Scientific Research and Essays Vol. 6(21), pp. 4469-4477, 30 September, 2011
Available online at http://www.academicjournals.org/SRE
DOI: 10.5897/SRE10.1022
ISSN 1992-2248 ©2011 Academic Journals

Full Length Research Paper

Meteorological drought analysis using artificial neural networks

M. Erol Keskin[1]*, Özlem Terzi[2], E. Dilek Taylan[1] and Derya Küçükyaman[1]

[1]Faculty of Engineering-Architecture, Suleyman Demirel University, 32260, Isparta, Turkey.
[2]Faculty of Technical Education, Suleyman Demirel University, 32260, Isparta, Turkey.

Accepted 22 August, 2011

Droughts may be classified as meteorological, hydrological or agricultural. When meteorological drought appears in a region, agricultural and hydrological droughts follow. In this study, the standardized precipitation index (SPI) was applied for meteorological drought analysis at five stations located around the Lakes District, Turkey. Analyses were performed on 3-, 6-, 9- and 12-month-long data sets. The SPI drought classifications were modeled by artificial neural networks (ANN), which has the advantage that, in contrast to most of the time series modeling techniques, it does not require the model structure to be known *a priori*. Comparison of the observed values and the modeling results shows a better agreement with SPI-12 and ANN models. While the mean square error (MSE) values varied between 0.061 and 0.153 for training stage, they varied between 0.09 and 0.147 for testing stage of SPI-12 values. Also, the highest R2 values obtained as 0.930 for training stage and 0.923 for testing stage of Sütçüler station between SPI-12 and ANN models.

Key words: Drought, standardized precipitation index (SPI), artificial neural networks.

INTRODUCTION

Drought analysis is important in water resources management, planning and for long term economic and social planning. In Turkey, total useable water amount reaches 110 billion m^3 which is the sum of annual mean stream flow and ground water potential. Considering that the population of Turkey is 70 million, about 1570 m^3 of water is available per capita annually, which indicates that in future Turkey may encounter serious problems especially in drought periods (Keskin et al., 2009).

Türkeş (1999) examined regional and historical changes of precipitation and drought index series as climatic factors for 1930 to 1993. He indicated that extreme drought is seen in Southeast and Middle Anatolia regions due to climatic effects while in the Mediterranean and Aegean regions due to human effects. Pongracz et al. (1999) used as input El Niño Southern Oscillation (ENSO) for regional drought estimation and formed a fuzzy based model with modified Palmer Index estimation for the case study of a typical Great Plains state, Nebraska. Fowler and Kilsby (2002) investigated effect of weather condition on historical drought records by using Lamb weather tips, since 1881. Three main drought tips are determined as east drought, Pennine or west drought and other region's drought in Yorkshire. Hisdal and Tallaksen (2003) studied meteorological and hydrological drought characteristics by using monthly precipitation and runoff values for Denmark. Decrease of runoff is found infrequent in longer period than reduction in precipitation. Bonaccorso et al. (2003) performed drought analysis by using Standardized Precipitation Index (SPI) in 43 gauges over Sicily and selected periods that are determined when the index affected results sensitively. Sırdas and Şen (2003) defined connection between drought duration and amplitude by drawing an optimum line and formed drought maps for Turkey. Tsakiris and Vangelis (2004)

*Corresponding author. E-mail: merol@mmf.sdu.edu.tr.

Table 1. SPI drought categories (Mishra and Desai, 2006).

Value of SPI	Drought category
0 to –0.99	Mild drought
–1.00 to –1.49	Moderate drought
–1.50 to –1.99	Severe drought
≤ –2	Extreme drought

formed drought monitoring system using SPI with geographical information system in the island of Crete. Dinpashoh et al. (2004) separated Iran into seven regions. Six regions had homogeny characteristics and one had different pattern. Drought analysis was performed depending on fifty seven different climatic and geographic variables. Paulo et al. (2005) determined occurrence probability of each drought categories by using Markov chain. SPI was applied for sixty eight-year precipitation data for several sites of Alentejo, a drought prone region of southern Portugal. It was stated that drought monitoring with stochastic models and early warning system were feasible. Moreira et al. (2006) performed drought study using sixty eight-year normalized precipitation data in Alentejo region in south Portugal. Performances of log-linear models were found good for comparison of drought categories among different periods. Mishra and Desai (2006) compared linear stochastic models, recursive multi-step neural network and direct multi-step neural network for drought forecasting in the Kansabati River Basin, which lies in the Purulia district of West Bengal, India. They obtained that recursive multi-step approach is best suited for one-month ahead prediction. Mishra et al. (2007) developed a hybrid model, combining a linear stochastic model and a non-linear artificial neural network model for drought forecasting in the Kansabati River Basin in India. They said that hybrid model forecasted droughts with greater accuracy. Keskin et al. (2009) experimented appropriate artificial intelligence modelling techniques for Lakes District. They used adaptive neural based inference system (ANFIS) and fuzzy logic techniques for nine stations. They suggested ANFIS had given better results than fuzzy logic. Instructions of European Union meteorological drought analysis is one of the hydrological studies performed with respect to continuous monitoring of water resources potential, for short, mid and long range management planning, preventing or diminishing of negative effects.

The aim of this paper is to develop suitable ANN models by considering the feed-forward back-propagation learning algorithm in the estimation of SPI values. For this purpose, meteorological drought analysis is formed for the stations Isparta (Centrum), Eğirdir, Uluborlu, Yalvaç and Sütçüler in the Lakes District, Turkey. For drought analysis, it is determined whether precipitation values are fitting gamma distribution and SPI method is used. Drought is investigated in 3-, 6-, 9-, 12- month periods and drought processes of region are searched. The obtained drought categories given by McKee et al. (1993) are modelled by ANN.

Standardized precipitation index (SPI)

Standardized precipitation index (SPI) was developed by Mckee et al. (1993). It transforms precipitation value to only one numerical value for defining drought of different categories. Computation of index is complex since precipitation may not adapt normal distribution in 12 months or shorter periods. Each of the data sets is fitted to the Gamma function to define the relationship of probability to precipitation. Once the relationship of probability to precipitation is established from the historic records, the probability of any observed precipitation data point is calculated and used along with an estimate of the inverse normal to calculate the precipitation deviation for a normally distributed probability density with a mean of zero and standard deviation of unity (Mckee et al., 1993). SPI values show linearly increasing or decreasing tendency with precipitation deficiency. Drought intervals in selected period are represented in Table 1. In drought evaluation based on SPI values, period is defined as "drought" if index is negative. While the first month index is negative it is accepted as starting of drought if it is positive then it is accepted as ending of drought (Kömüşcü et al., 2002). Drought intensity based on SPI is classified according to categories given in Table1.

A monthly precipitation data set is prepared for a period of m months, ideally for a continuous period of at least 30 years. A set of averaging periods are selected to determine a set of time sequence of period i months where i is 3-, 6-, 12-, 24-, or 48- month. The data set is moving in the sense that each month a new value is determined from the previous i months. Each of the data sets is fitted to the Gamma function to define the relationship of probability to precipitation. Once the relationship of probability to precipitation is established from the historic records, the probability of any observed precipitation data point is calculated and used along with an estimate of the inverse normal to calculate the precipitation deviation for a normally distributed probability density with a mean of zero and standard deviation of unity (Mckee et al., 1993). Time intervals may be different according to water resources condition in the region. These time intervals are chosen according to subjective logic when the effect of decrease in rainfall can be felt on available water resources (Kömüşcü et al., 2002).

Artificial neural networks

Neural networks are promising new generation of information processing systems that demonstrate the ability to learn, recall, and generalize from training patterns or data. Artificial neural networks (ANN) are systems that are deliberately constructed to make use of some organizational principles resembling those of the human brain.

ANN has a large number of interconnected processing elements (nodes or units) that usually operate in parallel and are configured in regular architectures. The collective behavior of an ANN, like a human brain, demonstrates the ability to learn, recall, and generalize from training patterns or data. ANN is inspired by modeling networks of real (biological) neurons in the brain. Hence, the processing elements in ANN are also called artificial neurons, or simply neurons. A typical neuron has three parts: the cell body or soma, where the cell nucleus is located, the dendrites, and the axon. Dendrites are tree like networks of nerve fiber connected to the cell body. An axon is a single, long, cylindrical connection extending from the cell body and carrying impulses (signals) from the neuron. The end of an axon splits into strands or a fine arborization. Each strand terminates in a small bulblike organ called a synapse, where the neuron introduces its signal to the neighboring neurons. The receiving ends of these junctions on the neighboring neurons can be found both on the dendrites and on the cell bodies themselves. There are approximately 10^4 synapses per neuron in a human brain.

The signals reaching a synapse, and received by dendrites are electric impulses. Such signal transmission involves a complex chemical process in which specific transmitter substances are released from the sending side of the junction. This raises or lowers the electric potential inside the body of the receiving cell. The receiving cell fires if its electric potential reaches a threshold, and a pulse or action potential of fixed strength and duration is sent out through the axon to the axonal arborization to synaptic junctions to other neurons. After firing, a neuron has to wait for a period of time called the refractory period before it can fire again. Synapses are excitatory if they let passing impulses cause the firing of the receiving neuron, or inhibitory if they let passing impulses hinder the firing of the neuron.

Although simplicity models a biological neuron as a binary threshold unit, a McCulloch-Pitts (1943) neuron has substantial computing potential. It can perform the basic logic operations NOT, OR, and AND when weights and thresholds are selected accordingly. Since any multivariable combinational function can be implemented by these basic logic operations, a synchronous assembly of such neurons is capable of performing universal computations, much like an ordinary digital computer (Lin and Lee, 1996).

The back-propagation learning algorithm is one of the most important historical developments in neural networks. It has reawakened the scientific and engineering community to the modelling and processing of many quantitative phenomena using neural networks. This learning algorithm is applied to multi-layer feed-forward networks consisting of processing elements with continuous and differentiable activation functions. Such networks associated with the back-propagation learning algorithm are also called back-propagation networks. Given a training set of input–output pairs, the algorithm provides a procedure for changing the weights in a backpropagation network to classify the given input patterns correctly. The basis for this weight update algorithm is simply the gradient-descent method as used for simple perceptrons with differentiable neurons (Terzi and Keskin, 2010).

For a given input–output pair, the back-propagation algorithm performs two phases of data flow. First, the input pattern is propagated from the input layer to the output layer and, as a result of this forward flow of data, it produces an actual output. Then the error signals resulting from the difference between output pattern and an actual output are back-propagated from the output layer to the previous layers for them to update their weights (Lin and Lee, 1995).

MATERIALS AND METHODS

Study area and data

The Lakes district lies at south of the Mediterranean in Turkey (Figure 1). Its surface area is 8.933 km^2 and it is located between 30°20´ to 31°33´N and 37°18´ to 38 30´S. The mean altitude of district is 1050 m.

As a consequence of climatologic analysis of long period observations, both Mediterranean and terrestrial climate types are seen, while Mediterranean climate is observed in the south of this region (that is, Sütçüler), terrestrial climate appears in the north (Yalvaç). The mean annual precipitation is 551.8 mm/m^2. The important part of precipitation is in the winter and spring months (72.69%). Summer and autumn months are rather dry (29.31% of total precipitation).

In this study, monthly mean precipitation values of Isparta (Centrum), Eğirdir, Uluborlu, Yalvaç and Sütçüler and Lakes district, Turkey have been obtained from Turkish State Metrological Service for meteorological drought analysis. Record intervals of precipitation values are listed in Table 2.

Application

Instructions of European Union meteorological drought analysis is one of the hydrological studies. Meteorological drought is performed with respect to continuous

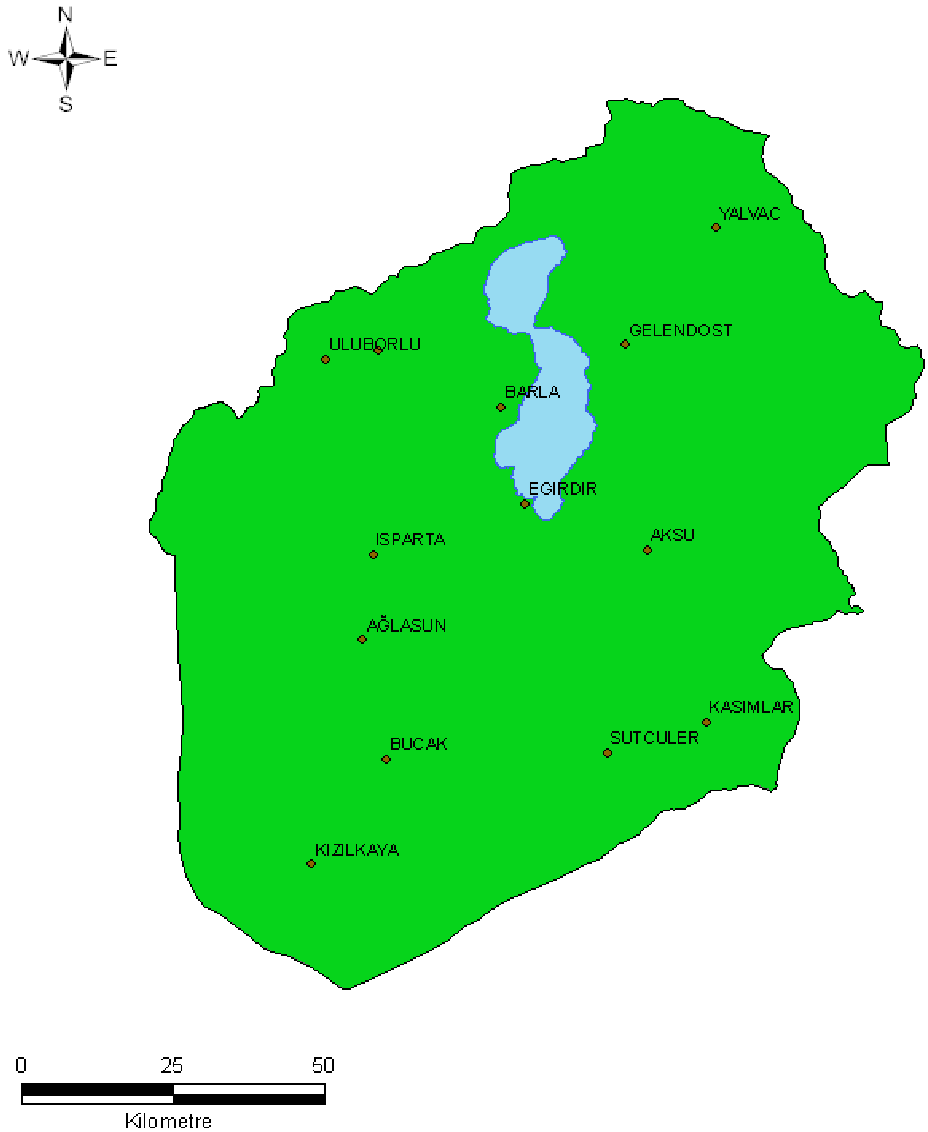

Figure 1. The selected sites in Lakes district.

Table 2. Precipitation stations for the Lakes District.

Station	Record interval
Isparta (Centrum)	1964–2006
Eğirdir	1964–2006
Uluborlu	1964–2006 (except 1968)
Yalvaç	1964–2004
Sütçüler	1975–1992

monitoring of water resources potential, for short, mid and long range management planning, preventing or diminishing of negative effects by conventional methods and artificial intelligence methods mostly used in recent years. Keskin et al. (2009) used ANFIS and fuzzy logic techniques for modeling drought. In this study, it is examined that ANN can be used to model drought.

It investigated which probability distribution fits into the

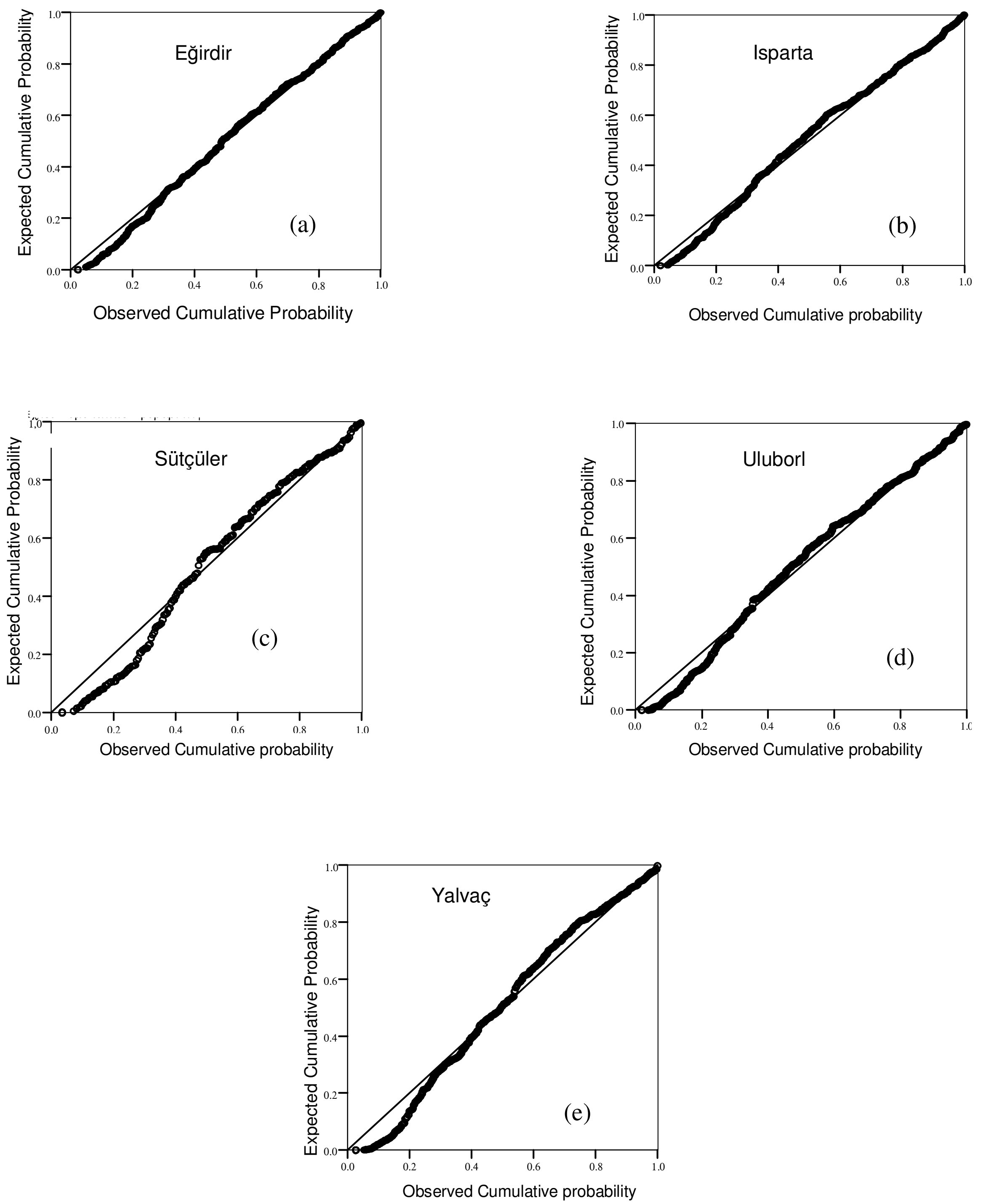

Figure 2. Gamma probability distributions of monthly mean precipitation for (a) Eğirdir, (b) Isparta, (c) Sütçüler, (d) Uluborlu and (e) Yalvaç stations.

precipitation data of station for identification of observed precipitation probabilities. Guttman (1999) examined the effect on the SPI values computed from different probability models for The United States and defined that the Pearson Type III distribution is the "best" universal model. However, Compo et al. (2007) said that seasonal mean precipitation is fitted significantly non-Gaussian in semi-arid regions of descent. In this study, gamma distribution is tested for stations in region and it is seen that gamma distribution fits to precipitation values in Figure 2.

SPI values of stations are computed for 3-, 6-, 9-, 12-

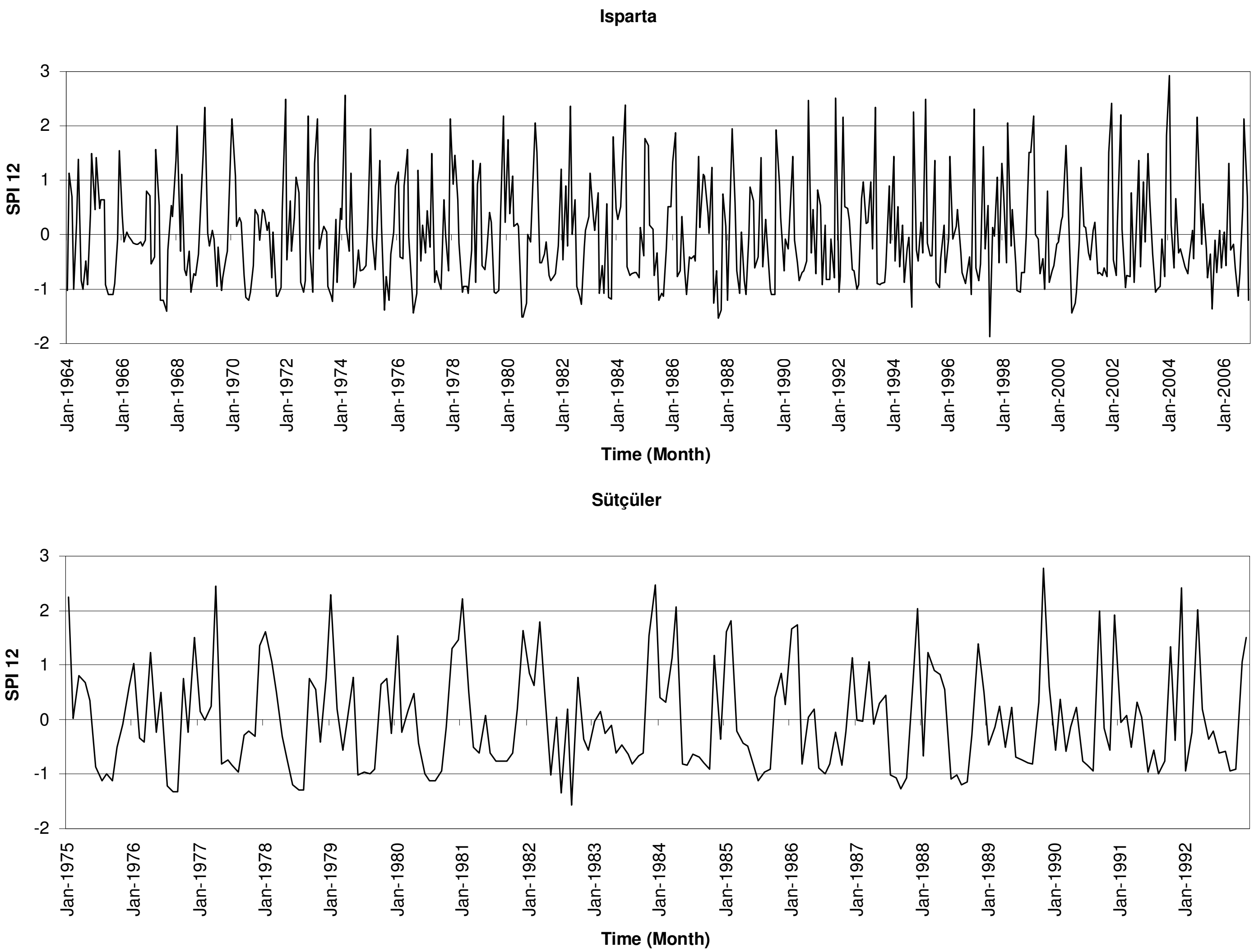

Figure 3. SPI time series for 12 months for (a) Isparta, (b) Sütçüler stations.

month periods and then categorized as in Table 1. Representative SPI time series for Isparta and Sütçüler stations are given for 12-month in Figure 3.

SPIs are estimated by ANN for 3-, 6-, 9-, 12-month periods by considering P_{t-1} and P_t as inputs in the model, where P_{t-1} is the precipitation for the previous period and P_t is the precipitation for the estimation period. For proper modelling, data is divided into two sections as training and testing set. 80% of total data is reserved for training and the remaining 20% for testing. The Neural Networks Toolbox of MATLAB R2008b was used for ANN models.

The adequacy of the ANN is evaluated by considering the coefficient of determination (R^2) and mean square error (MSE) based on the drought estimation errors as:

$$R^2 = 1 - \left(\left(\sum_{i=1}^{n} (D_{i(observed)} - D_{i(model)})^2 \right) \Big/ \left(\sum_{i=1}^{n} (D_{i(observed)} - D_{(mean)})^2 \right) \right) \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (D_{i(observed)} - D_{i(model)})^2 \quad (2)$$

Where n is the number of observed data, $D_{i(observed)}$ and $D_{i(model)}$ are observed SPI values and ANN results with respect to the monthly mean observed drought values D_{mean}. The resulting R^2 and MSE values are given in Tables 3 and 4.

In this study, ANN(i,j,k) indicates a feed forward back propagation network architecture with i, j and k neurons in input, hidden and output layers, respectively. Herein, i is 2; whereas k is 1. j is experimented for different hidden layer neuron in order to decide about the best ANN model alternative. Prior to execution of the model, standardization of the data, X_i, (i = 1, 2, …,n) is done according to the following expression such that all data values fall between 0 and 1.

Table 3. R^2 values of ANN models.

Stations	ANN models								
	Training stage								Testing stage:(12 Months)
	3 Months		6 Months		9 Months		12 Months		
Isparta (Centrum)	(2,5,1)	0.493	(2,3,1)	0.777	(2,3,1)	0.846	(2,3,1)	0.852	0.852
Eğirdir	(2,6,1)	0.469	(2,3,1)	0.837	(2,2,1)	0.850	(2,3,1)	0.882	0.882
Uluborlu	(2,5,1)	0.516	(2,6,1)	0.817	(2,2,1)	0.846	(2,3,1)	0.888	0.867
Yalvaç	(2,3,1)	0.464	(2,5,1)	0.765	(2,3,1)	0,829	(2,4,1)	0.830	0.907
Sütçüler	(2,5,1)	0.643	(2,3,1)	0.862	(2,3,1)	0.883	(2,3,1)	0.930	0.923

Table 4. MSE values of ANN models.

Stations	Training stage							
	3 Months		6 Months		9 Months		12 Months	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Isparta (Centrum)	1.670	0.276	0.186	0.193	0.136	0.306	0.133	0.136
Eğirdir	0.353	0.327	0.135	0.160	0.133	0.271	0.106	0.117
Uluborlu	0.322	0.028	0.152	0.184	0.137	0.207	0.099	0.126
Yalvaç	0.354	0.275	0.194	0.233	0.177	0.280	0.153	0.090
Sütçüler	0.304	1.185	0.113	0.162	0.101	0.171	0.061	0.147

$$x_i = (X_i - X_{\min})/(X_{\max} - X_{\min}) \tag{3}$$

Where x_i is the standardized value but X_{max} and X_{min} are the maximum and minimum measurement values. Such standardization procedure renders the data also into dimensionless form.

In Table 3, the highest R^2 values belong to 12-month period. ANN models which have different hidden layers have better R^2 and lower MSE than the other periods models for both training and testing, as shown in Table 1. The performance of the ANN models belong to 12-month period suggests that SPI values could be easily estimated from available data using the ANN approach for Isparta, Sütçüler, Eğirdir, Ulubaorlu and Yalvaç stations. Testing stage R^2 (Equation. 1) values are calculated for 12- month period. Comparison of the ANN and SPI values shows a better agreement between the ANN model estimations and SPI values. In order to test the performance of ANN model, representative Isparta and Sütçüler stations are plotted with results of SPI method in Figure 4 for the testing stage.

The straight lines in these graphs show 45° lines, which show the perfect correspondence between the observed and model values. However, the results from ANN model are given as a set of scatter points, which are randomly scattered around the perfect line. This point indicates the success in the overall model without any bias. On the other hand, in order to appreciate the SPI-12 observations and ANN model SPI-12 outputs the two time series are presented in Figure 5, which indicates obviously that they follow each other closely within less than 10% overall average relative error. Examining the time series given in Figure 5, for Isparta and Sütçüler stations, it is shown that ANN models have the ability to better predict drought categories (SPI <0) than wet categories.

Conclusions

In the present study, drought prediction for the Lakes District, Turkey is performed using artificial neural networks (ANN). First, time series of standardized precipitation index (SPI) are constituted for different periods based on monthly mean precipitation values for determination of drought categories for wet and drought periods. Then, different ANN models are developed to predict SPI categories for 3-, 6-, 9- and 12-month periods. The comparison shows that there is a better agreement between the results of the ANN models and SPI values for only 12-month periods as the highest R^2 and the lowest MSE values are obtained for 12 month periods for each station. The prediction of the drought also need more meteorological parameters than ANN models, and the drought could be easily estimated from available precipitation data using the ANN approach. It is seen that there is a good agreement between results of ANN model and SPI values. As a result, analyzing of the regional meteorological drought using artificial neural networks method developed in this study will be helpful in continuous monitoring of the water resources potential, planning of the short, medium and long-term management projects, preventing or minimizing adverse

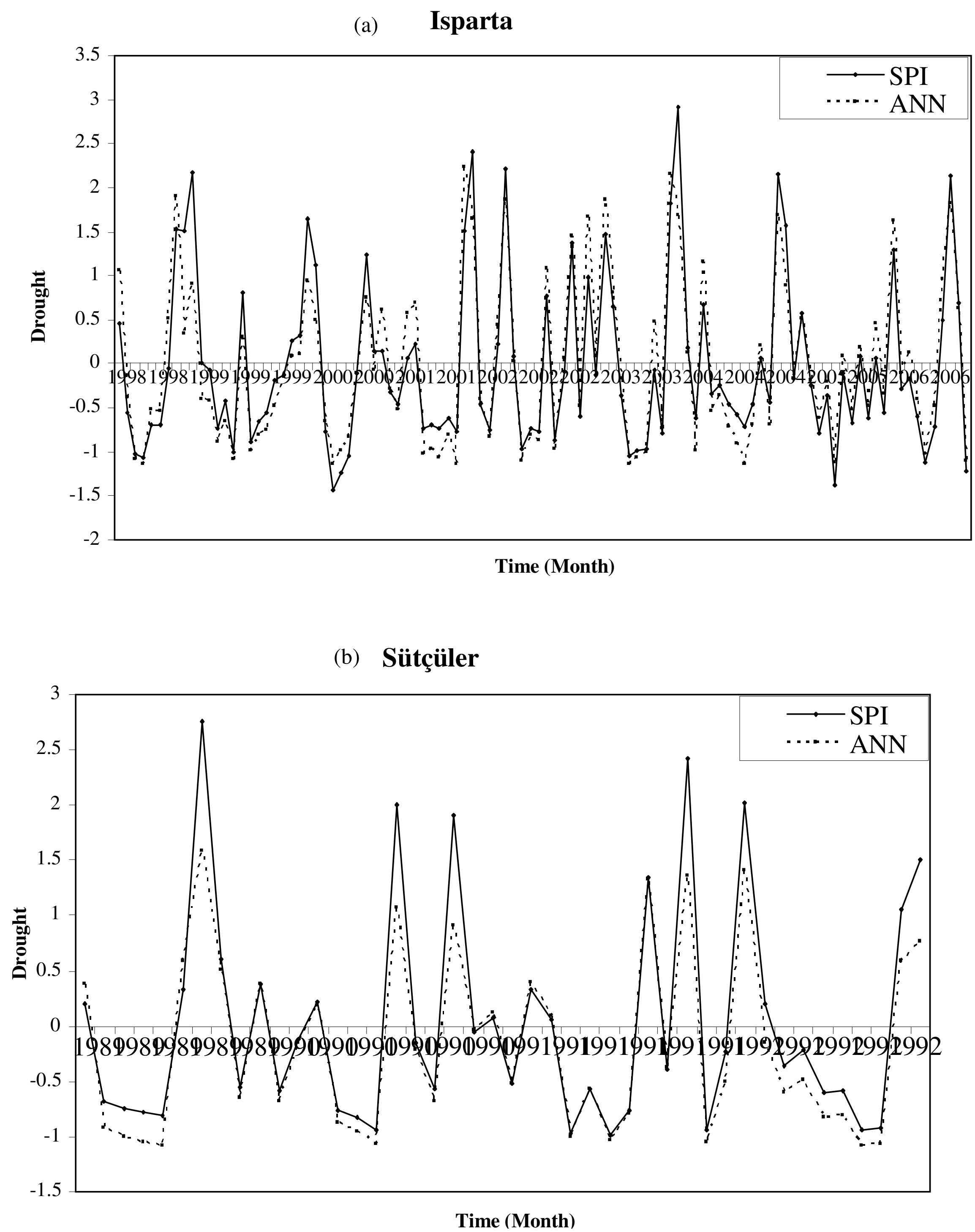

Figure 5. Plot of monthly drought prediction with ANN model for (a) Isparta, (b) Sütçüler stations.

impacts to be encountered at the hydrological studies. The developed ANN models are found suitable to predict effectively for peak point predictions. Drought prediction is a real problem for local administrations and water resources planners. Hence, this paper presents an applicable approximation to this planning stage by using ANN approach.

ACKNOWLEDGEMENTS

This research was supported by the Scientific and

Technological Research Council of Turkey, Project no. 106Y300.

REFERENCES

Bonaccorso B, Bordi I, Cancelliere A, Rossi G, Sutera A (2003). Spatial variability of drought: an analysis of the SPI in Sicily. Water Resour. Manage., 17: 273-296.

Compo GP, Sardeshmukh PD, Smith CA (2007). Precipitation probability distributions in drought-prone regions. 21st Conference on Hydrology, San Antonio, TX, pp. 14-18.

Dinpashoh Y, Fakheri-Fard A, Moghaddam M, Jahanbakhsh S, Mirnia M (2004). Selection of variables for the purpose of regionalization of Iran's precipitation climate using multivariate methods. J. Hydrol., 297: 109-123.

Fowler HJ, Kilsby JG (2002). A weather-type approach to analysing water resource drought in the Yorkshire Region from 1881 to 1998. J. Hydrol., 262: 177-192.

Guttman NB (1999). Accepting the Standardized Precipitation Index; a calculation algorithm. J. Am. Water Resour. Assoc., 35(2): 311-322.

Hisdal H, Tallaksen LM (2003). Estimation of regional meteorological and hydrological drought characteristics. J. Hydrol., 281: 230-247.

Keskin ME, Terzi Ö, Taylan ED, Küçükyaman D (2009). Meteorological drought analysis using data-driven models for Lakes District, Turkey. Hydrol. Sci. J. Sci. Hydrol., 54(6): 1114-1124.

Kömüşcü AÜ, Erkan A, Turgut E (2002). The geographic analysis of formation of drouhtness in Turkey used Standardized Precipitation Index Method. Edition of Turkish State Meteorological Service, Ankara, Turkey.

Lin CT, Lee CSG (1996). Neural fuzzy systems. Prentice Hall P.T.R., Upper Saddle Stream, NJ 07458.

Lin CT, Lee CSG (1995). Neural Fuzzy Systems. Prentice Hall: Englewood Cliffs, New Jersey, USA.

McCulloch WS, Pitts W (1943). A logical calculus of ideas immanent in nervous activity. Bull. Math. Biophys. **5**: 115-133.

McKee TB, Doesken NJ, Kleist J (1993). The relationship of drought frequency and duration to time scales. Eight Conferences on Applied Climatology, Anaheim, California.

Mishra AK, Desai VR, Singh VP (2007). Drought forecasting using a hybrid stochastic and neural network model. J. Hydrol. Eng., 12(6): 626-638.

Mishra AK, Desai VR (2006). Drought forecasting using feed-forward recursive neural network. Ecol. Modeling., 198: 127-138.

Moreira EE, Paulo AA, Pereira LS, Mexia JT (2006). Analysis of SPI drought class transitions using loglinear models. J. Hydrol., 331: 349-359.

Paulo AA, Ferreira E, Coelho C, Pereira LS (2005). Drought class transition analysis through markov and loglinear models, an approach to early warning. Agric. Water Manage., 77: 59-81.

Pongracz R, Bogardi I, Duckstein L (1999). Application of fuzzy rule-based modeling technique to regional drought. J. Hydrol., 224: 100-114.

Sırdaş S, Şen Z (2003). The modelling of meteorological drought and a case study of Turkey. ITU J. (in Turkish). 2(2): 95-100.

Terzi Ö, Keskin M E (2010). Comparison Of Artificial Neural Networks And Empirical Equations To Estimate Daily Pan Evaporation. Irrig. Drain., 59: 215-225.

Tsakiris G, Vangelis H (2004). Towards a drought watch system based on spatial SPI. Water Resour. Manage., 18: 1-12.

Türkeş M (1999). Vulnerability of Turkey to desertification with respect to precipitation and aridity conditions. Turkish J. Eng. Environ. Sci., 23: 363-380.