WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Drought Forecasting using Palmer Drought Severity Index with Wavelet Transform Technique and Machine Learning Methods

Ali Alkan[a]

[a] *Civil Engineering Department, Eskisehir Technical University, Eskisehir, Turkiye*
DOI: https://doi.org/10.55248/gengpi.2023.4158

A B S T R A C T

Drought is a climatic phenomenon characterised by persistent and anomalous water scarcity caused by weather anomalies. Drought is a slowly evolving phenomenon and affects a wide range of areas. The drought indices are one of the ways to monitor and measure drought. Therefore, the Palmer drought severity index (PDSI) was used as an operational and a valid model. This study calculated Palmer Drought Severity Index (PDSI) on a 12-month time scale with meteorological data for the Ceyhan Basin between January 1989 and July 2020. The calculated PDSI values were modelled with Discrete Wavelet Transform (DWT) technique using Artificial Neural Network (ANN) and Support Vector Machine (SVM) methods at different training rates and drought forecasts were made. The forecasting success rate with the models created using machine learning methods was statistically evaluated. This study determined that machine learning methods could be applied in drought forecasting in drought monitoring and mitigation.

Keywords: PDSI, Artificial Neural Network, Support Vector Machine, Hybrid Model, Drought, Discrete Wavelet Transform

1. Introduction

The protection and management of water resources have maintained their importance throughout human history. Nowadays, basin planning and drought prevention studies have gained more importance with the effect of global warming. This situation makes it indispensable to consider the status of our existing water resources. Drought is a natural disaster that seriously affecting a wide range of socio-economic life, especially agriculture, social life, and industry, where long-term lack of precipitation such as snow and sleet leads to water scarcity. There are many definitions of drought. Drought definitions in the literature are organized into six categories related to meteorology, climatology, atmosphere, agriculture, hydrology, and water management. All these views agree that drought is a situation in which a lack of precipitation over a while is caused by a lack of moisture (Bhalme&Mooley, 1980; Wilhite & Glantz, 1985).

Drought indices allow us to comprehensively combine and analyse thousands of recorded data, such as precipitation and runoff, in a single large table. The drought index value is usually a single number and is much more useful for decision-making than raw data. This allows us to compare with drought classifications. Nowadays, indices e.g., Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI) and Palmer Drought Severity Index (PDSI) are commonly used in drought analysis.

The Palmer Drought Severity Index (PDSI) is the one of most outstanding meteorological drought index withinside the United States. Palmer (1965) developed PDSI to quantify the cumulative deviation in the supply and demand (relative to local average conditions) of atmospheric moisture at the surface. It incorporates previous rainfall, moisture supply and moisture demand into a hydrological calculation system based on the classic work of Thornthwaite(1948). Palmer used a two-layer bucket-type model for soil moisture calculations and, on the basis of limited data from the central United States, made some assumptions about the water-holding capacity of the field and moisture transfer between layers (Heim, 2002; Palmer, 1965).

In line with the information in the *Handbook of Drought Indicators and Indices*(Svoboda & Fuchs, 2016) used for drought indices in Turkiye, Standardized Precipitation Index (SPI), Percent of Normal Index (PNI) and the Palmer Drought Severity Index (PDSI) are used. In addition to these methods, Turkish State Meteorological Service described the Aydeniz Drought Method on its official webpage (Turkish State Meteorological Service, n.d.).

Huang et al. (2022) systematically analysed spatio-temporal variations in the characteristics of drought events at the sub-regional scale in the Mongolian Plateau using the Palmer drought theory. Wambua et al. (2017) analysed spatial and temporal drought variability in the Tana River basin using the Palmer Drought Severity Index (PDSI). Fung et al. (2019) performed meteorological drought forecasts using SPEI-based wavelet-boosting-support vector regression, multi-input wavelet-fuzzy-support vector regression, and weighted wavelet-fuzzy-support vector regression. Ozger et al. (2011) developed a model to forecast PDSI data using a wavelet-fuzzy logic model based on meteorological variables. Ashok et al. (2012) examined the relationship between meteorological variables and the PHDI using the Palmer hydrological drought index (PHDI) using a wavelet-Bayes regression

model that increases the modelling power of a simple Bayesian regression model. Belayneh A. et al. (2013) conducted drought forecasting with Standardized Precipitation Index (SPI) based Support Vector Regression (SVR), Artificial Neural Network (ANN) and Wavelet-Artificial Neural Network (W-ANN) machine learning methods. Shamshirband et al. (2020) developed a drought forecasting model using Standardized Runoff Index (SSI), Standardized Precipitation Index (SPI) and Standardized Precipitation Evapotranspiration Index (SPEI) indices with Gene Expression Programming (GEP), Support Vector Regression (SVR) and M5 model trees (MT). Mokhtarzad et al. (2017) conducted drought forecasts with Adaptive Neuro-Fuzzy Interface System (ANFIS), Support Vector Machine (SVM) and Artificial Neural Network (ANN) on the basis of Standardized Precipitation Index (SPI). Khan et al. (2020) made drought forecasts with k-Nearest Neighbour (KNN), Artificial Neural Network (ANN) and Support Vector Machine (SVM) machine learning methods based on Standardized Precipitation Evaporation Index (SPEI).

In this study, aimed to assess the ability of hybrid machine learning models to forecast droughts represented by the PDSI index in the Ceyhan Basin. In order to forecast 12-month time scale PDSI values, hybrid models were created with Discrete Wavelet Transform Technique, Support Vector Machine and Artificial Neural Network. The hybrid models created at different training, test, and validation rates were statistically evaluated.

2. Materials and Methods

2.1 Study Area and Data

The Ceyhan Basin lies between 36° 55′ and 38° 72′ north latitudes and 35° 45′ and 37° 81′ east longitudes in Eastern Mediterranean region of Turkiye. The basin constitutes nearly 2.73% of Turkiye with an area of 21,391 km^2. The Ceyhan Basin extends from the Gulf of Iskenderun to the Southeastern Taurus Mountains, consists of steep mountainous terrain and vast alluvial lands (Republic of Turkiye Ministry of Agriculture and Forestry General Directorate of Water Management, 2019). The Mediterranean climate prevails in the basin, with hot, dry summers and mild, wet winters. Ceyhan Basin borders the Euphrates Basin to the east and northeast, the Seyhan Basin to the west and northwest, and the Asi Basin to the south (See to Fig. 1.).

Drought analysis of the Ceyhan Basin was performed using meteorological data from the General Directorate of Meteorology between January 1989 and July 2020 with various indices. The data were sourced from meteorological stations with the numbers defined by the Turkish State Meteorological Service namely 17355 Osmaniye, 17866 Goksun, 17868 Afsin, 17870 Elbistan, 17908 Kozan, 17960 Ceyhan and 17979 Yumurtalik, respectively.

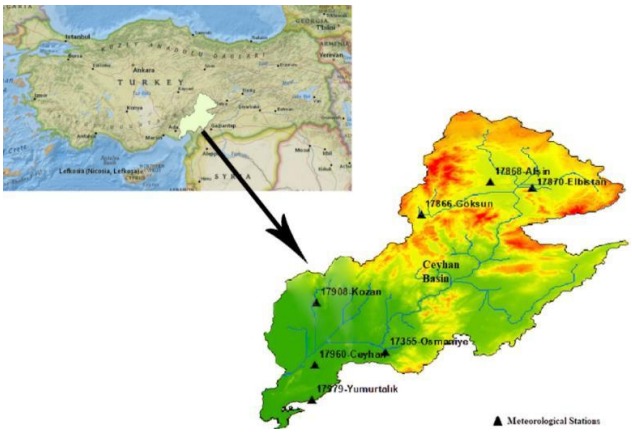

Fig. 1 - Spatial distribution of meteorological stations in Ceyhan Basin.

2.2 The Palmer Drought Severity Index (PDSI)

Palmer (1965) developed the Palmer drought severity index and it can be calculated monthly or weekly. Since this index is used to monitor meteorological drought, it is also called the meteorological drought index. Although the Palmer drought severity index is a meteorological index, it also responds to hydrological and agricultural drought due to the parameters in its algorithm (Akbas, 2014). Palmer, while expressing the drought measure of a place, takes into account which of the class ranges the calculated drought severity value falls within. These class intervals specified by Palmer are given in Table 1 (Karl, 1986).

Table 1 - Classification of dry and wet periods according to PDSI.

PDSI	Classification
≥ 4.00	Extremely Wet
3.00 – 3.99	Very Wet
1.50 – 2.99	Moderately Wet
(-1.49) – (1.49)	Near Normal Wet
(-1.50) – (-2.99)	Moderately Drought
(-3.00) – (-3.99)	Very Drought
≤ – 4.00	Extremely Drought

The first step in calculating PDSI is determining the climatic water balance using long-term monthly precipitation and temperature data as input. Palmer uses an empirical approach to describe the nature of moisture accumulation by separating the soil in two layers. This is the layer where rain falls, and evaporation occurs. Evaporation loss in the upper layer is presumed to occur at the potential level. As long as the moisture in the surface layer remains constantly saturated or evaporates completely, the substrate is not changed. In order to fulfil the water requirement, first potential evapotranspiration must occur, then the soil must be saturated, and as a result, surface runoff must occur. In this method, taking into account the potential evapotranspiration calculated with Thornthwaite, it derives coefficients from the ratios of the actual runoff to the potential values of the monthly averages of the evaporation and evapotranspiration losses occurring in the soil, derives the CAFEC (Climatically Appropriate For Existing Condition) precipitation from these coefficients, uses a series of empirical equations according to the difference of this precipitation from the actual precipitation and determines the severity of drought (Alley, 1984).

In Palmer's method, in addition to precipitation, four surface water data, namely evapotranspiration (*ET*), soil water recharge (*R*), runoff (*RO*) and water loss to the soil layers (*L*), were considered. The complementary potential values of these values are potential evapotranspiration (*PE*), potential recharge (*PR*), potential runoff (*PRO*), and potential loss (*PL*), respectively (Dai, 2011; Wells et al., 2004). Palmer defined CAFEC values for each month *j* as follows:

$$\alpha_j = \frac{\overline{ET_j}}{\overline{PE_j}}, \beta_j = \frac{\overline{R_j}}{\overline{PR_j}}, \gamma_j = \frac{\overline{RO_j}}{\overline{PRO_j}}, \delta_j = \frac{\overline{L_j}}{\overline{PL_j}} \quad (1)$$

where the coefficients describe the ratio of long-term average values between a water flux and its potential value. The CAFEC precipitation ($\hat{P}$), representing the amount of precipitation in order to maintain a normal soil moisture level over a given period, is then generated with the CAFEC potential values as follows (Dai, 2011; Wells et al., 2004):

$$\hat{P} = \alpha_j PE + \beta_j PR + \gamma_j PRO - \delta_j PL \quad (2)$$

2.3 Discrete Wavelet Transform (DWT)

In time series analysis, it is very significant to acquire frequency information. The basis of the wavelet transform is a classical frequency analysis method called the Fourier transform. The Fourier transform technique allows time series to be represented by continuous sine waves and to reach frequency dimensions. Wavelet transforms have been developed over time to overcome the flaws of the Fourier transforms, such as their inability to represent dynamic sequences, their inability to analyse frequency and time information simultaneously, and the inability of the sine wave geometry to capture sudden jumps. The main principle of the wavelet transform is the shifting and scaling operations developed with the wavelet transform, which are also found in the windowing technique (Basakin et al., 2019).

A wavelet is a function that processes the frequency information of a signal at a resolution mapping to its scale. The main idea of the wavelet transform is to analyse signals according to a specified scale (Serroukh et al., 2000). Thanks to their sharp geometry, compressibility, and finite energy properties, wavelet functions determine the changes in time series better than Fourier functions. Wavelet analysis is performed using the time-scale domain. The most important feature of wavelet analysis is its ability to perform local analysis. Compared to traditional methods, wavelet analysis can compress or decompress the signal without distorting the original signal. The Discrete Wavelet Transform technique is used in this study. The mother wavelet functions used in the wavelet transform technique can be calculated by means of the equation depending on the translation and scale parameters. The scale parameter a matches the 1/frequency expression in the Fourier transform. Therefore, small-scale wavelets are useful in determining the detail components with high-frequency information, and large-scale wavelets with high temporal information are useful in determining the trend character of the series. Wavelets are defined as a set of $\psi_{a,b}(t)$ functions obtained by scaling and translating the mother wavelet function $\psi(t)$ (Shao et al., 2000).

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\ \psi\left(\frac{t-b}{a}\right), \qquad a\ \neq 0, \qquad a, b \in R \quad (3)$$

where a is a scale parameter measures the degree of compression, and b is a translation parameter that defines the position of the wavelet (Debnath & Shah, 2015). The wavelet transform is the projection of the function $f(t)$ on the wavelet, i.e., the inner product of the wavelet function $f(t)$ and $\psi_{a,b}(t)$. In the Discrete Wavelet Transform (DWT) analysis technique, wavelet functions are used to decompose the original time series into sub-time series and the signals are analysed in both the time and frequency domains of their processes and relationships devoid of losing information about the moment of the analysed element occurrence (Daubechies, 1990). Labat et al. (2000) developed a mathematical overview of the wavelet transform and a review of its applications. Therefore, the Discrete Wavelet Transform is favoured in most of the estimation problems of water resources systems on account of its shorter computational time (Cannas et al., 2005). For practical applications, hydrologists can access a discrete-time signal instead of a continuous-time signal so that a discrete wavelet may be represented as (Addison et al., 2001):

$$g_{i,j}(t) = \left(\frac{1}{\sqrt{a_0^i}}\right) g\left(\frac{\left(t - j\, b_0 a_0^i\right)}{a_0^i}\right) \quad (4)$$

where i and j are integer values and $b0$ and $a0$ are the location parameter with the specified precise expansion step, respectively. The most common values for a_0 and b_0 are 2 and 1, respectively (Cohen & Kovacevic, 1996).

2.4 Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) are an information processing approach similar to the structure and functioning of the brain. This approximation technique was developed by McCulloch and Pitts (1943) and then gradually advanced in calibration methodologies (Rumelhart et al., 1986). The relational capabilities of ANNs make them more robust to missing and erroneous data, as the knowledge of the relationships among variables is

dispersed over a large number of network connections (Wray et al., 1994). The ANN technique also has the advantage that the modeler does not need to precisely define the intermediate relationships (physical processes) between inputs and outputs (Anctil et al., 2004; Dawson et al., 2006). Thus, ANNs are suitable for analysing complex processes, such as drought forecasting, where the relationships between a large number of input and output variables need to be investigated. Neurons are organised into interconnected groups called layers, with the network structure shown in Figure 2. ANNs contain the following elements, respectively:

The input layer is where data is presented to the network.

Hidden layers are where the data is processed.

The output layer is where the results are found. The given inputs are generated. A neuron calculates the output response on the basis of a weighted sum of all inputs in accordance with an activation function.

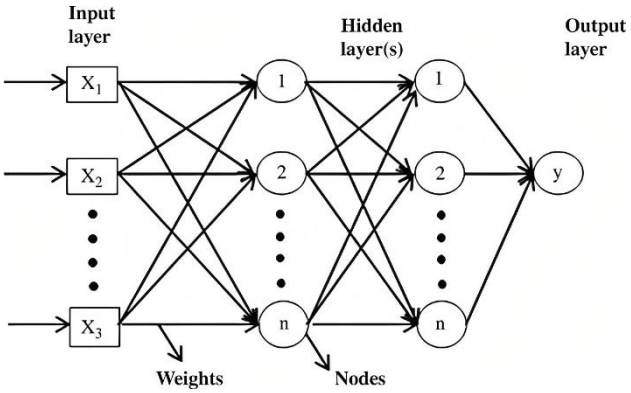

Fig. 2 - Structure of Artificial Neural Network (Sairamya et al., 2019).

The multilayer feed-forward ANN, referred to as Multilayer Perceptron (MLP), has been adopted in numerous applications and studies in water resources management and drought (Lange et al., 1997). An MLP comprises a single input layer, one or more hidden layers consisting of computational nodes and only one output layer. Expression equation for a single output neuron as follows (Kim & Valdés, 2003):

$$\hat{y}_k = f_o\left[\sum_{j=1}^{M} w_{kj} \cdot f_h\left(\sum_{i=1}^{N} w_{ji} x_i + w_{jo}\right) + w_{ko}\right] \quad (5)$$

wheref_o is the activation function for the output layer neuron. f_h is the activation function for the hidden layer neuron; N is the number of input layer variables. M is the number of hidden layer neurons. w_{jo} connects from the degree of j to the neuron of the hidden layer. w_{ji} is the weight connecting the neuron of order i in the input layer and the neuron of order j in the hidden layer. w_{kj} is the weight connecting the neuron of order j in the hidden layer. Output layer w_{ko} is the forecasted output layer neuron judgment at step $\hat{y}_k$ of degree k.

2.5 Support Vector Machine (SVM)

Vapnik(1995) developed the Support Vector Machine (SVM) technique. Support Vector Machines are prominent as an effective learning technique for solving classification and regression problems in various fields such as finance, hydrology and computational biology categorisation in recent years. This is partially regards to the built-in mechanisms to ensure good generalisation leading to accuracy estimations, the capability to train relatively quickly on large datasets using new mathematical optimisation techniques, the use of kernel functions to model non-linear distributions, and most prominently the possibility of theoretical analysis using computational learning theory (Shah, 2007). The main purpose of the statistical learning methods is to find the definition of unknown dependencies between measurements and certain properties of objects. This method assumes that measurements, also known as input variables, can be observed across all relevant objects. Conversely, properties of objects or output variables are generally only available for a small subset of objects known as instances. The objective of estimating the dependence between the elements of the input and output variable is in order to determine the value of the output variable for any of the objects of interest. It is about trying to estimate the value of a function f: $\mathbb{R}^N \longmapsto \{\pm 1\}$ that can accurately categorise new instances on the basis of past observations in pattern recognition (Ertekin, 2009). After estimation of this function, a linear boundary line is used in the separation process and can draw numerous lines. It uses a range for the line that provides the best separation, called the margin. When drawing the margin range, the aim is to find the largest range. In order to find the maximum margin, the following equations must be solved. The structure of SVM is given in Figure 3.

$$Minimisation: min\frac{1}{2}\|w\|^2 \quad (6)$$

$$y_i(w \cdot x_i + b) \geq +1, \quad \forall i \quad (7)$$

where b is the bias, w is the weight vector and x is a point on the hyperplane.

In solving Equation 6, a nonlinear optimisation problem technique is encountered, and this problem is solved using the Lagrange function. Since the problem is quadratic, the optimisation process always results in a global maximum. Then the maximal margin classifier as follows:

$$f(x) = sgn\left(\sum_{i=1}^{l} y_i a_i \langle x_i x\rangle + b\right) \quad (8)$$

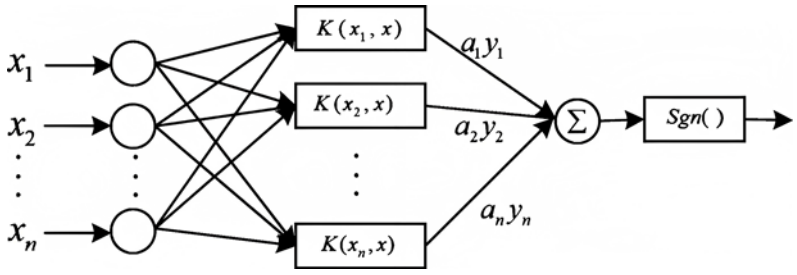

Fig. 3 - Structure of Support Vector Machine (Zhou et al., 2021).

In SVMs, kernel functions e.g., sigmoid, polynomial, radial basis function and linear are used to find the optimal boundary for estimations and solutions between variables with complex relationships. In this study, radial basis function (RBF) is utilised. It can be calculated as follows:

$$K(x_i, x_j) = exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \gamma > 0 \quad (9)$$

where γ is the optimisation parameter related to the kernel width, $x_i$and x_j are vectors in the input space. Also, γ is corresponds to $1/2\sigma^2$, where σ designates the Gaussian noise level of the standard deviation (Gocic et al., 2015).

2.6 Model Performance Assessment

The statistical measures of mean absolute error (MAE) and the Pearson correlation coefficient (R) were used in the training, testing and validation phases to comprehensively assess the accuracy and performance of the drought forecasting models.

$$R = \frac{\sum_{j=1}^{N}(Y_j - \bar{Y})(Y_j - \hat{Y})}{\sqrt{\left[\sum_{j=1}^{N}(Y_j - \bar{Y})^2\right]\left[\sum_{j=1}^{N}(Y_j - \hat{Y})^2\right]}} \quad (10)$$

$$\text{MAE} = \frac{1}{N}\sum_{j=1}^{N}\left|Y_j - \hat{Y}\right| \quad (11)$$

where $\hat{Y}$ is the estimated value of the main variable; Y_j is the observed value in the validation and estimated sets, N indicates the number of samples.

The Pearson correlation coefficient (R) can take a value between -1 to +1. If R is +1, the two variables have a purely positive linear relationship with each other; if R is -1, the two variables have a purely negative linear relationship. If the R value is between 0.5 to 1, it indicates a strong relationship between the variables, and if it is between 0 to 0.5, it indicates a weak relationship between the variables. R = 0 means that the two variables are not related to each other (Tasdemir, 2010; Zhou et al., 2017)

MAE measures the standard deviation of the difference between estimated and observed values. The lower MAE values, the less deviation between estimations and observations (Feng et al., 2011).

3. Results and Discussion

The PDSI drought forecasts were conducted with the Discrete Wavelet Transform (DWT) technique and ANN, SVM methods in accordance with precipitation, minimum and maximum temperature values and 12-month time scale PDSI values obtained from 7 meteorological stations of the Ceyhan Basin between January 1989 and July 2020. The analyses of these methods are abbreviated as DWT-SVM, and DWT-ANN. Potential Evapotranspiration (PET) was calculated from minimum and maximum temperature values. Precipitation (P) and PET values were utilised as input. According to the 12-month PDSI analyses values, the skewness coefficient was 0.765, the standard deviation was 1.964, the median was -0.613 and the arithmetic mean was -0.113 determined. As a result of the analysis, the highest PDSI value determined as 4.62 and the lowest value determined as -5.24. In line with these values, it is observed that the series does not have a symmetric distribution and is not a normally distributed series.

Table 2 - MAE, R Results of DWT-ANN forecasting.

Model	Phase	Ratio	MAE	R	Set	MAE	R
DWT-ANN-1	Training	%70	0,843	0,788	%100	0,810	0,803
	Test	%10	0,763	0,832			
	Validation	%20	0,714	0,848			
DWT-ANN-2	Training	%70	0,843	0,740	%100	0,868	0,756
	Test	%15	0,874	0,797			
	Validation	%15	0,973	0,771			
DWT-ANN-3	Training	%70	0,862	0,756	%100	0,839	0,778
	Test	%20	0,781	0,851			
	Validation	%10	0,779	0,795			
DWT-ANN-4	Training	%70	0,887	0,751	%100	0,872	0,753

	Test	%30	0,837	0,745			
DWT-ANN-5	Training	%80	0,796	0,810	%100	0,814	0,809
	Test	%10	0,939	0,854			
	Validation	%10	0,970	0,750			
DWT-ANN-6	Training	%80	0,862	0,756	%100	0,839	0,778
	Test	%20	0,781	0,851			

The models were created with different training, testing and validation rates. In addition, sets of models without validation were also created. The test, validation and training series in the models are selected randomly within the whole series, not between specific years. Support Vector Machine models are built using the Radial Basis Function (RBF). Artificial Neural Network models are designed as Multi-Layer Perceptron (MLP). R values range between 0.753-0.905 in the models.

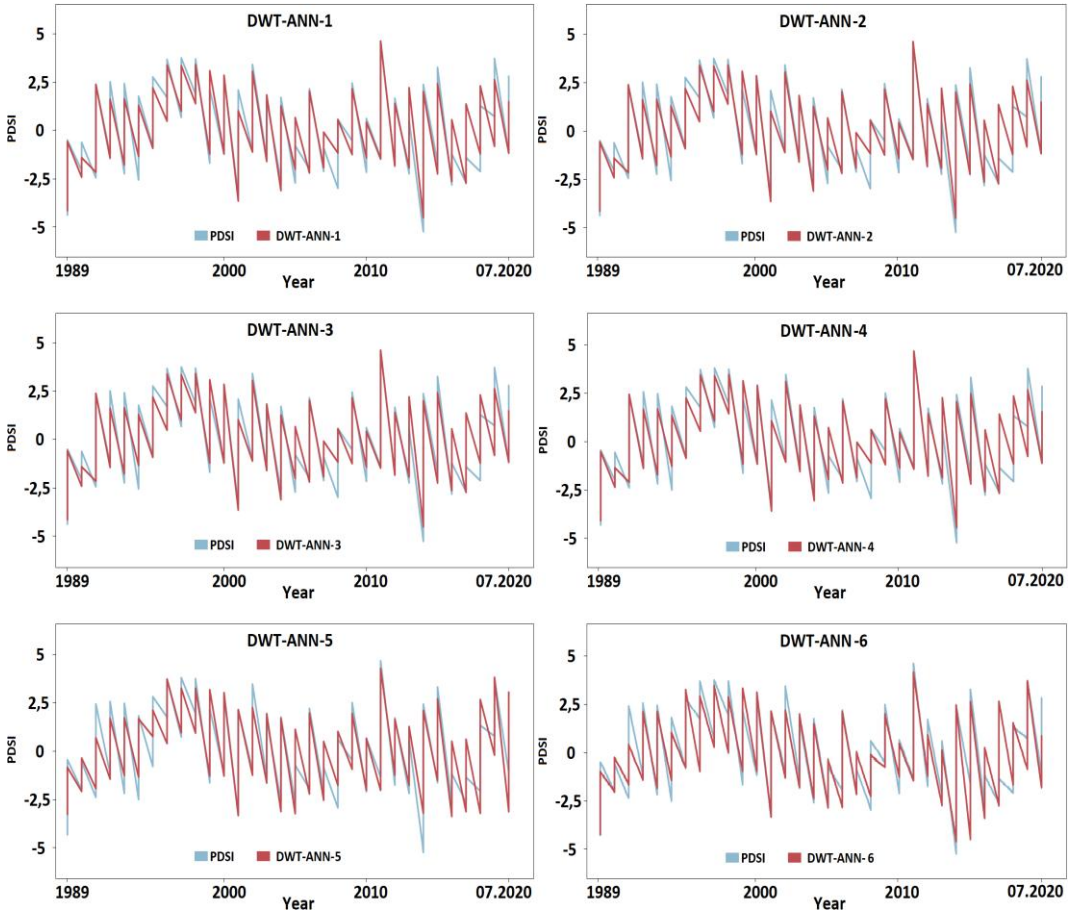

Fig. 4 - Temporal distributions of DWT-ANN drought forecast models.

Table 2 demonstrates the estimation results of PDSI drought index values with DWT-ANN at 12-month time scale. Figure 4 demonstrates the annual drought plots of PDSI forecasts and DWT-ANN drought models on a 12-month time scale. MAE values range between 0.810-0.872 and R values range between 0.753-0.809. In the test phase of the DWT-ANN-6 model, the MAE value increased to 1.009 and caused the performance value of this model to be low. Accordingly, among the models considered in DWT-ANN, the DWT-ANN-5 model with an R value of 0.809 showed higher performance. It is concluded that DWT-ANN-4 has the lowest performance in the model group with an R value of 0.753. In the DWT-ANN-2 training process, the lowest performance R value was 0.740.

Table 3 - MAE, R Results of DWT-SVM forecasting.

Model	**Phase**	**Ratio**	**MAE**	**R**	**Set**	**MAE**	**R**
DWT- SVM -1	Training	%70	0,544	0,844	%100	0,575	0,844
	Test	%10	0,673	0,843			
	Validation	%20	0,642	0,848			
DWT- SVM -2	Training	%70	0,558	0,839	%100	0,574	0,836
	Test	%15	0,513	0,858			
	Validation	%15	0,685	0,810			
DWT- SVM -3	Training	%70	0,636	0,828	%100	0,602	0,832

	Test	%20	0,570	0,810			
	Validation	%10	0,422	0,918			
DWT- SVM -4	Training	%70	0,652	0,822	%100	0,607	0,830
	Test	%30	0,511	0,849			
DWT- SVM -5	Training	%80	0,382	0,913	%100	0,405	0,905
	Test	%10	0,527	0,885			
	Validation	%10	0,498	0,879			
DWT- SVM -6	Training	%80	0,474	0,870	%100	0,447	0,878
	Test	%20	0,324	0,916			

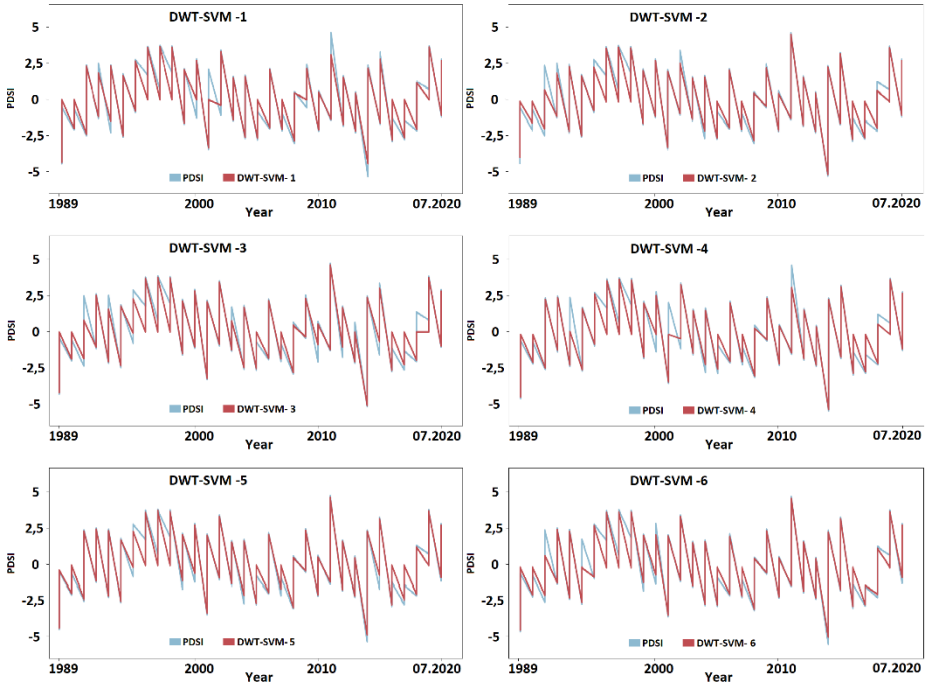

Fig. 5 - Temporal distributions of DWT-SVM drought forecast models.

Table 3 demonstrates the estimation results of PDSI drought index values with DWT-SVM at 12-month scale. Figure 5 demonstrates the annual drought graphs of PDSI forecasts and DWT-SVM drought models for 12-month scale. MAE values range between 0.405-0.607 and R values range between 0.830-0.905. In DWT-SVM analyses, DWT-SVM-5 performed better with an R value of 0.905 among the models considered. It is concluded that DWT-SVM-4 shows low performance among the model group with an R value of 0.830. It is observed that DWT-SVM models have lower MAE values than DWT-ANN models in the training, testing and validation phases. DWT-SVM models were found to have higher correlation coefficient values than DWT-ANN models.

4. Conclusion

Drought forecasting models were created with DWT-ANN and DWT-SVM at the PDSI 12-month time scale in this study. This study shows the variation between the performance of drought hybrid prediction models using wavelet transform and machine learning methods. The models in the study were analysed in terms of performance criteria and it was concluded that the models can be used in drought forecasting. Among the models, it is observed that validated models have higher performance values than non-validated models. In drought analysis using Discrete Wavelet Transform method, SVM models were found to have higher performance and lower error rate than ANN models. In this study, DWT-SVM models gave better results in forecasting than DWT-ANN models. Among the drought forecasting models, the DWT-SVM-5 model has a high performance value with an R value of 0.905. Among all the models in the study, the DWT-ANN-4 model performed low performance with an R value of 0.753.

It is very important to forecast drought in advance and to take the necessary measures to minimise the possible negative impacts of drought. In terms of drought management planning and functioning, future long-term drought forecasts should be considered as an important input parameter. It is thought that it would be useful to apply drought forecasts in other basins by using different drought indices in order to provide more comprehensive information within the scope of preliminary assessment and measures to be taken for drought.

References

Addison, P. S., Murray, K. B., & Watson, J. N. (2001). Wavelet Transform Analysis of Open Channel Wake Flows. Journal of Engineering Mechanics, 127(1), 58–70. https://doi.org/10.1061/(ASCE)0733-9399(2001)127:1(58)

Akbas, A. (2014). Distribution of climatological drought probabilities in Turkey. Turkish Geographical Review, 63, 1–8. https://doi.org/10.17211/tcd.08368

Alley, W. M. (1984). The Palmer Drought Severity Index: Limitations and Assumptions. Journal of Climate and Applied Meteorology, 23(7), 1100–1109. https://doi.org/10.1175/1520-0450(1984)023<1100:TPDSIL>2.0.CO;2

Anctil, F., Michel, C., Perrin, C., &Andreassian, V. (2004). A soil moisture index as an auxiliary ANN input for stream flow forecasting. Journal of Hydrology, 286(1–4), 155–167. https://doi.org/10.1016/j.jhydrol.2003.09.006

Basakin, E. E., Ekmekcioglu, O., &Ozger, M. (2019). Drought Analysis with Machine Learning Methods. Pamukkale University Journal of Engineering Sciences, 25(8), 985–991. https://doi.org/10.5505/pajes.2019.34392

Belayneh, A., &Adamowski, J. (2013). Drought forecasting using new machine learning methods. Journal of Water and Land Development, 18, 3–12. https://doi.org/10.2478/jwld-2013

Bhalme, H. N., &Mooley, D. A. (1980). Large-Scale Droughts/Floods and Monsoon Circulation. Monthly Weather Review, 108(8), 1197–1211. https://doi.org/10.1175/1520-0493(1980)108<1197:LSDAMC>2.0.CO;2

Cannas, B., Fanni, A., Sias, G., Tronci, S., &Zedda, M. K. (2005). River flow forecasting using neural networks and wavelet analysis. Geophysical Research Abstracts, 7, 08651.

Cohen, A., & Kovacevic, J. (1996). Wavelets: the mathematical background. Proceedings of the IEEE, 84(4), 514–522. https://doi.org/10.1109/5.488697

Cortes, C., &Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018

Dai, A. (2011). Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. Journal of Geophysical Research, 116(D12), D12115. https://doi.org/10.1029/2010JD015541

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 36(5), 961–1005. https://doi.org/10.1109/18.57199

Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., &Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. Journal of Hydrology, 319(1–4), 391–409. https://doi.org/10.1016/j.jhydrol.2005.07.032

Debnath, L., & Shah, F. A. (2015). Wavelet Transforms and Their Applications. Birkhäuser Boston. https://doi.org/10.1007/978-0-8176-8418-1

Ertekin, S. (2009). Learning in Extreme Conditions: Online and Active Learning with Massive, Imbalanced and Noisy Data. Pennsylvania State University.

Feng, Y., Zhang, W., Sun, D., & Zhang, L. (2011). Ozone concentration forecast method based on genetic algorithm optimized back propagation neural networks and support vector machine data classification. Atmospheric Environment, 45(11), 1979–1985. https://doi.org/10.1016/j.atmosenv.2011.01.022

Fung, K. F., Huang, Y. F., & Koo, C. H. (2019). Coupling fuzzy–SVR and boosting–SVR models with wavelet decomposition for meteorological drought prediction. Environmental Earth Sciences, 78(24), 693. https://doi.org/10.1007/s12665-019-8700-7

Gocic, M., Motamedi, S., Shamshirband, S., Petković, D., Ch, S., Hashim, R., &Arif, M. (2015). Soft computing approaches for forecasting reference evapotranspiration. Computers and Electronics in Agriculture, 113, 164–173. https://doi.org/10.1016/j.compag.2015.02.010

Heim, R. R. (2002). A Review of Twentieth-Century Drought Indices Used in the United States. Bulletin of the American Meteorological Society, 83(8), 1149–1166. https://doi.org/10.1175/1520-0477-83.8.1149

Huang, Y., Liu, B., Zhao, H., & Yang, X. (2022). Spatial and Temporal Variation of Droughts in the Mongolian Plateau during 1959–2018 Based on the Gridded Self-Calibrating Palmer Drought Severity Index. Water, 14(2), 230. https://doi.org/10.3390/w14020230

Karl, T. R. (1986). The Sensitivity of the Palmer Drought Severity Index and Palmer's Z-Index to their Calibration Coefficients Including Potential Evapotranspiration. Journal of Climate and Applied Meteorology, 25(1), 77–86. https://doi.org/10.1175/1520-0450(1986)025<0077:TSOTPD>2.0.CO;2

Khan, N., Sachindra, D. A., Shahid, S., Ahmed, K., Shiru, M. S., & Nawaz, N. (2020). Prediction of droughts over Pakistan using machine learning algorithms. Advances in Water Resources, 139, 103562. https://doi.org/10.1016/j.advwatres.2020.103562

Kim, T.-W., & Valdés, J. B. (2003). Nonlinear Model for Drought Forecasting Based on a Conjunction of Wavelet Transforms and Neural Networks. Journal of Hydrologic Engineering, 8(6), 319–328. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:6(319)

Labat, D., Ababou, R., &Mangin, A. (2000). Rainfall–runoff relations for karstic springs. Part II: continuous wavelet and discrete orthogonal multiresolution analyses. Journal of Hydrology, 238(3–4), 149–178. https://doi.org/10.1016/S0022-1694(00)00322-X

Lange, N., Bishop, C. M., & Ripley, B. D. (1997). Neural Networks for Pattern Recognition. Journal of the American Statistical Association, 92(440), 1642–1645. https://doi.org/10.2307/2965437

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics, 5(4), 115–133. https://doi.org/10.1007/BF02478259

Mishra, A. K., & Singh, V. P. (2012). Simulating Hydrological Drought Properties at Different Spatial Units in the United States Based on Wavelet–Bayesian Regression Approach. Earth Interactions, 16(17), 1–23. https://doi.org/10.1175/2012EI000453.1

Mokhtarzad, M., Eskandari, F., JamshidiVanjani, N., &Arabasadi, A. (2017). Drought forecasting by ANN, ANFIS, and SVM and comparison of the models. Environmental Earth Sciences, 76(21), 729. https://doi.org/10.1007/s12665-017-7064-0

Ozger, M., Mishra, A. K., & Singh, V. P. (2011). Estimating Palmer Drought Severity Index using a wavelet fuzzy logic model based on meteorological variables. International Journal of Climatology, 31(13), 2021–2032. https://doi.org/10.1002/joc.2215

Palmer, W. C. (1965). Meteorological Drought. In U.S. Weather Bureau, Res. Pap. No. 45.

Republic of Turkiye Ministry of Agriculture and Forestry General Directorate of Water Management. (2019). Ceyhan HavzasiKuraklikYonetimiPlani [Ceyhan Basin Drought Management Plan].

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536. https://doi.org/10.1038/323533a0

Sairamya, N. J., Susmitha, L., Thomas George, S., &Subathra, M. S. P. (2019). Hybrid Approach for Classification of Electroencephalographic Signals Using Time–Frequency Images With Wavelets and Texture Features. In Intelligent Data Analysis for Biomedical Applications (pp. 253–273). Elsevier. https://doi.org/10.1016/B978-0-12-815553-0.00013-6

Serroukh, A., Walden, A. T., & Percival, D. B. (2000). Statistical Properties and Uses of the Wavelet Variance Estimator for the Scale Analysis of Time Series. Journal of the American Statistical Association, 95(449), 184–196. https://doi.org/10.2307/2669537

Shah, R. S. (2007). Support Vector Machines for Classification and Regression. McGill University.

Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., Kargar, K., Mosavi, A., Nabipour, N., & Chau, K. W. (2020). Predicting Standardized Streamflow index for hydrological drought using machine learning models. Engineering Applications of Computational Fluid Mechanics, 14(1), 339–350. https://doi.org/10.1080/19942060.2020.1715844

Shao, X., Pang, C., &Su, Q. (2000). A novel method to calculate the approximate derivative photoacoustic spectrum using continuous wavelet transform. Fresenius' Journal of Analytical Chemistry, 367(6), 525–529. https://doi.org/10.1007/s002160000404

Svoboda, M., & Fuchs, B. (2016). Handbook of Drought Indicators and Indices (pp. 155–208). https://doi.org/10.1201/9781315265551-12

Tasdemir, S. (2010). Determination of body measurements on the holstein cows by digital image analysis method and estimation of their live weight. Selcuk University.

Thornthwaite, C. W. (1948). An Approach toward a Rational Classification of Climate. Geographical Review, 38(1), 55–94. https://doi.org/10.2307/210739

Turkish State Meteorological Service. (n.d.). KuraklikAnalizi [Drought Analysis]. Retrieved January 24, 2023, from https://mgm.gov.tr/veridegerlendirme/kuraklik-analizi.aspx?d=yontemsinif

Wambua, R. M., Mutua, B. M., &Raude, J. M. (2017). Analysis of spatial and temporal drought variability in a tropical river basin using Palmer Drought Severity Index (PDSI). International Journal of Water Resources and Environmental Engineering, 9(8), 178–190. https://doi.org/10.5897/IJWREE2017.0723

Wells, N., Goddard, S., & Hayes, M. J. (2004). A Self-Calibrating Palmer Drought Severity Index. Journal of Climate, 17(12), 2335–2351. https://doi.org/10.1175/1520-0442(2004)017<2335:ASPDSI>2.0.CO;2

Wilhite, D. A., & Glantz, M. H. (1985). Understanding: the Drought Phenomenon: The Role of Definitions. Water International, 10(3), 111–120. https://doi.org/10.1080/02508068508686328

Wray, B., Palmer, A., &Bejou, D. (1994). Using Neural Network Analysis to Evaluate Buyer ‐ Seller Relationships. European Journal of Marketing, 28(10), 32 ‐ 48. https://doi.org/10.1108/03090569410075777

Zhou, F., Zou, L., Liu, X., Zhang, Y., Meng, F., Xie, C., & Zhang, S. (2021). Microlandform classification method for grid DEMs based on support vector machine. Arabian Journal of Geosciences, 14(13), 1269. https://doi.org/10.1007/s12517-021-07596-0

Zhou, T., Wang, F., & Yang, Z. (2017). Comparative Analysis of ANN and SVM Models Combined with Wavelet Preprocess for Groundwater Depth Prediction. Water, 9(10), 781. https://doi.org/10.3390/w9100781.