Italian Journal of Agrometeorology
Rivista Italiana di Agrometeorologia

Firenze University Press
www.fupress.com/ijam

Citation: Ö. Baydaroğlu Yeşilköy, K. Koçak, L. Şaylan (2020) Prediction of commonly used drought indices using support vector regression powered by chaotic approach. *Italian Journal of Agrometeorology* (2): 65-76. doi: 10.13128/ijam-970

Received: June 18, 2020

Accepted: October 17, 2020

Published: January 25, 2021

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Prediction of commonly used drought indices using support vector regression powered by chaotic approach

Özlem Baydaroğlu Yeşilköy[1,*], Kasim Koçak[2], Levent Şaylan[2]

[1] *Altınbaş University, School of Engineering and Natural Sciences, Department of Civil Engineering, İstanbul/Turkey*
[2] *İstanbul Technical University, Faculty of Aeronautics and Astronautics, Department of Meteorological Engineering, İstanbul/Turkey*
*Corresponding author. E-mails: ozlembaydaroglu@gmail.com, kkocak@itu.edu.tr; saylan@itu.edu.tr

Abstract. An effective water resources management requires accurate predictions of possible risks. Drought is one of the most devastating phenomena that has a certain risk of occurrence. Understanding the variability of the drought indices is of great importance in determining the spatiotemporal behavior of the drought phenomenon. Moreover, determination of the variability and short-term prediction of the drought indices enables us to take necessary steps in hydrological and agricultural issues. In this study, drought indices have been predicted via Support Vector Regression, SVR. This method originated from a linear regression method in a high dimensional feature space. SVR necessitates a special input matrix. In this study, this matrix has been constructed on the basis of Chaotic Approach, CA. Commonly used drought indices are used in the prediction stage. These indices consist of monthly Palmer Drought Severity Index, PDSI, Palmer Hydrological Drought Index, PHDI, Palmer Z-Index, ZNDX, Modified Palmer Drought Severity Index, PMDI, and Standard Precipitation Index, SPI. One-step ahead prediction has been realized for a 36-month period. Most results show that predictions of the drought indices using SVR are quite promising.

Keywords. Drought indices, prediction, phase space reconstruction, machine learning.

1. INTRODUCTION

It is a well-known fact that drought is one of the most important natural disasters in the world. Severity and duration of droughts in different regions of the world are expected to increase in the future due to climate change. Most countries will be affected by drought at different levels depending on the risk factors they have (Carrao et al., 2016). Drought is a natural disaster with the power to produce significant social and economic consequences. For this reason, better management of water resources will gain more importance in the future. In drought studies, the first step is the determination of drought's level and its variation and predicting the future value of the drought indices. Drought indices are used for determination and classification of the drought (Yihdego et al., 2019; Hao et al., 2016).

In literature, there are some studies about prediction of the drought indices. Liu et al. (2009) have implemented the Markov Chain Model to predict the PDSI. They constructed a one-step prediction model which produces different prediction accuracy such as 96.68%, 64.45%, 52.6% and 0.40% for normal, slight, moderate, and severe drought conditions, respectively. Mehta et al. (2014) have used the Model for Interdisciplinary Research on Climate 5 (MIROC5) and Statistical Forecast System (SPS) in order to predict the Self-Calibrating Palmer Drought Severity Index (SC-PDSI). Although correlation coefficients and root mean square errors are big enough with regard to the time period and grid spacing, the accuracy of the prediction is within the acceptable limits. Cutore et al. (2009) employed neural networks and climate indices so as to predict the Palmer Indices. They found that the North Atlantic Oscillation (NAO) series are uncorrelated with the Palmer Indices for winter months while the European Blocking (EB) are correlated with the Palmer Indices both for winter and autumn months. Belayneh et al. (2014) have used a wavelet-neural network and wavelet-support vector regression approaches to predict long-term SPI series. The best prediction result is produced from the coupled wavelet neural network.

The SVR which is an advanced, state-of-the-art prediction method can be seen as a specific implementation of Support Vector Machines (SVMs) to the regression problem (Vapnik, 1995; Cortes and Vapnik, 1995). The SVM is a kind of statistical learning machine which is widely used in an area of classification (Vapnik and Lerner, 1963; Vapnik and Chervonenkis, 1964). SVR transforms input space which is formed from the observations into high dimensional feature space by way of a kernel function and performs a linear regression in this space.

There are many applications of SVR about prediction of many variables. Yu et al. (2006) have used SVR for real-time flood stage prediction. This study has demonstrated that a SVR has strong prediction performance. Santamaria-Bonfil et al. (2016) have developed a method based on the SVR to predict wind speeds for wind farms. They showed that the SVR is more appropriate for short term wind speed and wind power values prediction than persistence and autoregressive models. SVR has been employed to predict hourly O_3 concentrations by Ortiz-Garcia et al. (2010) with accurate prediction results. Baydaroğlu and Koçak (2014) have used SVR algorithm to predict evaporation values. The results show that SVR-based predictions are very successful with high determination coefficients as 83% and 97% for univariate and multivariate time series embeddings, respectively. Moreover, river flow prediction using hybrid models of the SVR with Wavelet Transform (WT), Singular Spectrum Analysis (SSA) and CA is realized by Baydaroğlu et al. (2017). The SVR-WT combination has resulted in the highest coefficient of determination and the lowest mean absolute error. Granata et al. (2016) have applied SVR for a simulation of rainfall-runoff processes in two experimental basins and compared with EPA's stormwater management model. The hydrograph shape and the time to peak are correctly modelled by these approaches. It can be said that the SVR shows considerable potential for applications to the problems of urban hydrology.

The main idea behind this study is to predict drought indices consisting of many nonlinear variables with high accuracy. Prediction of drought indices enables decision makers to monitor all components of hydrologic cycle, gain simple information about different kinds of droughts which are complex phenomena, determine economic impacts, risks and changes on agricultural productivity, plan irrigation facilities and water distribution. In the prediction part of the study, SVR has been employed. It requires a special input data format. Chaotic Approach (CA) is utilized in order to prepare the input data set for SVR. For this purpose, a phase space is reconstructed by using Embedding Theorem (Takens, 1981). According to this theorem, time delays and embedding dimensions are determined from the time series in question. In this study, False Nearest Neighbour (FNN) (Kennel et al., 1992) and Mutual Information Function (MIF) (Fraser and Swinney, 1986) have been implemented to determine the embedding dimension and the time delay, respectively.

In the organization of this paper, Section 2 introduces the methods and material used in the study, Section 3 discusses the results obtained in the study.

2. MATERIAL AND METHODS

The following are employed in this study: (1) Determining the optimum embedding dimensions and time delays (2) A phase space is constructed with these embedding parameters in order to prepare input matrix for SVR (3) Prediction of the drought indices using SVR.

FNN algorithm is based on the definition of true and false neighboring points in a phase space. Percentages of the false neighboring points in a successively higher dimensional phase space provide to develop this algorithm to choose an optimum embedding dimension. The Mutual Information Function (MIF) can be considered as a nonlinear counterpart of the Autocorrelation Function (ACF).

2.1. Data

The American Meteorological Society (1997) categorizes droughts as meteorological or climatological, agricultural, hydrological and socioeconomic droughts. Wilhite and Glantz (1985) are expressed numerically some indices based on historical climate records such as temperature and precipitation.

In this study, US Palmer Drought Severity Index (PDSI), Palmer Hydrological Drought Index (PHDI), Palmer Z-Index (ZNDX), Modified Palmer Drought Severity Index (PMDI) and Standard Precipitation Index (SPI) are considered. These indices were calculated from monthly data between June 1929 and December 2015 provided by National Oceanic and Atmospheric Administration (NOAA) (downloaded from https://www7.ncdc.noaa.gov/CDO/ CDO DivisionalSelect.jsp)

The PDSI is widely and operationally used for determination of drought status (Palmer, 1965). This index is based on a soil moisture balance between supply and demand. In other words, it is a function of precipitation, temperature, and available water content of the soil (Palmer, 1965; Alley, 1984). It is generated indicating the severity of wet or dry spells. This index is above +4 and below -4. Negative values indicate dry spells while positive values denote wet spells.

The PHDI is a monthly hydrological drought index used to assess long-term moisture supply and indicates the severity of a wet or dry spell like the PDSI. This index considers the information about precipitation as inflow, outflow and storage. Increased irrigation, new reservoirs, and added industrial water use are not typically included in the computation of this index. The index generally ranges from - 6 to +6. The major disadvantage of this index is that it does not take the long-term precipitation trend in consideration (Karl and Knight, 1985).

The ZNDX essentially measures the moisture anomaly. It specifies a deviation from the normal monthly PDSI. This index can respond to a month of above-normal precipitation, even during periods of drought.

The PMDI is derived from the PDSI having the difference with respect to transition periods between dry and wet spells. The PMDI is based on a weighting factor for wet and dry indices (Heddingshaus and Sabol, 1991). An index value is selected as the PDSI drought index. This selection is realized by the program regarding probabilities (see https://www7.ncdc. noaa.gov/CDO/CDODivisionalSelect.jsp).

The Surface Water Supply Index (SWSI) is calculated by using the components of precipitation, snowpack (in winter), stream flow (in summer) and reservoir storage inputs. Monthly data are used for the computation of the SWSI. McKee et al. (1993) stated that the SPI is another well-known and frequently used meteorological drought index in application. This index is a transformation of the probability of observing a given amount of precipitation in given months (see https://www7.ncdc. noaa.gov/CDO/CDODivisionalSelect.jsp).

The Global Historical Climatology Network (GHCN) Daily dataset is the source of station data. GHCN-Daily contains several major observing networks in North America. The primary network is the National Weather Service (NWS) Cooperative Observing (COOP) program. These data update every day from a variety of data streams with quality checks. Moreover the data are reconstructed each weekend from its data source components (see detail https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn).

Drought indices used in the study are given in the Figures 1 (a) to (k). These figures clearly reveal the erratic behavior of the drought indices.

2.2. Phase Space Reconstruction

Prediction of drought indices by way of SVR requires a special set of input data matrix. In this study, chaos theory is utilized to construct the input data matrix. The phase spaces of drought indices have been reconstructed using the most appropriate embedding parameters.

To estimate optimum embedding dimensions, a method proposed by Cao (1997) has been applied to all series. For this purpose, nonlinearTseries package in R-Studio has been used (Cao, 1997; Arya and Mount, 1993; Arya et al., 1998). Embedding dimension values are determined from the points which $E1(d)$ and $E2(d)$ stay constant together.

To reconstruct the phase space we need the proper time delays, . In the application, this parameter is extracted from both autocorrelation and mutual information functions (MIF). In this study, we have chosen the MIF to decide the proper time delay because of its flexibility in measuring both linear and nonlinear inner-dependences in a given series. To estimate optimum time delay, the fractal package in R-Studio has been used (Kantz and Schreiber, 1997; Bassingthwaighte et al., 1994; Fraser and Swinney, 1986; Casdagli et al., 1991).

The details of phase space reconstruction from a univariate or single variable time series is given in Packard et al. (1980). A phase space can be reconstructed by using delay coordinate method. Coordinates of the phase space are spanned by the variables which are necessary to specify the time evolution of the system (Koçak et al., 2004). With regard to Embedding Theorem (Takens, 1981), a phase space can be constructed from a one-dimensional

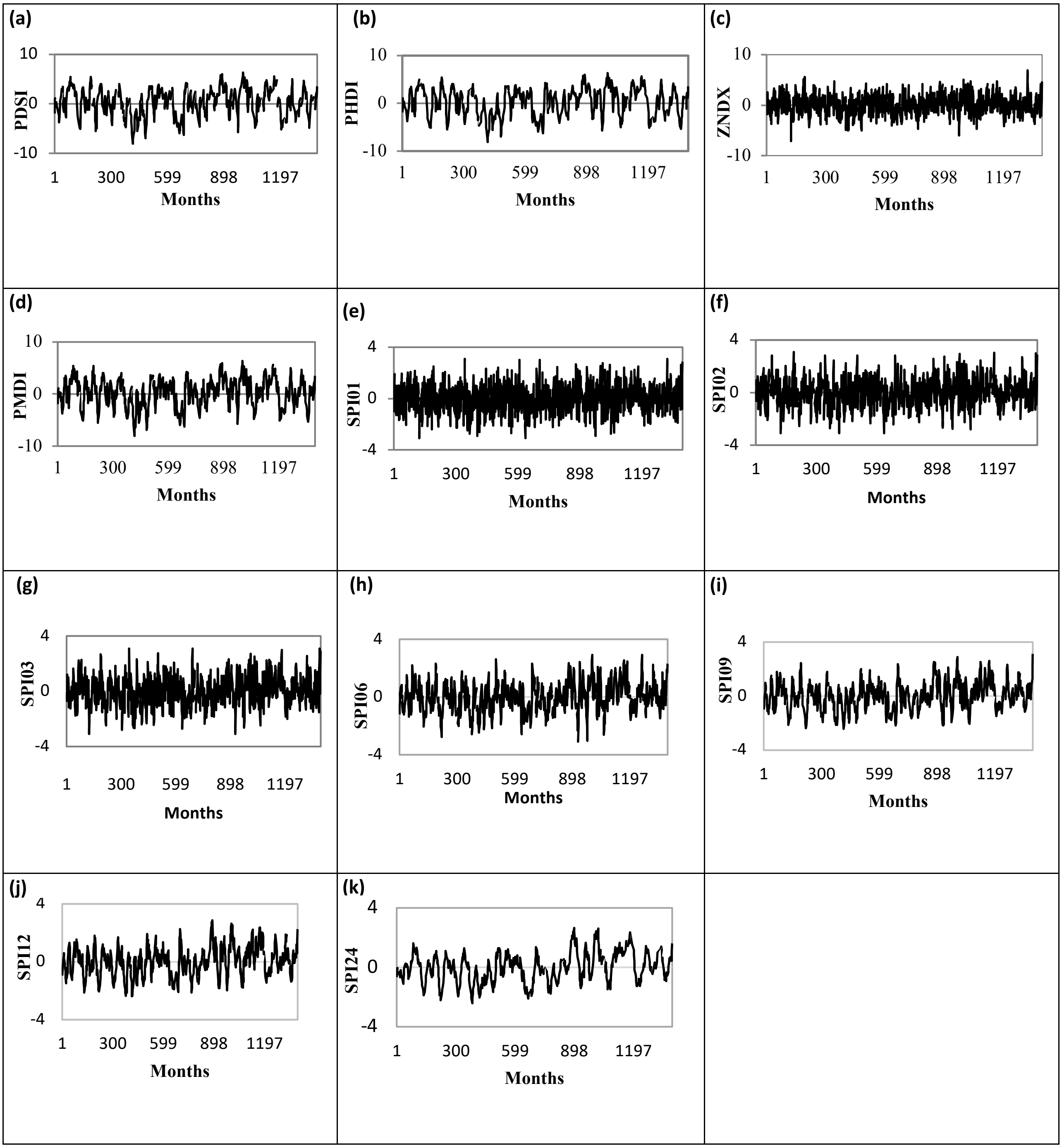


Fig. 1. Monthly (a) PDSI (b) PHDI (c) ZNDX (d) PMDI e) 1-month SPI (f) 2-month SPI (g) 3-month SPI (h) 6-month SPI (i) 9-month SPI (j) 12-month SPI (k) 24-month SPI.

series. Let's consider a time series $x_i \in \mathbb{R}$, $i=1,2,\ldots,N$, then the reconstruction procedure for the purpose of prediction is given as

$$X_i=(x_i,x_{i-\tau},\ldots,x_{i-(d-1)\tau}) \in \mathbb{R}^d,\ i=1,2,\ldots,N-(m-1)\tau \qquad (1)$$

where X_i is an d-dimensional phase space vector, τ is a time delay and m is an embedding dimension.

In the d-dimensional space, phase space vectors describe an object which is topologically equivalent to the attractor of the physical system (Porporato and Ridolfi, 1997).

2.2.1. False Nearest Neighbor (FNN) Method

To reconstruct a phase space, embedding dimensions and time delays should be determined from a given time series. In literature, there are some methods like False Nearest Neighbor (Kennel et al., 1992) and Grassberger-Procaccia (GP) methods (Grassberger and Procaccia, 1983).

In this study, another algorithm proposed by Cao (1997) has been implemented to determine the minimum embedding dimension. For a time series $x_1, x_2, \ldots, x_N$, the time delay vectors reconstructed,

$$X_i(d)=(x_i, x_{i+\tau}, \ldots, x_{i+(d-1)\tau}),\ i=1,2,\ldots,N-(d-1)\tau \tag{2}$$

Similar to FNN algorithm,

$$a(i,d)=\frac{\|X_i(d+1)-X_{n(i,d)}(d+1)\|}{\|X_i(d)-X_{n(i,d)}(d)\|},\quad i=1,2,\ldots,N-d\tau \tag{3}$$

where || || is some measurement of Euclidean distance and the maximum norm of the distance

$$\|X_k(d)-X_l(d)\|=\max_{0\le j\le d-1}|x_{k+j\tau}-x_{l+j\tau}| \tag{4}$$

$X_i(d+1)$ is the ith reconstructed vector with embedding dimension $d+1$, i.e., $X_i(d+1)=(x_i, x_{i+\tau}, \ldots, x_{i+d\tau})$; $n(i,d)$, $(1\le n(i,d)\le N-d\tau)$, is an integer such that $x_{n(i,d)}(d)$ is the nearest neighbour of $x_i(d)$ in the d-dimensional reconstructed phase space in the sense of distance || ||.

The threshold value should be determined by the derivative of the underlying signal, therefore, for different phase point i, $a(i,d)$ should have different threshold values. Also, different time series may have different threshold values. For this problem, Cao (1997) defines the mean value of all $a(i,d)$'s

$$E(d)=\frac{1}{N-d\tau}\sum_{i=1}^{N-d\tau}a(i,d) \tag{5}$$

To investigate its variation from d to $d+1$

$$E1(d)=\frac{E(d+1)}{E(d)} \tag{6}$$

$E1(d)$ stops changing when d is greater than some value d_0 if the time series comes from an attractor. Then, the d_0+1 is the minimum embedding dimension.

Another quantity which is useful to distinguish deterministic signals from stochastic signals can be developed by using $E^*(d)$

$$E^*(d)=\frac{1}{N-d\tau}\sum_{i=1}^{N-d\tau}|x_{i+d\tau}-x_{n(i,d)+d\tau}| \tag{7}$$

Finally, $$E2(d)=\frac{E^*(d+1)}{E^*(d)} \tag{8}$$

For time series data from a random set of numbers, $E1(d)$, in principle, will never attain a saturation value as d increases. In practical computations, it is difficult to resolve whether the $E1(d)$ is slowly increasing or has stopped changing if d is sufficiently large. Because available observed data samples are limited, it may happen that $E1(d)$ stops changing at some d although the time series is random. To solve this problem, $E2(d)$ can be considered. For random data, because the future values are independent of the past values, $E2(d)$ will be equal to 1 for any d in this case. However, for deterministic data, $E2(d)$ is certainly related to d, as a result, it cannot be a constant for all d; in other words, there must exist some d's such that $E2(d)\neq 1$.

2.2.2. Mutual Information Function (MIF)

Another important embedding parameter for reconstructing a phase space is a time delay. Various methods such as Mutual Information Function (MIF) (Fraser and Swinney, 1986), autocorrelation function, Cross Autocorrelation (CAC) (Palit et al., 2013), C-C Method (Kim et al., 1999) based on correlation integral can be used to determine proper time delay. MIF which is a nonlinear counterpart of autocorrelation function is the most commonly used approach in nonlinear time series analysis. The time corresponding to the first minimum of the MIF gives the optimum value for the time delay.

2.2.3. Support Vector Regression

Vapnik and Lerner (1963) and Vapnik and Chervonenkis (1964) have developed SVM algorithm. SVM is a state-of-the-art method which provides a good generalization capability to dynamics of a given process thanks to Structural Risk Minimization (SRM) approach. SVR, an application of SVM to the regression problem, is based on the computation of a linear regression function in a multidimensional feature space.

For a set of k examples $[(x_1,y_1),(x_2,y_2),\ldots,(x_k,y_k)]$, each generated from an unknown probability distribution $P(x,y)$ where x_i are the input vectors and y_i are the corresponding output values $(i=1,2,\ldots,k)$, the best approximate function of the possible smallest risk can be given as (Liong and Sivapragasam, 2002; Yu et al., 2006),

$$R(\alpha)=\int(y-f(x,\alpha))^2 dP(x,y) \tag{9}$$

where $f(x,\alpha)$ is a class of functions and α is the parameter of this function.

$R(\alpha)$ cannot be calculated because $P(x,y)$ is unknown. For this reason, an induction principle for risk minimization should be taken into consideration. The approach which replaces the probable smallest risk $R(\alpha)$ by the empirical risk $R_{emp}(\alpha)$ is given by

$$R_{emp}(\alpha) = \frac{1}{k}\sum_{i=1}^{k}\left(y - f(x,\alpha)\right)^2 \quad (10)$$

This approach is named Empirical Risk Minimization (ERM) induction principle. However, ERM does not guarantee a small $R(\alpha)$ if the number of training data is limited. Therefore, SRM principle based on statistical learning theory has been developed by Vapnik (1999).

The SRM principle theoretically minimizes $R(\alpha)$ based on the simultaneous minimization of both the emprical risk and the confidence interval Ω (Yu et al., 2006). The bound on $R(\alpha)$ is given by

$$R(\alpha) \leq R_{emp}(\alpha) + \Omega\left(\frac{k}{h}\right) \quad (11)$$

where the parameter h is called the VC-dimension (Vapnik, 1995) of a set of functions. It can be seen as the capacity or the complexity of a set of functions.

The learning machine is given a training data set $\{x_i, y_i\}$, i=1,…,k, $y_i \in \mathbb{R}, x \in \mathbb{R}^D$. The regression function assumed for this data set is a linear regression on the hyperplane

$$y_i = wx_i + b \quad (12)$$

where w is the weight vector, y_i is the element of {+1,-1} and b is the bias or deviation. In case of real observations, most of the processes exhibit nonlinearity. Therefore, linear approaches may not be practical and effective. When considering nonlinearity, the input data, x, in the input space is mapped to high dimensional feature space using nonlinear function $\phi(x)$ and the decision function is given by

$$f(w,b) = w.\phi(x) + b \quad (13)$$

If data are not separated linearly, then slack variables are inserted to the optimization problem. Then the regression problem can be converted into the following convex optimization problem (Yu et al., 2006).

$$\min_{w,b,\xi^+,\xi^-} \frac{1}{2}w^2 + C(\xi^+ + \xi^-) \quad (14)$$

subject to $y_i - (w.\phi(x_i) + b) \leq \varepsilon + \xi^+$; $(w.\phi(x_i) + b) - y_i \leq \varepsilon + \xi^-$; $\xi^+, \xi^- \geq 0$, i=1,2,…,k

where ξ^+ and ξ^- are slack variables that indicate the upper and lower training errors subject to an error tolerance ε. C is the penalty factor which is a balance between the training error and model complexity. To find Lagrange multipliers α_i^+ and α_i^- which are necessary to solve convex optimization problem, the following function is maximized

$$L_D = \sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-)y_i - \epsilon\sum_{i=1}^{L}(\alpha_i^+ - \alpha_i^-) - \frac{1}{2}\sum_{i,j}(\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)\,\phi(x_i).\phi(x_j) \quad (15)$$

subject to $\sum_{i=1}^{k}(\alpha_i^+ - \alpha_i^-) = 0$; $0 \leq \alpha_i^+ \leq C$, i=1,2,…,k; $0 \leq \alpha_i^- \leq C$, i=1,2,…,k

In Eq. (15), $\phi(x_i).\phi(x_j)$ is a kernel function, $K(x_i,x_j)$. There are various kernel functions such as linear, polynomial, radial basis and sigmoid. The application of SVR requires the selection of an appropriate kernel function. Radial Basis Function (RBF) is the most commonly used kernel function because of its flexibility in applications (see Baydaroğlu and Koçak, 2014; Baydaroğlu et al., 2017; Yu and Liong, 2007; Belayneh et al., 2014; Ortiz-Garcia et al., 2010). Besides, it has a strong learning ability and is able to reduce computational complexity of the training process and improve the generalization performance of SVR (Li and Xu, 2005). In the study, RBF has been chosen as the kernel function

$$K(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2g^2}} \quad (16)$$

where g is the width of radial basis function. The final decision function of nonlinear SVR can be given by

$$f(x_i) = \sum_{i=1}^{k}(\alpha_i^+ - \alpha_i^-)\,K(x_i, x_k) + b \quad (17)$$

where x_k is a new entry for the estimation.

2.3. Prediction

In general, prediction from a time series can be performed by using three different approaches. These are one-step prediction, direct multi-step prediction and indirect multi-step prediction (Wang et al, 2010). Let η shows the prediction horizon or predefined time interval used for how far ahead the model predicts the future. Then the abovementioned prediction methods can be expressed as given below.

a. One-step prediction (η=1)

$$X(t+1) = F(X(t)) \quad (18)$$

b. Direct multi-step prediction ($\eta > 1$)

The prediction value of $t + \eta$ can be estimated according to the historical measurement data, and the prediction equation is

$$X(t + \eta) = F[X(t)] \quad (19)$$

c. Indirect multi-step prediction ($\eta > 1$)

Using iteration method to process one-step prediction model, and the prediction equation is

$$X(t + \eta) = F(F\{\cdots F[X(t)]\}) \quad (20)$$

In application, one-step prediction is the most frequently used method due to its simplicity and its high accuracy. Thus, throughout this study, one-step prediction was preferred to other prediction approaches and run for 36 months (three years). This time period is enough to evaluate the accuracy of the prediction method.

2.4. Performance Criteria

In this study, Mean Absolute Error (MAE) (Willmott and Matsuura, 2005), Relative Error (RE) (Golub and Charles, 1996) and coefficient of determination (R^2) (Steel and Torrie, 1960) are calculated to estimate the prediction performance.

Let N, x_i, y_i, $\bar{x}$, $\bar{y}$ denote the total number of observations, observed values, forecasts, the arithmetic means of the observed and forecasted values, respectively.

$$MAE = \frac{1}{N}\sum|y_i - x_i| \quad (21)$$

$$RE = \frac{|y_i - x_i|}{x_i} \quad (22)$$

$$R^2 = \frac{\{\sum(x_i-\bar{x})(y_i-\bar{y})\}^2}{\sum(x_i-\bar{x})^2(y_i-\bar{y})^2} \quad (23)$$

3. RESULTS

The recent studies on climate change indicate that drought will be one of the most important challenges in the future. To facilitate the drought related analysis, some drought indices have been developed to judge numerically whether current conditions fall within the drought limits or not. Thus, predictions of drought indices will enable us to take some necessary precautions to mitigate the possible effects of drought related damages.

All drought indices have been predicted for a 36-month prediction period by way of one-step ahead prediction. Prediction of drought indices is a very complicated task since this kind of data show quite high variability. Therefore, SVR that is a state-of-the-art method when comparing other counterparts is chosen to predict drought indices.

Tab. 1. Optimum time delays and embedding dimensions for drought indices.

Drought Indices	Time Delays	Embedding Dimensions	Drought Indices	Time Delays	Embedding Dimensions
PDSI	17	9	SP02	2	10
PHDI	17	9	SP03	2	9
ZNDX	17	9	SP06	6	10
PMDI	2	10	SP12	14	8
SP01	2	10	SP24	20	8

Tab. 2. SVR parameters used in the predictions of the drought indices.

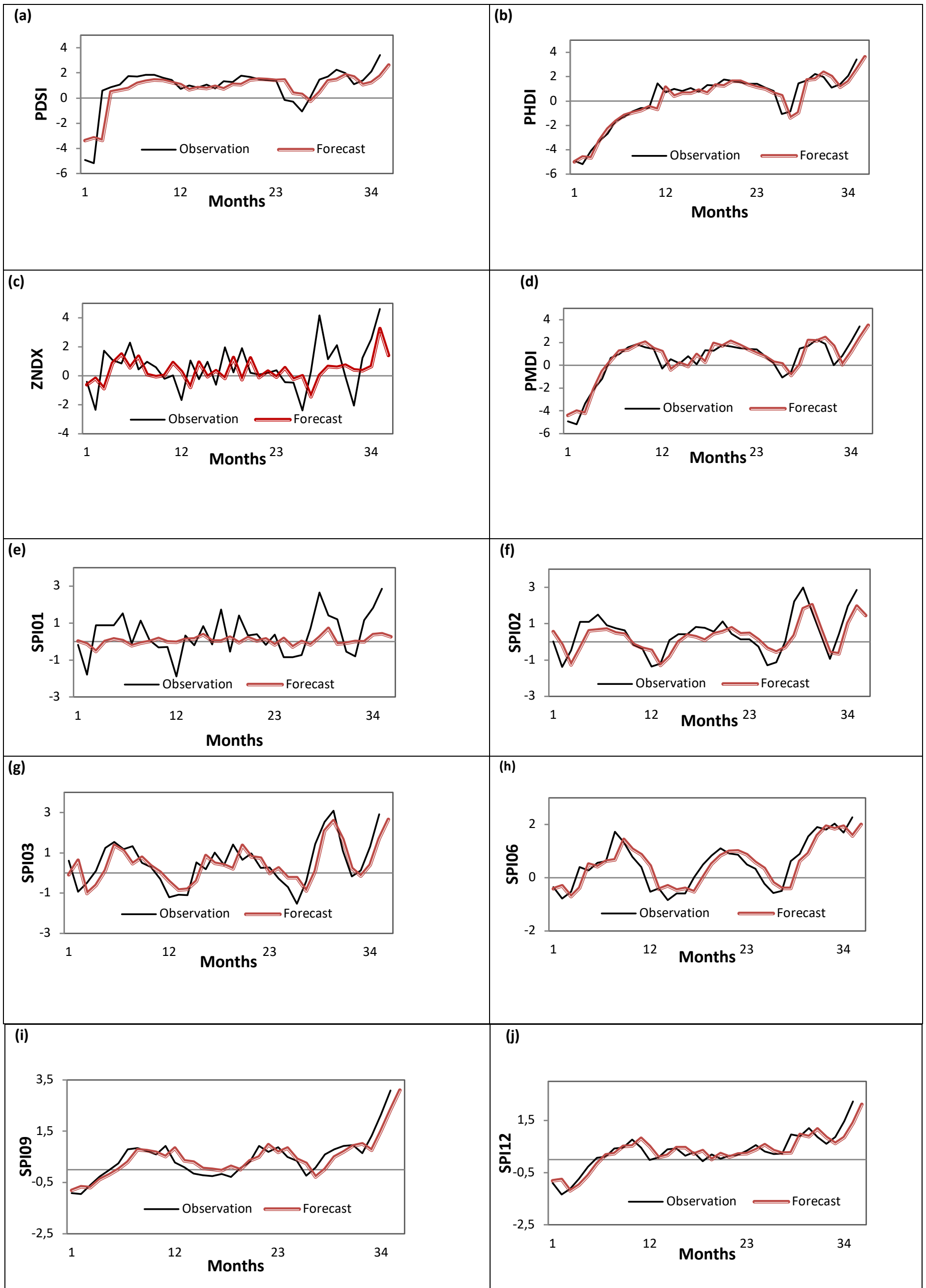
SVR Parameters	C	ε	g
PDSI	0.3536	0.0044	0.0060
PHDI	8.0000	0.0229	0.7711
ZNDX	197.4029	0.0001	0.5452
PMDI	1.8340	0.0499	1.6818
SP01	256.0000	0.0004	0.0482
SP02	98.7015	0.0031	0.4585
SP03	256.0000	0.1294	0.1621
SP06	1.4142	0.1187	1.0000
SP09	279.1699	0.1996	0.0078
SP12	256.0000	0.1088	0.0060
SP24	58.6883	0.0593	0.3855

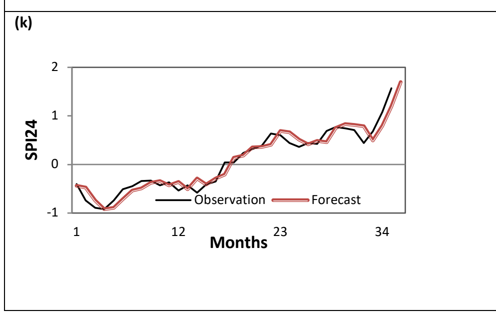
To prepare an input matrix for SVR, optimum embedding parameters have been estimated as seen in Tab. (1).

After the proper determination of embedding parameters phase spaces for the drought indices can be reconstructed. Then by using these phases, the input data matrix can be formed properly. In this case, any input data matrix consists of two parts. One is a training data set and the other is a test data set. In addition, the input data set also includes the target data column. This column indicates the next or future values of the predicted variable.

From embedding parameters given in Tab. 1, phase spaces are reconstructed.

Tab. 2 shows the SVR parameters used in the predictions of drought indices. As mentioned before, the parameter C given in the second column of the table controls the trade-off between the slack variable penalty and the size of the margin. SVR uses a more sophisticated penalty function than the SVM, not allocating a penalty if the predicted value is less than a distance ε away from the actual value.





(l)

Indices	MAE	RE	R^2
PDSI	0.004	0.36	0.88
PHDI	0.002	0.24	0.95
ZNDX	↑ **0.010**	↑ **0.78**	↓ **0.30**
PMDI	0.002	0.25	0.94
SPI01	↑ **0.016**	↑ **0.88**	↓ **0.24**
SPI02	0.008	0.68	0.55
SPI03	0.008	0.65	0.60
SPI06	0.004	0.51	0.74
SPI09	0.002	0.39	0.86
SPI12	0.002	0.40	0.84
SPI24	0.001	0.37	0.86

Fig. 2. Prediction results of (a) PDSI (b) PHDI (c) ZNDX (d) PMDI (e) SPI01 (f) SPI02 (g) SPI03 (h) SPI06 (i) SPI09 (j) SPI12 (k) SPI24 (l) Prediction performance of the drought indices.

The last parameter, *g*, is the width of the Gaussian kernel function. These parameters are calculated for their valid ranges which are determined by using a Fortran code and optimum values of parameters given in Tab. 2 are used for the implementation of SVR.

The application of SVR requires the selection of an appropriate kernel function. In the study, RBF has been chosen as the kernel function. RBF is the most commonly used kernel function because of its flexibility in applications (see Baydaroğlu and Koçak, 2014; Baydaroğlu et al., 2017; Yu and Liong, 2007; Belayneh et al., 2014; Ortiz-Garcia et al., 2010). Besides, it has a strong learning ability and is able to reduce computational complexity of the training process and improve the generalization performance of SVR (Li and Xu, 2005).

Predictions of the last 36-month period of the drought indices are given in Fig. 2(a) to (k). In this fig., normalized prediction and indices values have been used. Besides, statistics of the performance criteria of the predictions can be seen in Fig. 2(l).

As stated before, drought is the most devastating natural phenomenon that causes serious loss of life and property. Therefore, prediction of drought occurrence is extremely important for decision makers. In the current literature, there are various drought indices and each one is developed for different purposes. Instead of focusing on a specific drought index, most of the well-known drought indices are considered in the prediction process

The importance of drought indices' prediction originates from the cruciality of drought prediction. Calculation of drought indices consists of many meteorological and hydrologic parameters. Therefore, prediction of these indices has a complex and difficult challenge. Since accurate drought prediction enables people to take necessary precautions for agricultural sustainability, disaster management and plan water management, agricultural activities, a hybrid prediction method based on SVR powered by Chaos Theory has been used in this study. It is a well-known fact that there is an inverse relationship between variability and prediction accuracy. Although drought indices have high variability, high generalization capability of SVR leads to predictions with high accuracy.

Fig. 2 (l) shows prediction performance criteria, MAE, RE, R^2. High MAE and RE values and low R^2 values indicate medium and low prediction accuracy and vice versa.

If $RE=0$, prediction is perfect while $RE=1$ means that only the average value is predicted. From Fig. 2 (l), it is seen that ZNDX and SP01 have the highest relative errors. Namely, these indices have been predicted on their average levels. Similarly, it can be easily seen that ZNDX and SPI01 are indices which are difficult to predict as seen similarly on MAEs and R^2 values of ZNDX and SPI01.

Willmott and Matsuura (2005) indicate that MAE is a good natural measure of average error. Fig. 2 (l) shows that maximum MAE values of all indices belong to ZNDX and SPI01. Moreover, ZNDX and SPI01 have the smallest R^2 values.

As a result, PDSI, PHDI, PMDI, SP09, SP12 and SP24 can be predicted with high accuracy when considering all performance criteria. These indicators show that all Palmer drought indices can be predicted accurately using SVR.

In summary, the results show that the proposed hybrid method can accurately predict drought indices but ZNDX and SPI01. The low predictability of these indices is quite understandable. Because ZNDX is limited to just one month, it is a short-term drought measure without memory from previous months (see https://www.cisa.sc.edu/atlas/glossary.html). Similarly, it may be difficult to obtain the required number of data for high prediction accuracy of SPI01 since it shows the anomalies of the observed total precipitation for a month. Obviously, prediction errors decrease as the numbers of the month increase for SPIs. Predictions for Palmer Indices and long-term SPIs are very promising.

4. CONCLUSION

One of the most important features that distinguish the drought phenomenon from other natural disasters is of its slowly developing feature over time. The second important feature of drought events is that once it has started, its devastating effects continue for a relatively long period of time. Drought has the ability to deeply affect all sectors and leaves permanent marks on the life of society. For this reason, it is a natural disaster that deserves further study. There are many drought indices which are widely used in the monitoring of drought events. Therefore, in order to minimize the negative effects of drought events, it is of vital importance to realize accurate predictions of drought indices.

SVMs were developed on the basis of statistical learning. It had outstanding advantages in theory and it realized the nonlinear mapping of the high-dimensional space by kernel function, and it was used to solve nonlinear classification and regression estimation problems (Li et al., 2019). On the other hand, Chaos Theory has made a great success in many fields of pure and applied sciences, especially in atmospheric sciences, climatology, economy, fluid mechanics, hydrology, medical sciences, etc. In this study, the two state of the art techniques were applied to the prediction of drought indices.

SVR is a product of SVM developed to be applied to the regression problems. In SVR, the most important step of the application is to form the input matrix. In the current literature, there is no standard procedure in the construction of input matrix. CA enables the scientists to use a standard method to construct the input matrices which plays an important role in the success of the SVR method.

In the application stage, the above mentioned methods were applied to five frequently used drought indices for prediction purposes. Except for ZNDX and SPI01, which fluctuate almost randomly, the results are very encouraging. The results show clearly that the proposed hybrid method produces very accurate one-step predictions of drought indices. The proposed method has the ability to produce the prediction higher than one-step. Undoubtedly, such predictions will provide significant flexibility to the decision makers in terms of taking the necessary measures on time.

ACKNOWLEDGEMENT

This article was presented in American Meteorological Society (AMS)-27th Conference on Weather Analysis and Forecasting in Chicago, July 2nd, 2015.

REFERENCES

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

American Meteorological Society, (1997) Meteorological Drought-Policy statement. Bulletin of the American Meteorological Society 78: 847-849.

Arya, S., Mount, D. M., (1993). Approximate nearest neighbor searching. Proc. 4th Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA'93) 271-280.

Arya, S., Mount, D.M., Netanyahu, N. S., Silverman, R., Wu, A. Y., (1998). An optimal algorithm for approximate nearest neighbor searching fixed dimensions. Journal of the ACM 45 891-923. doi: 10.1145/293347.293348

Bassingthwaighte, J. B., & Raymond, G. M. (1994). Evaluating rescaled range analysis for time series. *Annals of biomedical engineering*, *22*(4), 432-444.

Baydaroğlu, Ö., Koçak, K., (2014). SVR-based prediction of evaporation combined with chaotic approach. Journal of Hydrology 508 356-363. doi: 10.1016/j.hydrol.2013.11.008

Baydaroğlu, Ö., Koçak, K., Duran, K. (2017). River flow prediction using hybrid models of support vector regression with the wavelet transform, singular spectrum analysis and chaotic approach. Meteorology and Atmospheric Physics 1-11. doi: 10.1007/s00703-017-0518-9

Belayneh, A., Adamowski, J., Khalil, B., Ozga-Zielinski, B., (2014). Long-term SPI Drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression model. Journal of Hydrology 508 418-429. doi: 10.1016/j.hydrol.2013.10.052

Cao, L., (1997). Practical method for determining the minimum embedding dimension of a scalar time series. Physica D: Nonlinear Phenomena 110 1 pp. 43-50. doi: 10.1016/S0167-2789 (97)00118-8

Carrão, H., Naumann, G., & Barbosa, P., (2016). Mapping global patterns of drought risk: An empirical framework based on sub-national estimates of hazard, exposure and vulnerability. *Global Environmental Change*, *39*, 108-124.

Casdagli, M. (1989). Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, *35*(3), 335-356.

Cortes, C., Vapnik, V., (1995). Support vector networks. Machine Learning vol. 20 pp. 273-297. doi: 10.1007/BF00994018

Cutore, P., Di Mauro, G., Cancelliere, A., (2009). Forecasting Palmer Index using neural networks and climate indexes. Journal of Hydrologic Engineering Vol. 14 No.6. doi: 10.1061/(ASCE)HE.1943-5584.0000028#sthash.7USLO2fJ.dpuf

Fraser, A. M., Swinney, H. L., (1986). Independent coordinates for strange attractors from mutual information. Phys. Rev. A Vol. 33 Number 2. doi: 10.1103/PhysRevA.33.1134

Golub, G., Charles, F. V. L, (1996). Matrix Computations-Third Edition. Baltimore. The John Hopkins University Press. P.53. ISBN 0-8018-5413-X

Granata, F., Gargano, R., De Marinis, G., (2016). Support vector regression for rainfall-runoff modeling in urban drainage: A comparison with the epa's storm water management model. Water 8(3) 69. doi: 10.3390/w8030069

Grassberger, P., Procaccia, I., (1983). Estimation of the Kolmogorov entropy from a chaotic signal. Physical Review A 28(4): 2591-2593. doi: 10.1103/PhysRevA.28.2591

Hao, Z., Hao, F., Singh, V. P., Xia, Y., Ouyang, W., & Shen, X., (2016). A theoretical drought classification method for the multivariate drought index based on distribution properties of standardized drought indices. *Advances in water resources*, *92*, 240-247.

Heddinghaus, T. R., & Sabol, P. (1991, September). A review of the Palmer Drought Severity Index and where do we go from here. In *Proc. 7th Conf. on Applied Climatology* (pp. 242-246). Boston, MA: American Meteorological Society

Kantz, H., & Schreiber, T. (1997). *Nonlinear time series analysis*. Cambridge University press.

Karl, T. R., Knight, R. W., (1985). Atlas of Monthly Palmer Hydrological Drought Indices (1931–1983) for the Contiguous United States. Historical Climatology Series 3–7 National Climatic Data Center Asheville North Carolina.

Kennel, M. B., Brown, R., Abarbanel, H. D. I., (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys. Rev. A 45 3403-3411. doi: 10.1103/PhysRevA.45.3403

Kim, H. S., Eykholt, R., Salas, J. D., (1999). Nonlinear dynamics, delay times and embedding windows. Physica D 127 48-60. doi: 10.1016/S0167-2789(98)00240-1

Koçak, K., Şaylan, L., Eitzinger, J., (2004). Nonlinear prediction of near-surface temperature via univariate and multivariate time series embedding. Ecological Modelling 1-7. doi: 10.1016/S0304-3800(03)00249-7

Li, X., Luo, A., Li, J., & Li, Y. (2019). Air pollutant concentration forecast based on support vector regression and quantum-behaved particle swarm optimization. *Environmental Modeling & Assessment*, *24*(2), 205-222.

Li, P. C., Xu, S. H., (2005). Support vector machine and kernel function characteristic analysis in pattern recognition. Computer Engineering and Design vol. 26 pp. 302-304.

Liong, S., Sivapragasam, C., (2002). Flood stage forecasting with support vector machines. Journal of the American Water Resources Association Vol.38 No.1. doi: 10.1111/j.1752-1688.2002.tb01544.x

Liu, X., Ren, L., Yuan, F., Yang, B., (2009). Meteorological drought forecasting using Markov Chain Model. 2009 International Conference on Environmental Science and Information Application Technology. doi: 10.1109/ESIAT.2009.19

McKee, T. B., Doesken, N. J., Kleist, J., (1993). The Relationship of Drought Frequency and Duration to Time Scales. Proceedings of the Eighth Conference on Applied Climatology American Meteorological Society: Boston: 179–184.

Mehta, V. M., Wang, H., Mendoza, K., Rosenberg, N. J., (2014). Predictability and prediction of decadal hydrological cycles: A case study in Southern Africa. Weather and Climate Extremes 3 47-53. doi: 10.1016/j.wace.2014.04.002

Ortiz-Garcia, E. G., Salcedo-Sanz, S., Perez-Bellido, A. M., Portilla-Figueraz, J. A., Prieto, L., (2010). Prediction on hourly O_3 concentrations using support vector regression algorithms. Atmospheric Environment 44 4481-4488. doi: 10.1016/j.atmosenv.2010.07.024

Packard, N. H., Crutchfield, J. P., Farmer, J. D., Shaw, R. S., (1980). Geometry from a time series. Physical Review Letters 45 (9). doi: 10.1103/PhyRevLett.45.712

Palit, S. K., Mukherjee, S., Bhattacharya, D. K., (2013). A high dimensional delay selection for the reconstruction of proper phase space with cross auto-correlation. Neurocomputing 113 49-57. doi: 10.1016/j.neucom.2013.01.034

Palmer, W. C., (1965). Meteorological Drought. Res. Paper No.45 58pp. Dept. of Commerce Washington D.C.

Porporato, A., Ridolfi, L., (1997). Nonlinear Analysis of near-wall turbulence time series. Applied Scientific Research 57: 235-261. doi: 10.1007/BF02506062

Santamaria-Bonfil, G., Reyes-Ballesteros, A., Gershenson, C., (2016). Wind speed forecasting for wind farms: A method based on support vector regression. Renewable Energy 85 790-809. doi: 10.1016/j.renene.2015.07.004

Steel, R. G. D., Torrie, J. H., (1960). Principles and Procedures of Statistics with Special Reference to the Biological Sciences. McGraw Hill.

Takens, F., (1981). Detecting strange attractors in turbulence. In: Rand, DA Young, LS (Eds.) Lecture Notes in Math. Springer-verlag pp. 366-381.

Vapnik, V., Lerner, A., (1963). Pattern recognition using generalized portrait method. Automation and Remote Control 24: 774-780.

Vapnik, V., Chervonenkis, A., (1964). A note on one class of perceptrons. Automation and Remote Control vol. 25.

Vapnik, V. N., (1995). The nature of statistical learning theory. Springer New York.

Vapnik, V. N., (1999). An overview of statistical learning theory. Neural Networks IEEE Transactions on Vol.10 No.5 988-999. doi: 10.1109/72.788640

Wang Y, Wu D.L., Guo C.X., Wu Q.H., Qian W.Z. and Yang J., (2010). Short-term wind speed prediction using support vector regression. IEEE PES General Meeting, Providence, RI, pp. 1-6, doi: 10.1109/PES.2010.5589418.

Wilhite, D. A., Glantz, M. H., (1985). Understanding the drought phenomenon: The role of definitions. Water International 10 (3): 111–120. doi: 10.1080/02508068508686328

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, *30*(1), 79-82.

Yihdego, Y., Vaheddoost, B., & Al-Weshah, R. A. (2019). Drought indices and indicators revisited. *Arabian Journal of Geosciences*, *12*(3), 69.

Yu, P., Chen, S., Chang, I., (2006). Support vector regression for real time flood stage forecasting. Journal of Hydrology 328 704-716. doi: 10.1016/j.jhydrol.2006.01.021

Yu, X., Liong, S., (2007). Forecasting of hydrologic time series with ridge regression in feature space. Journal of Hydrology 332 290-302.

ftp://ftp.ncdc.noaa.gov/pub/data/cirs/climdiv/drought-readme.txt (accessed in May, 2014)

https://www.cisa.sc.edu/atlas/glossary.html (accessed in October, 2020)

https://cran.r-project.org/web/packages/nonlinearTseries/nonlinearTseries.pdf

https://cran.r-project.org/web/packages/fractal/index.html

https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn (accesed in October, 2020)