the observed and predicted data. Outcomes showed that at Tabriz station, the ANN3 (third input combination that are air temperatures and solar radiation used by ANN) with RMSE of 0.701, MAE of 0.525, R of 0.990 and NSE of 0.977 had better performances in comparison with WANN, SVR and WSVR. So, the wavelet transforms did not have positive effects in increasing the precision of ANN and SVR predictions at Tabriz station. Also, approximately the same trend was seen at Antalya station. In other words, ANN5 (fifth input combination that are air temperatures, relative humidity and solar radiation used by ANN) with RMSE of 0.923, MAE of 0.697, R of 0.962 and NSE of 0.898 had a more accurate predictions among others. Conversely, wavelet transform reduced the prediction errors of SVR at Antalya station. So, the WSVR5 with RMSE of 1.027, MAE of 0.728, R of 0.950 and NSE of 0.870 predicted evaporation rates of Antalya station more precisely than other SVR models. As a conclusion, results from the current study proved that ANN provided reasonable trends for evaporation modeling at both Tabriz and Antalya stations.

ARTICLE HISTORY

Received 24 October 2018
Accepted 28 December 2018

KEYWORDS

Data-driven technique; estimation; meteorological parameters; pan evaporation

1. Introduction

Evaporation is playing a prominent role in the processes of hydrologic cycle. Therefore, estimation of this phenomenon is necessary as it affects water resources, especially in arid and semi-arid areas like Iran. A considerable amount of precipitation is allocated to evaporation and the sum of evaporation from the land surface is approximately 61% of the entire global precipitation (Chow, Maidment, & Mays, 1988). Every year, several millions of cubic meters of fresh water that has collected with high costs, efforts and difficulties evaporated from tank of dams and reservoirs. Therefore, it is of great significance to estimate the evaporation amount for water

resource monitoring. In regions with hot weather and low rainfall, evaporation can lead to losing the noticeable amount of water, hence it causes decreasing in the water surface elevation. Typically, there are two different (direct and indirect) approaches for estimation of evaporation from various variables. Measuring pan is the most prevalent direct method for specifying the value of evaporation (Stanhill, 2002). Class A pan, as the most frequently used pan, has 4 ft (122 cm) in diameter and 10 inch (25 cm) deep and is located about 6 inch (15 cm) above the soil surface. Although calculating evaporation by using pans appears to be having precise answers but using indirect approaches for approximating pan

evaporation (PE) seems to be easier and cost-effective due to the fact that installing pans and meteorological stations are the expensive process. Indirect methods for estimation of PE are based on using meteorological variables. Empirical and semi-empirical models have been completely used for calculating PE from meteorological variables such as wind speed, sunshine hours, relative humidity, radiation, atmosphere pressure and the main factor, temperature. Because of these several variables, estimating evaporation rate becomes more complex as it is totally non-linear. A number of scholars have tried to predict evaporation from different meteorological variables (Chang, Sun, & Chung, 2013; Kisi, 2006; Lin, Lin, & Wu, 2013). In recent years, machine learning approaches including Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Fuzzy Genetic (FG), Support Vector Regression (SVR) and the integrated models of these methods with wavelet or other data preprocessing approaches have been effectively applied in water related fields such as water resource engineering, prediction of suspended sediment load in rivers, forecasting monthly stream flow, estimating friction factor in irrigation pipes, PE modeling, etc. (Chau, 2017; Choubin et al., 2018; Choubin, Malekian, Samadi, Khalighi Sigaroodi, & Sajedi Hosseini, 2017; Fotovatikhah et al., 2018; Jain & Srinivasulu, 2004; Keshtegar, Piri, & Kisi, 2016; Kisi, 2005; Kisi, Genc, Dinc, & Zounemat-Kermani, 2016; Nourani, Alami, & Daneshvar Vousoughi, 2015; Rajaei, Mirbagheri, Zounemat-Kermani, & Nourani, 2009; Samadianfard et al., 2018; Samadianfard, Delirhasannia, Kisi, & Agirre-Basurko, 2013; Samadianfard, Nazemi, & Sadraddini, 2014; Samadianfard, Sattari, Kisi, & Kazemi, 2014; Taormina, Chau, & Sivakumar, 2015; Wang, Xu, Chau, & Chen, 2013). Goyal, Bharti, Quilty, Adamowski, and Pandey (2014) applied ANN, ANFIS, least squares-SVR (LS-SVR) and fuzzy logic to increase the precision of PE estimation using various meteorological variables, like daily rainfall, sunshine hours, minimum and maximum air temperatures and humidity. The outcomes exhibited that the LS-SVR and fuzzy logic can be utilized effectively in predicting PE values. Guven and Kisi (2013) compared the precision of five different models in modeling PE named: ANFIS, ANN, FG, Stephens–Stewart (SS) and Linear Genetic Programming (LGP) methods. The results revealed that LGP model has a better accuracy in comparison to other models and can be utilized successfully to estimate PE. Kisi and Tombul (2013) applied the FG method in estimation of PE in two different stations inside Turkey and compared the accuracy of the method with the ANFIS, ANN and SS approaches. Finally, it was concluded that the FG method has the closest estimation in comparison to other models. Tabari, Marofi, and

Sabziparvar (2009) examined the precision of ANN and Multivariate Non-Linear Regression (MNLRL) methods for PE estimation in a semi-arid area. The consequences indicated that the ANN approach provided the finest estimates of PE. ANN and genetic algorithms were utilized by Kim and Kim (2008) for PE and evapotranspiration modeling. The outcomes indicated the specific capabilities of both studied models for the precise approximation of PE and evapotranspiration values. Recently, Shafaei and Kisi (2015) applied ANFIS, SVR and autoregressive moving average (ARMA) models, separately, then combined the models with wavelet, known as WANFIS, WSVR, and WARMA models. The efficiency of the aforementioned two groups (single form of models and combined forms) is compared with each other. Obtained results discovered that the combined models give better accuracy in anticipating lake levels in the study zone in comparison to single models.

In previous researches, it has been tried to improve the prediction accuracy of PE estimates. So, it can be comprehended that the current gap is developing and implementing new methods for decreasing the prediction errors. So, in the current research, the potential of two machine learning approaches, namely ANN and SVR, and the integrated wavelet of these models, Wavelet-ANN (WANN) and Wavelet-SVR (WSVR), used to forecast evaporation rates in Tabriz and Antalya stations and the efficiencies of them were compared with each other. The goals of the study were (i) evaluating the performance of aforementioned models in the estimation of PE, and (ii) investigating the role of climate in predicting PE by considering humid and arid climates. The rest of the paper is structured as follow. Section 2 provides the characteristics of study area, description of implemented methods and evaluation parameters. Additionally, section 3 discussed the obtained results and finally, the conclusion is presented in section 4.

2. Materials and methods

2.1. Study area

The monthly data of two synoptic stations in two different countries, Tabriz Station (latitude 38°05'N, longitude 46°17'E) in Iran and Antalya Station (latitude 36°42'N, longitude 30°44'E) in Turkey that are in approximately same latitudes, were used in the current study (Figure 1). Tabriz, which is situated at northwestern of Iran, is a mountainous basin with a semi-arid climate and cold winters. Besides, Antalya, which is placed in the southern region of Turkey, has Mediterranean climate that characterized by temperate wet winters and hot arid summer. The elevations from sea level are 1350 and 64 m



Figure 1. The location of Tabriz and Antalya stations (URL1).

Table 1. The monthly statistical parameters of each meteorological station.

Statistic	T ($^{\circ}\text{C}$)		RH (%)		W (m/s)		R_s (MJ/m^2 day)		PE (mm)	
	Tabriz	Antalya	Tabriz	Antalya	Tabriz	Antalya	Tabriz	Antalya	Tabriz	Antalya
Minimum	−6.474	7.800	25.137	45.500	1.117	1.800	6.243	5.588	0.000	1.600
Maximum	28.769	32.250	82.460	67.500	5.617	4.900	30.008	30.245	15.448	12.400
Mean	13.092	19.874	51.058	55.800	3.385	2.756	18.011	17.608	5.425	5.568
Standard deviation	9.848	7.327	14.406	4.223	0.934	0.517	7.220	6.523	4.956	2.672
Variation coefficient	0.751	0.368	0.282	0.076	0.275	0.187	0.400	0.370	0.912	0.479
Skewness	−0.064	−0.001	0.181	0.123	0.161	0.817	0.022	−0.132	0.299	0.455
Kurtosis	−1.322	−1.408	−1.193	−0.219	−0.553	1.206	−1.403	−1.262	−1.358	−0.887
Correlation with PE	0.963	0.912	−0.892	−0.342	0.730	−0.269	0.927	0.843	1.000	1.000

for the Tabriz and Antalya stations, respectively. In summary, as it can be proved from Table 1, Antalya has a hot and wet climate. In contrary, Tabriz has cold and dry weather. Meteorological variables that were used for this study are air temperatures (T), solar radiation (R_s), wind speed (W), relative humidity (RH) and PE with the time period of 1992–2016 for Tabriz station and 1986–2006 for Antalya station. Table 1 represents the monthly statistical parameters of the applied variables for both stations. It can be seen in Table 1 that the mean values of PE at Tabriz and Antalya stations are 5.425 and 5.568 mm, respectively. Furthermore, it is obvious that the minimum PE in Tabriz station is zero because in winters when the temperature reaches below zero, water freezes in the pan and the volume of evaporation becomes negligible. Most of the variables indicate normal distributions because they have low skewness values, except wind speed in Antalya station which has the maximum skewness and shows a skewed distribution. This might be due to the location of this station in a coastal area (Kisi, 2009). Correlations are generally higher for the Tabriz station and temperature and PE show a higher variation in Tabriz when compared

to Antalya. Table 1 also shows the correlations between PE and other parameters. In both stations, T and R_s have first and second high correlations with PE, respectively. Moreover, the highest inverse correlations are between RH and PE in both stations. The observed PE data for both Tabriz and Antalya stations are shown in Figure 2.

2.2. Artificial neural networks

In preceding decades, the ANN approach was considered by many researchers as a dependable tool of computations. From the initial neural model which has been introduced by McCulloch and Pitts (1943), researchers developed several different models. Additionally, most activities have been concentrated on back-propagation and its extensions (Salas, Markus, & Tokar, 2000). This algorithm (Rumelhart & McClelland, 1986) is utilized in feed-forward ANNs, meaning that the neurons are ordered in several layers, and deliver their signals forward, then the errors are propagated backward. The network obtains inputs by input layer neurons, and the network output is obtained from the output layer

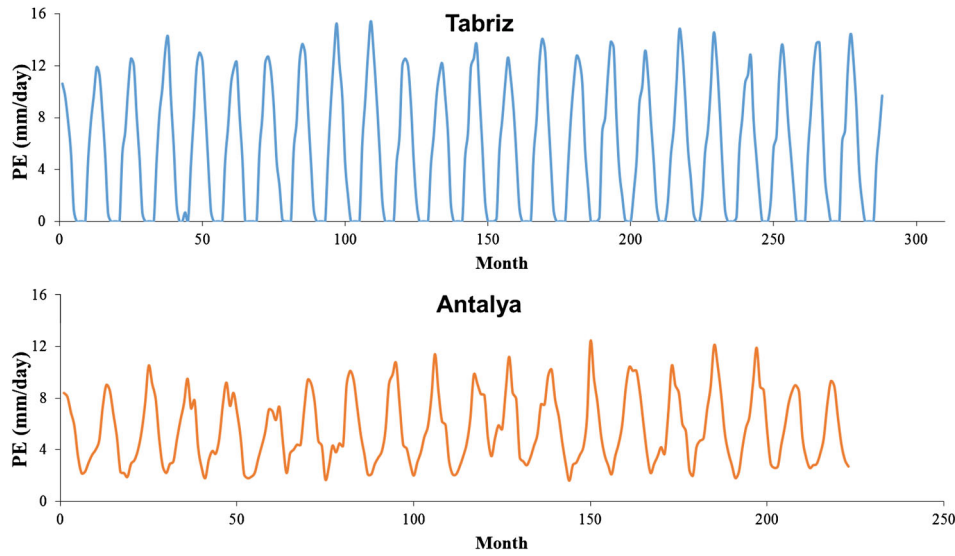


Figure 2. The observed PE for Tabriz and Antalya stations.

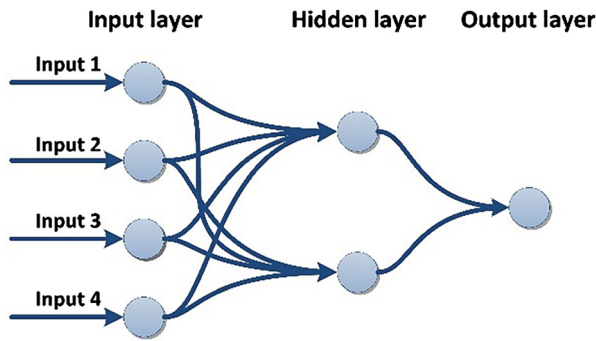


Figure 3. The typical configuration of the ANN model.

neurons. The number of hidden layers is typically determined by trial and error. ANNs are extremely talented of modeling and simulating linear and non-linear systems. In the current study, input layers were determined using a trial-error process and the number of neurons in each layer were selected to be among 1 and 10. Figure 3 shows a typical configuration of the ANN model.

2.3. Support vector regression

The support vector machine (SVM) is an extensively used and popular estimator which its fundamental idea was introduced by Vapnik (1995, 1998). Then, a regression technique, namely SVR, was developed based on SVM models in order to solve complex problems, effectively. This model was constructed on the structural risk minimization. For this purpose, ε – the insensitive loss function is identified which means that the model allows tolerating errors up to ε in the training data sets. Therefore,

the SVR seeks for a linear function formulated as follows:

$$P(x) = F^T x + L, \quad (1)$$

where F and L denote the coefficients of the weight vector of the linear expression. This linear regression problem can be explained as the following optimization problem:

$$\begin{aligned} &\text{Min } \frac{1}{2} \|F\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \text{ Subject to} \\ &\times \begin{cases} F^T x + L - y_i \leq \varepsilon + \xi_i^* \\ y_i - F^T x - L \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, N \end{cases} \end{aligned} \quad (2)$$

The constant $C > 0$ is a positive trade-off factor for the grade of the experimental error. C and ε are predefined parameters. Moreover, ξ_i and ξ_i^* that is called as slack variables. For solving this problem, the SVR uses kernel trick (Smola & Scholkopf, 2004). In this study, four different kernel functions, including Pearson VII function-based, radial basis function (RBF), polynomial and normalized polynomial are used. For building optimum SVR model, the parameters of SVR such as the kernel function, C and ε must be selected cautiously. Figure 4 indicates the schematic configuration of the SVR model.

2.4. Wavelet transform

Developed by Grossman and Morlet (1984), the wavelet transform (WT) has been broadly applied in different fields of science and engineering. It is mathematical transformations which are used for eliciting further information from the data which is not currently obvious in its crude form. This transformation can determine

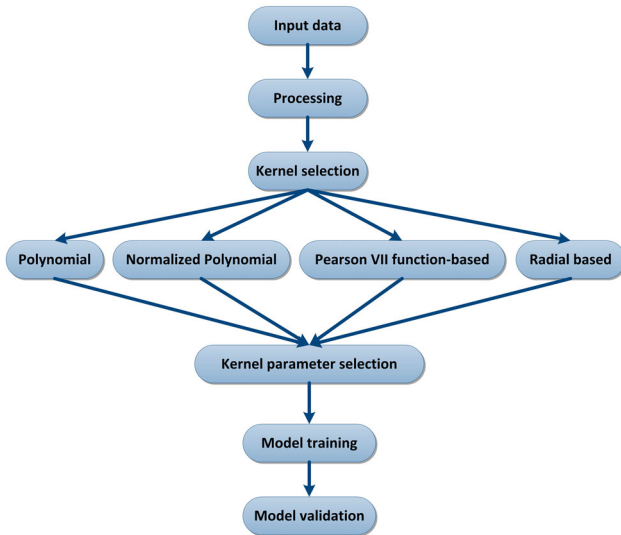


Figure 4. Schematic configuration of SVR model.

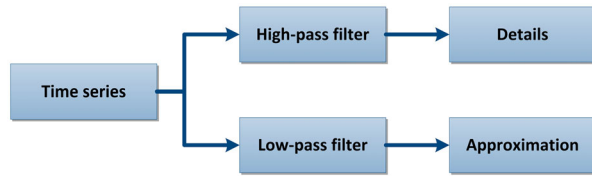


Figure 5. Schematic structure of decomposition by DWT.

trends in the data such as discontinuities, breakdown points and minimum and maximum. Also, the ability of WT to be applied to any types of time series has made it a beneficial tool. As a consequence, this tool seems to be an effective substitute for the Fourier transform in saving local, non-periodic and multi-scaled phenomena (Goyal et al., 2014). The WT generates some sub-series from the main time series data in order to create wavelet data-driven models. The WT appears in different forms such as continuous (CWT) and discrete (DWT). The calculation of the CWT approach requires large data and high calculation time. Conversely, the DWT needs less calculation time and is simpler to apply than the CWT. Two sets of filters, including low and high passes, are employed by DWT to decompose the main time series (Figure 5) (Nourani, Alami, & Aminfar, 2009). Choosing wavelet types and level of decomposition are the key points of WT. These factors should determine carefully. So, trial and error procedure is frequently utilized for finding appropriate mother wavelet type. Different wave families were examined in the current research and Daubechies with relatively better performance was selected for PE estimation. Furthermore, an empirical survey by Nourani, Hosseini Baghanam, Adamowski, and Kisi (2014) recommended that the optimum enactment of WT occurs when the decomposition level is

selected equal to the integer part of $[\log(N)]$. So, this parameter was selected equal to two for PE estimation.

2.5. Wavelet artificial neural networks and wavelet support vector regression

In this investigation, wavelet analysis was combined with ANN and SVR for monthly pan evaporation prediction as presented in Figure 5. Above mentioned wave types were evaluated and finally the measured monthly time series, H were decomposed into multi-frequency time series including details ($H_{D1}; H_{D2}; \dots; H_{Dn}$) and approximation (H_a) by optimum DWT. All of the decomposed components were used as inputs for both ANN and SVR separately. The output is known as prediction shown in Figure 6.

2.6. Evaluation parameters

For evaluating the precision of the studied models, four different evaluation functions are used including root mean squared error (RMSE), mean absolute error (MAE), correlation coefficient (R) and Nash–Sutcliffe efficiency (NSE). The mentioned functions are defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (p_i - o_i)^2}{n}}, \quad (3)$$

where p_i is the i^{th} predicted value by models, o_i is the i^{th} observed value, and n is an entire number of sample data. The MAE, R and NSE functions can be stated as:

$$\text{MAE} = \frac{\sum_{i=1}^n |p_i - o_i|}{n}, \quad (4)$$

$$R = \frac{\sum_i (o_i - \bar{o})(p_i - \bar{p})}{\sqrt{\sum_{i=1}^n (o_i - \bar{o})^2 \sum_{i=1}^n (p_i - \bar{p})^2}}, \quad (5)$$

$$\text{NSE} = 1 - \frac{\sum_i (o_i - p_i)^2}{\sum_i (o_i - \bar{o})^2}, \quad (6)$$

where \bar{o} and \bar{p} are an average of observed and predicted data, respectively.

Moreover to the statistical parameters defined in Equations (3–6), Taylor diagram (Taylor, 2001) was used to validate the accuracy of considered models. Basically, the mentioned diagram is a detailed illustration of the observed and predicted data (e.g. IPCC, 2007). Taylor presented a solo illustration for presenting different evaluation parameters, instantaneously. Remarkably, these are capable for highlighting the correctness of predictive models using several points on a polar plot. In this diagram, correlation coefficient, standard deviation and

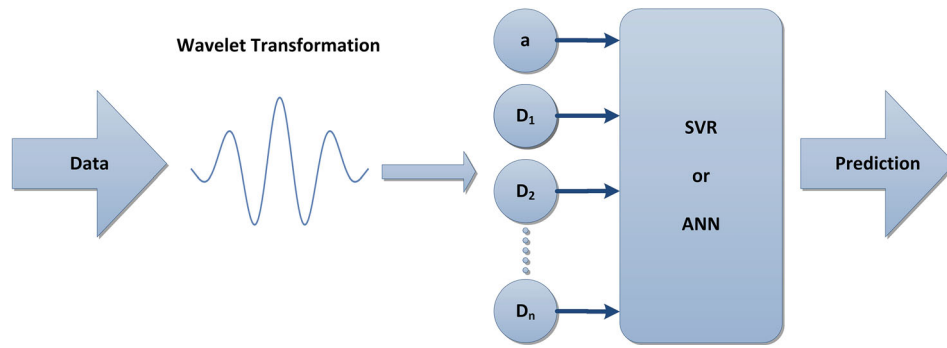


Figure 6. Schematic representation of WANN or WSVR models.

RMSE values among predicted and observed values can be presented for better recognition of variations (Gleckler, Taylor, & Doutriaux, 2008; Taylor, 2001). Additionally, the mentioned diagram can be plotted by a developed code in Wolfram Mathematica (version 11.0.1.0).

3. Results and discussion

The performance of considered methods, namely ANN, WANN, SVR and WSVR, to forecast PE at two different stations, Tabriz (northwest of Iran) and Antalya (southwest of Turkey), was compared. In this research, all models were partitioned, where 70% of the data was used to train and the 30% was used to test the mentioned models. For Tabriz station, monthly data from 1992 to 2008 were considered as the training data set, while the measured data from 2009 to 2016 served as the testing data set. Furthermore, for Antalya station, training and testing data set periods were from 1986 to 1999 and 2000 to 2006, respectively. Based on meteorological variables, seven different input combinations were considered in the current study. Because T has the highest correlation with PE in Tabriz and Antalya stations, it has been used in all input combinations to increase prediction accuracy. As given in Table 2, the evaluated input combinations are: (1) T and RH, (2) T and W , (3) T and R_s , (4) T , RH and W , (5) T , RH and R_s , (6) T , W and R_s , (7) T , RH, W and R_s . The results of statistical parameters including RMSE, MAE and R values for prediction of PE at Tabriz

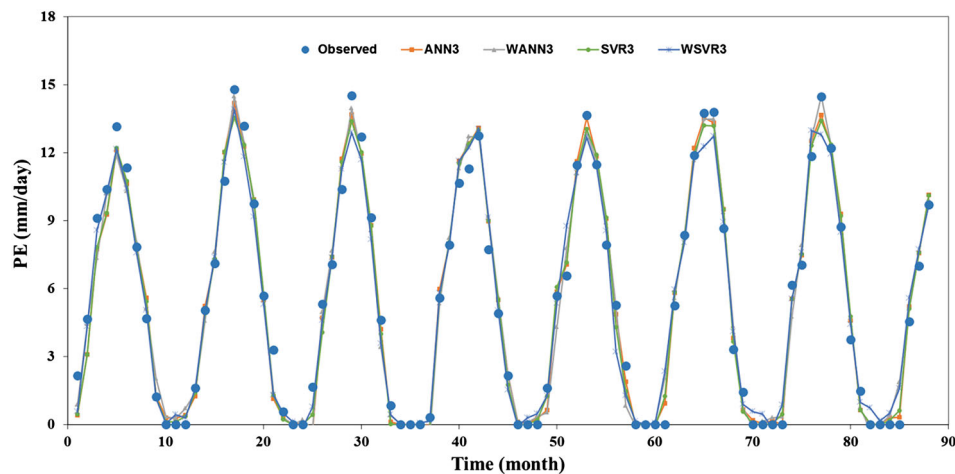
and Antalya stations for mentioned methods are illustrated in Table 3. For Tabriz station, ANN3 with RMSE of 0.701, MAE of 0.525, R value of 0.990 and NSE of 0.977 had the best performance in comparison with other ANN scenarios and also with other methods such as WANN, SVR and WSVR. Moreover, SVR3 ranked the second best for predicting the PE at Tabriz station with RMSE of 0.712, MAE of 0.544, R value of 0.990 and NSE of 0.977. Also, ANN5 was selected as third best at Tabriz station with RMSE of 0.736, MAE of 0.546, R value of 0.989 and NSE of 0.976. Furthermore, it is clear from Table 3 that WT did not have a substantial positive effect in reducing the prediction error at Tabriz station. In other words, the precision of ordinary ANN and SVR methods were higher than correspondent ones with WT. Thus, it can be concluded that the meteorological parameters of temperature and radiation had the most effects in the exact prediction of PE at Tabriz station. Also, relative humidity had a positive effect in scenario V for reducing the prediction errors. Somehow different trend was seen for Antalya station. At this station, ANN5 with input parameters of temperature, radiation and relative humidity, with RMSE of 0.923, MAE of 0.697, R value of 0.962 and NSE of 0.895 had more accurate prediction in comparison with other methods and scenarios. Then, WANN5 was selected as the second best with the same input parameters, with RMSE of 0.927, MAE of 0.682, R value of 0.968 and NSE of 0.894. In addition, ANN6 was selected as the third best with input parameters of temperature, radiation and wind speed. So, it is clear that, similar to the Tabriz station, temperature and radiation had the most positive effects for presenting an accurate prediction of pan evaporation. Also, relative humidity increased the prediction precision in Antalya station. It is maybe due to the fact that Antalya is next to the seaside, therefore it could be concluded that relative humidity is an important factor in increasing the accuracy of estimated pan evaporation, adjacent to the seaside. So, it can be comprehended that the ANN has the undeniable capability in reducing the

Table 2. Input parameters for implementation of studied models.

Model	Input parameters	Output parameter
1	T , RH	PE
2	T , W	PE
3	T , R_s	PE
4	T , RH, W	PE
5	T , RH, R_s	PE
6	T , W , R_s	PE
7	T , RH, W , R_s	PE

Table 3. Statistical meters for various ANN, WANN, SVR and WSVR models in test stage.

Model	Tabriz				Antalya			
	MAE (mm)	RMSE (mm)	<i>R</i>	NSE	MAE (mm)	RMSE (mm)	<i>R</i>	NSE
ANN1	0.719	0.928	0.983	0.966	0.858	1.144	0.923	0.839
ANN2	0.650	0.826	0.987	0.971	0.873	1.100	0.944	0.851
ANN3	0.525	0.701	0.990	0.977	0.860	1.145	0.962	0.838
ANN4	0.737	0.954	0.983	0.965	0.917	1.192	0.937	0.825
ANN5	0.546	0.736	0.989	0.976	0.697	0.923	0.962	0.895
ANN6	0.584	0.737	0.989	0.975	0.883	1.102	0.965	0.850
ANN7	0.641	0.816	0.987	0.972	0.776	1.000	0.967	0.877
WANN1	0.660	0.877	0.984	0.968	0.817	1.104	0.930	0.850
WANN2	0.739	0.913	0.983	0.967	0.891	1.112	0.939	0.847
WANN3	0.594	0.787	0.987	0.974	0.836	1.103	0.957	0.850
WANN4	0.715	0.921	0.984	0.968	0.909	1.181	0.937	0.828
WANN5	0.636	0.818	0.986	0.972	0.682	0.927	0.968	0.894
WANN6	0.636	0.786	0.987	0.973	0.921	1.113	0.967	0.847
WANN7	0.630	0.815	0.986	0.971	0.878	1.092	0.968	0.853
SVR1	0.715	0.944	0.982	0.963	0.956	1.189	0.915	0.826
SVR2	0.626	0.852	0.986	0.971	0.902	1.145	0.939	0.838
SVR3	0.544	0.712	0.990	0.977	0.900	1.221	0.948	0.816
SVR4	0.691	0.931	0.984	0.966	0.894	1.137	0.939	0.841
SVR5	0.593	0.757	0.988	0.975	0.765	1.070	0.951	0.859
SVR6	0.584	0.779	0.988	0.973	1.015	1.251	0.961	0.807
SVR7	0.594	0.786	0.988	0.972	0.903	1.150	0.963	0.837
WSVR1	0.771	0.999	0.979	0.956	0.902	1.205	0.915	0.821
WSVR2	0.839	1.049	0.979	0.955	0.895	1.142	0.940	0.839
WSVR3	0.633	0.838	0.986	0.970	0.928	1.254	0.947	0.806
WSVR4	0.885	1.151	0.973	0.944	0.910	1.157	0.940	0.835
WSVR5	0.671	0.871	0.985	0.966	0.728	1.027	0.950	0.870
WSVR6	0.727	0.946	0.982	0.960	1.021	1.244	0.959	0.809
WSVR7	0.728	0.960	0.982	0.958	0.860	1.097	0.961	0.852

**Figure 7.** Plots of observed and predicted values of PE in the test stage at Tabriz station.

prediction error of PE values. This may be due to the artificial architecture of ANN models. Figures 7 and 8 show the observed and predicted values of PE in the test stage at Tabriz and Antalya stations, respectively. It is obvious that the predictions of Tabriz station are in better agreement with observed data in comparison with Antalya station. It is deduced that in wet areas, the high degrees of humidity reduces the accuracy of predictions since it has the lowest correlation with PE. Also, Figures 9 and 10 show the observed (x -axis) and predicted (y -axis) PE values for the

best scenarios of each method in the test stage for both studied stations in the form of scatter plots. At Tabriz station, ANN3 seems to be better than the optimal scenarios of the ANN, WANN, SVR and WSVR. In other words, the estimates of ANN3 are closer to the exact line than the others. Also, for Antalya station, the predictions of ANN5 are less scattered than the best scenarios of other mentioned methods. So, the ANN3 and ANN5 with less scattered predictions seem to be the best models for Tabriz and Antalya stations, respectively. It is recognizable from

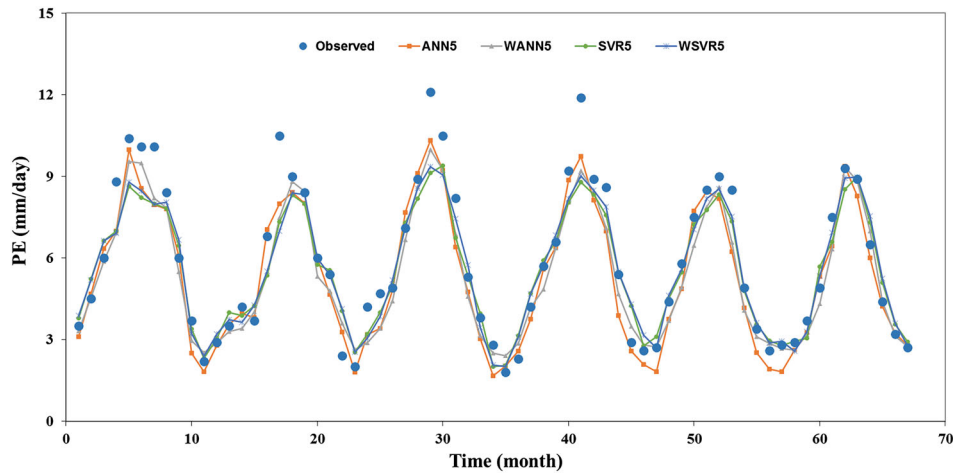


Figure 8. Plots of observed and predicted values of PE in the test stage at Antalya station.

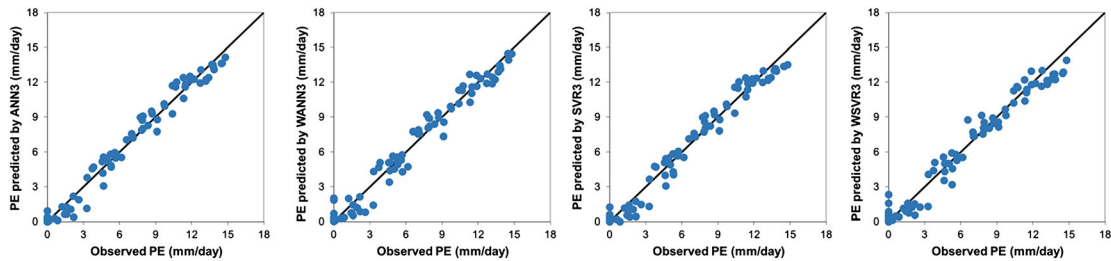


Figure 9. Scatter plots of considered models for PE estimation in the test stage at Tabriz station.

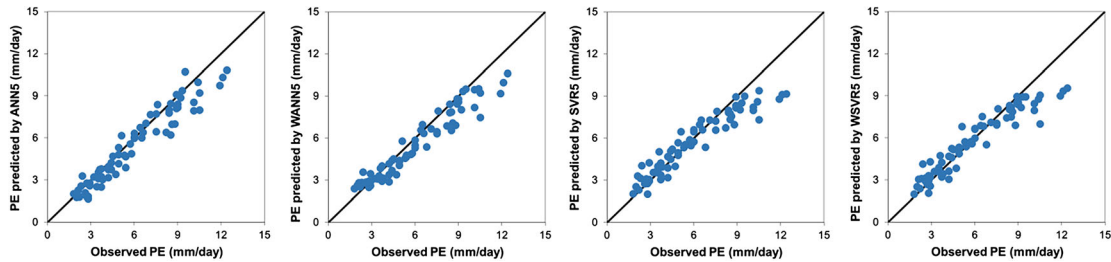


Figure 10. Scatter plots of considered models for PE estimation in the test stage at Antalya station.

Figure 10 that in Antalya station, as much as the observed amount of evaporation increases, the estimated amount deviates from the exact line. So, estimation of evaporation for great values has less accuracy in comparison to low values. However, Figure 9 shows that for Tabriz station this deviation is not noticeable. Moreover, the resulted statistical parameters of the mentioned models in the test period are presented in Table 3. Additionally, the overall results proved that the accuracy of the studied models in Tabriz station was higher than Antalya station. This may be due to the fact that the PE values in arid regions are lower than humid regions and it affects the predictions. The obtained results are parallel to the related literature. Kisi (2009) applied ANN for modeling pan evaporation data of inland (Fresno station) and

coastal stations (Los Angeles and San Diego stations) of USA and he found that the models generally performs better in the inland station when compared to coastal ones. Moreover, comparing the obtained results with the findings of Ghorbani, Kazempour, Chau, Shamshirband, and Taherei Ghazvinei (2018) showed that the RMSE value of ANN-3 in Tabriz station (0.701) is lower than the RMSE value of Hybrid MLP-QPSO Model in Talesh station.

Probability distribution of the observed and predicted data in the test period is presented in Figures 11 and 12. These figures represent the probability occurrence of a specific PE inside a particular interval (Al-Shammari et al., 2016). It can be comprehended that the probability distributions of the predicted values of the ANN3

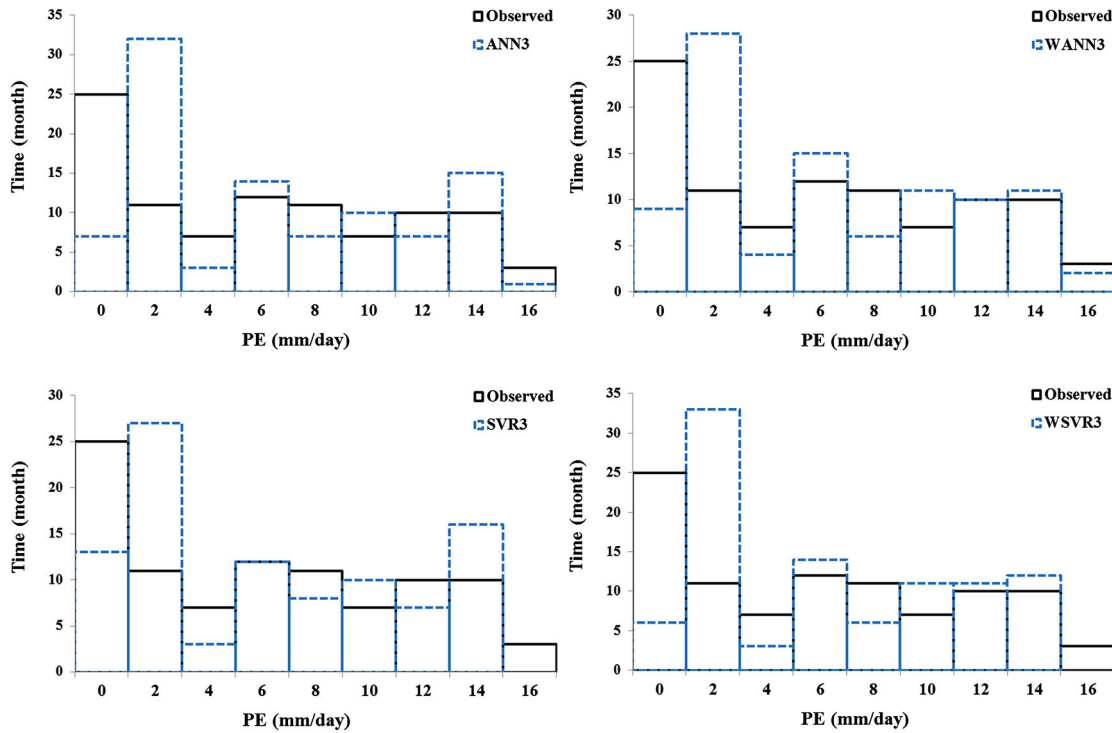


Figure 11. Histograms of pan evaporation at Tabriz station.

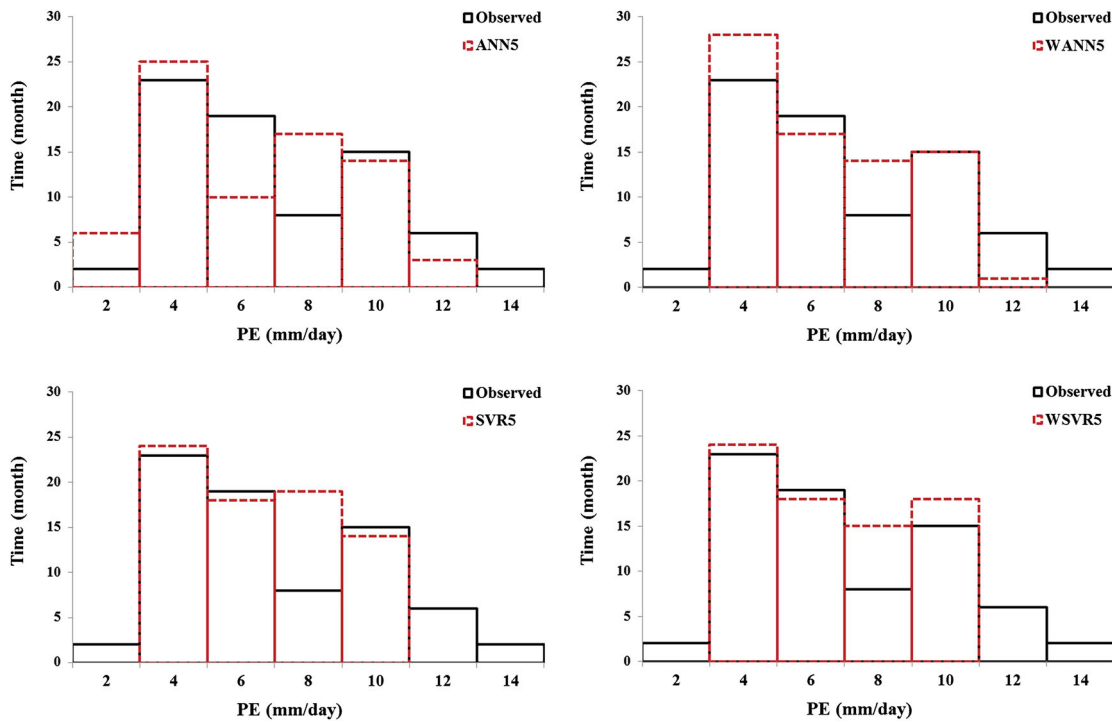


Figure 12. Histograms of pan evaporation at Antalya station.

and SVR3 (Tabriz station) and the ANN5 and WANN5 (Antalya station) are close to the observed values for most intervals.

Furthermore, Figure 13 illustrates Taylor diagrams for the best studied models. In these figures, the radial

length from the green point (i.e. the observed PE value) is a quantity of the RMSE (Taylor, 2001). Consequently, the more precise model is identified by the point with an R value of 1 and the minimum RMSE. It is evident from Figure 13 that ANN3 and ANN5 produced the

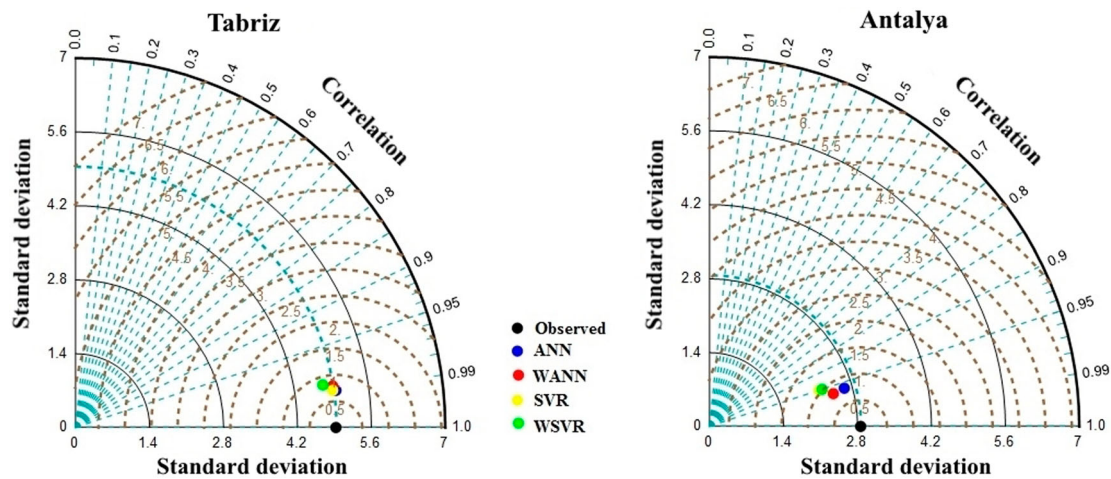


Figure 13. Taylor diagram of the predicted PE for Tabriz and Antalya stations in test period.

best prediction of PE in Tabriz and Antalya stations, respectively.

4. Conclusion

In the current study, the capabilities of ANNs, SVR and also their combination with WT in predicting PE in Tabriz (Iran) and Antalya (Turkey) stations have been examined. A time series data of monthly PE (1992–2016 for Tabriz and 1986–2006 for Antalya) were prepared and 70% of them was selected as training and remaining 30% was used as a testing period. Seven different scenarios including different effective meteorological parameters have been defined for analyzing the effectiveness of each parameter in increasing the prediction accuracies. Also, the performances of studied methods including ANN, WANN, SVR and WSVR, with seven different scenarios for each of them, were comprehensively examined using three statistical parameters, namely RMSE, MAE and R coefficients. Additionally, Taylor diagrams were used where the joint assessment of model correctness using R and RMSE values was carried out. The obtained results showed that at Tabriz station, ANN3, SVR3 and ANN5 had better predictions of PE in comparison to other methods. Also, ANN5, WANN5 and ANN6 proved as the best models for Antalya station. Moreover, the results showed that PE highly depends on temperature and radiation rather than other parameters. As a conclusion, ANN can provide reasonable predictions of PE in both Tabriz and Antalya stations. However, WT did not have a remarkable influence in reducing the prediction errors. Finally, to enrich the effectiveness of the current research, considering and testing the mentioned models in other locations with different climates can be carried out in the future works.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Al-Shammari, E. T., Mohammadi, K., Keivani, A., Hamid, S. H. A., Akib, S., Shamshirband, S., & Petković, D. (2016). Prediction of daily dewpoint temperature using a model combining the support vector machine with firefly algorithm. *Journal of Irrigation and Drainage Engineering*, 142, 04016013. doi:10.1061/(ASCE)IR.1943-4774.0001015
- Chang, F. J., Sun, W., & Chung, C. H. (2013). Dynamic factor analysis and artificial neural network for estimating pan evaporation at multiple stations in northern Taiwan. *Hydrological Sciences Journal*, 58(4), 813–825.
- Chau, K. W. (2017). Use of meta-heuristic techniques in rainfall-runoff modelling. *Water*, 9(3), 186. 6p.
- Choubin, B., Malekian, A., Samadi, S., Khalighi Sigaroodi, S., & Sajedi Hosseini, F. (2017). An ensemble forecast of semi-arid rainfall using large-scale climate predictors. *Meteorological Applications*, 24, 376–386.
- Choubin, B., Zehtabian, G., Azareh, A., Rafiei-Sardooi, E., Sajedi-Hosseini, F., & Kisi, O. (2018). Precipitation forecasting using classification and regression trees (CART) model: A comparative study of different approaches. *Environmental Earth Sciences*, 77(8), 314. doi:10.1007/s12665-018-7498-z
- Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied hydrology*. New York, NY: McGraw-Hill.
- Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K. W., Faizollahzadeh Ardabili, S., & Piran, J. (2018). Survey of computational intelligence as basis to big flood management: Challenges, research directions and future work. *Engineering Applications of Computational Fluid Mechanics*, 12(1), 411–437.
- Ghorbani, M. A., Kazempour, R., Chau, K. W., Shamshirband, S., & Taherei Ghazvinei, P. (2018). Forecasting pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: A case study in Talesh, Northern Iran. *Engineering Applications of Computational Fluid Mechanics*, 12(1), 724–737.

- Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research: Atmospheres*, 113, D06104. doi:10.1029/2007JD008972
- Goyal, M. K., Bharti, B., Quilty, J., Adamowski, J., & Pandey, A. (2014). Modelling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, fuzzy logic, and ANFIS. *Expert Systems with Applications*, 41, 5267–5276.
- Grossman, A., & Morlet, J. (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, 15, 723–736.
- Guyen, A., & Kisi, O. (2013). Monthly pan evaporation modelling using linear genetic programming. *Journal of Hydrology*, 503, 178–185.
- IPCC. (2007). Climate change 2007: The physical science basis. *Intergovernmental Panel on Climate Change*, 446, 727–728.
- 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 Resource Research*, 40(4), W04302.
- Keshtegar, B., Piri, J., & Kisi, O. (2016). A nonlinear mathematical modeling of daily pan evaporation based on conjugate gradient method. *Computers and Electronics in Agriculture*, 127, 120–130.
- Kim, S., & Kim, H. S. (2008). Neural networks and genetic algorithm approach for nonlinear evaporation and evapotranspiration modelling. *Journal of Hydrology*, 351, 299–317.
- Kisi, O. (2005). Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrological Sciences Journal*, 50(4), 683–696.
- Kisi, O. (2006). Daily pan evaporation modelling using a neuro-fuzzy computing technique. *Journal of Hydrology*, 329(3–4), 636–646.
- Kisi, O. (2009). Daily pan evaporation modelling using multi-layer perceptrons and radial basis neural networks. *Hydrological Processes*, 23, 213–223.
- Kisi, O., Genc, O., Dinc, S., & Zounemat-Kermani, M. (2016). Daily pan evaporation modeling using chi-squared automatic interaction detector, neural networks, classification and regression tree. *Computers and Electronics in Agriculture*, 122, 112–117.
- Kisi, O., & Tombul, M. (2013). Modelling monthly pan evaporations using fuzzy genetic approach. *Journal of Hydrology*, 477, 203–212.
- Lin, G. F., Lin, H. Y., & Wu, M. C. (2013). Development of a support-vector-machine-based model for daily pan evaporation estimation. *Hydrological Processes*, 27(22), 3115–3127.
- McCulloch, W. S., & Pitts, W. H. (1943). A logical calculus of the ideas immanent in neural nets. *Bulletin of Mathematical Biology*, 5, 115–133.
- Nourani, V., Alami, M. T., & Aminfar, M. (2009). A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation. *Engineering Applications of Artificial Intelligence*, 22(3), 466–472.
- Nourani, V., Alami, M. T., & Daneshvar Vossoughi, F. (2015). Wavelet-entropy data pre-processing approach for ANN-based groundwater level modelling. *Journal of Hydrology*, 524, 255–269.
- Nourani, V., Hosseini Baghanam, A., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet-artificial Intelligence models in hydrology: A review. *Journal of Hydrology*, 514, 358–377.
- Rajaei, T., Mirbagheri, S. A., Zounemat-Kermani, M., & Nourani, V. (2009). Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Science of the Total Environment*, 407, 4916–4927.
- Rumelhart, D., & McClelland, J. (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Salas, J. D., Markus, M., & Tokar, A. S. (2000). Streamflow forecasting based on artificial neural networks. In G. Rao & A. R. Rao (Eds.), *Chapter 4 in artificial neural networks in hydrology* (pp. 23–51). London: Kluwer Academic Publishers.
- Samadianfard, S., Asadi, E., Jarhan, S., Kazemi, H., Kheshtgar, S., Kisi, O., ... Abdul Manaf, A. (2018). Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths. *Soil and Tillage Research*, 175, 37–50.
- Samadianfard, S., Delirhasannia, R., Kisi, O., & Agirre-Basurko, E. (2013). Comparative analysis of ozone level prediction models using gene expression programming and multiple linear regression. *GEOFIZIKA*, 30, 43–74.
- Samadianfard, S., Nazemi, A. H., & Sadraddini, A. A. (2014). M5 model tree and gene expression programming based modeling of sandy soil water movement under surface drip irrigation. *Agriculture Science Developments*, 3(5), 178–190.
- Samadianfard, S., Sattari, M. T., Kisi, O., & Kazemi, H. (2014). Determining flow friction factor in irrigation pipes using data mining and artificial intelligence approaches. *Applied Artificial Intelligence*, 28, 793–813.
- Shafaei, M., & Kisi, O. (2015). Lake level forecasting using wavelet-SVR, wavelet-ANFIS and WaveletARMA conjunction models. *Water Resource Management*, 30(1), 79–97.
- Smola, A. J., & Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222.
- Stanhill, G. (2002). Is the Class A evaporation pan still the most practical and accurate meteorological method for determining irrigation water requirements. *Agricultural and Forest Meteorology*, 112(3–4), 233–236.
- Tabari, H., Marofi, S., & Sabziparvar, A. A. (2009). Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. *Irrigation Science*, 28(5), 399–406.
- Taormina, R., Chau, K. W., & Sivakumar, B. (2015). Neural network river forecasting through baseflow separation and binary-coded swarm optimization. *Journal of Hydrology*, 529(3), 1788–1797.
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106, 7183–7192.
- URL1. Retrieved from <https://www.google.com/maps/@33.7728123,50.3526526,2214625m/data=!3m1!1e3>.
- Vapnik, V. (1995). *The nature of statistical learning theory*. New York, NY: Springer Verlag.
- Vapnik, V. (1998). *Statistical learning theory*. New York, NY: John Wiley & Sons.
- Wang, W. C., Xu, D. M., Chau, K. W., & Chen, S. (2013). Improved annual rainfall-runoff forecasting using PSO-SVM model based on EEMD. *Journal of Hydroinformatics*, 15(4), 1377–1390.