2012 International Conference on Environmental, Biomedical and Biotechnology
IPCBEE vol.41 (2012) © (2012) IACSIT Press, Singapore

Forecasting Drought in Tel River Basin using Feed-forward Recursive Neural Network

S. Sangita Mishra+ and R. Nagarajan

Centre of Studies in Resources Engineering, Indian Institute of Technology, Bombay, Powai, Mumbai-400076, India

Abstract. Drought is the most complex and least understood natural hazard with varying patterns in space, time and intensity. Because of the spatial and temporal variability and multiple impacts of drought, it is necessary to forecast drought conditions reasonably well in advance by either few months or seasons. This paper describes an approach to drought forecasting, using the recursive feed forward neural network and predicts quantitative values of the Standardized precipitation index (SPI). A number of different ANN models for SPI with the lead times of 1 to 12 months have been tested at several stations in the Tel river basin of Odisha, India. The SPI with 3 month and 12 months lead times have forecasted better than other lead times at all rainfall stations. The best models have R^2 values of 0.93-0.99 for a lead time of 3 months. The models including potential evapotranspiration as an input have performed better than the models considering rainfall and SPI alone. The structure of the model inputs (previous rain, drought index and potential evapotranspiration) does not vary with the lead time, which makes the models convenient for operational purposes like management of water resources.

Keywords: Drought forecasting, SPI, Neural networks, Tel river basin.

1. Introduction

Drought is a normal feature of climate and its occurrence appears inevitable. However, much confusion remains within the scientific and policy-making community about its characteristics. Research has shown that the lack of a precise and objective definition of drought in specific situations has been an obstacle in understanding drought. The success of drought preparedness and mitigation depends, to a large extent, upon timely information on drought onset, development in time and spatial extent. This information may be obtained through continuous drought monitoring, which is normally performed using drought indices. Drought indices are continuous functions of rainfall and/or other water-related variable(s) or temperature. They reflect emerging drought severity and can be used to trigger drought contingency plans, if those are designed and supported with appropriate institutional structure and responsibilities. Indices like Palmer Drought Severity index (PDSI), Deciles or Standardized Precipitation Index (SPI) are well known and frequently used in drought monitoring. However, monitoring, although useful for identifying early signs of droughts, detects only events that are already happening. The major challenge is to predict the future drought periods and their extremity- i.e. to enhance the early warning capability of drought monitoring systems through drought forecasting.

In recent decades, artificial neural networks (ANNs) have shown great ability in modeling and forecasting nonlinear and non-stationary time series in hydrology and water resource engineering due to their innate nonlinear property and flexibility for modeling. Some of the advantages of ANNs are (ASCE, 2000a). (1) They are able to recognize the relation between the input and output variables without explicit physical considerations. (2) They work well even when the training sets contain noise and measurement errors. (3)

+ Corresponding author. Tel.: + 91-9920271659 fax: +91-022-25767681
E-mail address: Sangita_csre@iitb.ac.in

They are able to adapt to solutions over time to compensate for changing circumstances and (4) They possess other inherent information processing characteristics and once trained are easy to use.

Drought forecasting using the Artificial Neural Networks has emerged as the most convenient method in recent years. Mishra and Desai (2006, 2007) demonstrated the application of ANN for drought forecasting using feed forward recursive multilayer approach, direct approach and the back propagation algorithm. The results obtained from the models show that recursive multi-step approach is best suited for 1-month ahead prediction and the direct multistep approach outperformed recursive multistep approach when a lead time of 4 months or more was considered. Morid et al., 2007 have made an attempt to forecast drought in Iran using past combinations of rainfall and climatic indices like Southern Oscillation Index (SOI) and North Atlantic Oscillation Index (NAO) and two drought indices viz Standardized Precipitation Index (SPI) and Effective Drought Index (EDI). The best models in both cases have been found to include, among the others, a corresponding drought index value from the same month of the previous year. Both best models have the R^2 values of 0.66-0.79 for a lead time of 6 months, but it is also shown that the EDI forecasts are superior to those of the SPI for all lead times and at all rainfall stations. Juan (2008) presented a non-linear multivariate model based on an artificial neural network multilayer perception that includes a random component. The developed model was applied to generate monthly streamflows, which are used to obtain synthetic annual droughts. Marj et al., 2011 proposed a model for agricultural drought forecasting based on normalized difference vegetation index (NDVI), using Southern Oscillation Index (SOI) and North Atlantic Oscillation (NAO), artificial neural network (ANN). This model was applied to Ahar-chay Basin in Azerbaijan Province, which is located in the northwest of Iran. The results show that in spring (May, June and July (MJJ)) synthetic NDVI can be predicted using ANN, with the input of SOI and NAO indices of the preceding (1 year) spring period.

In the present study, the utility of ANN approach for medium and long term forecasting of both the likelihood of drought events and their severity was examined In Tel river basin of Odisha, India. Neural network models were developed using the recursive multi-step neural network approach (RMSNN) with SPI as the drought indicator. Different combinations of past rainfall and evapo-transpiration and SPI values were used for forecasting drought at different lead times.

2. Database and Methodology

The physical area considered in this study is the Tel river basin, a major tributary of Mahanadi river in the western part of Odisha in eastern India. The region has an area of 2756 km^2. The study area experiences tropical wet and dry climate where the wet season (June–September) is much shorter and receives low precipitation from the south-west monsoon than the normal and the rest months of the year are generally dry because it does not receive any precipitation from north-east monsoon, which is the main reason of drought occurrences in this region. For this study, seven raingauge stations (Bhawanipatna, Dharamgarh, Junagarh, T. Rampur, Kalampur, Koksara and Jayapatna) were considered and the monthly rainfall data for the period 1965-2010 from all these stations was used. SPI time series for multiple time scales were derived for the average rainfall over the basin and these SPI values were used as drought index for drought forecasting. The SPI was calculated by taking the difference of the precipitation from the mean for a particular time scale, then dividing it by the standard deviation.

$$SPI = (Xik - Xi)/ói \tag{1}$$

where

ói = Standardized deviation for the ith station
Xik = Precipitation for the ith station and kth observation
Xi = Mean precipitation for the ith station

The SPI may be computed with different time steps (1 month, 3 months ….24 months). Positive and negative SPI values indicate wet and dry conditions respectively. The 'drought' part of the SPI range is split into 'near normal' (0.99>SPI>-0.99), 'moderately dry' (-1.0>SPI>-1.49), 'severely dry' (-1.5>SPI> -1.99) and 'extremely dry' (SPI<-2) conditions. A drought event starts when SPI value reaches -1.0 and ends when

SPI becomes positive again. The spatio-temporal variation of SPI at different time scales and in Tel river basin was already explained by the authors earlier (Sangita and Nagarajan, 2011). In this study, SPI was used for drought forecasting with different lead times of 1-12 months using the recursive feed forward neural network and the Multi-Layer Perception (MLP) theory which is most popular for hydrological research.

In the network structure, the neurons are arranged in interconnected groups called layers. Every ANN include: (1) input layer(s) – where the data are introduced to the network, (2) hidden layer(s) – where data are processed, and (3) output layer(s) – where the results for the given inputs are produced. A neuron computes its output response based on the weighted sum of all its inputs according to an activation function. In the present work sigmoid function, $F(x) = 1/(1+e^{-x})$ was used, which is the most popular choice. Data sets were normalized before the training begins using the following equation:

$$Xn = (X_0 - X_{min})/(X_{max}-X_{min}) \tag{2}$$

where X_n and X_0 represent the normalized and original data. X_{min} and X_{max} represent the minimum and maximum value among original data.

The coefficient of correlation (R^2), root mean square error (RMSE) and mean absolute error (MAE) for each combination of input and hidden neurons was calculated. The combination having maximum coefficient of correlation and minimum root mean square error was chosen as optimal network.

3. Results and Discussion

In the present study, the neural network models were developed to forecast drought using the feed forward training with standard back propagation algorithm. The available data were split into two parts, the data set from 1965 to 1999 was used to estimate the model parameters and the data from 2000 to 2010 was used to check the forecast accuracy. The forecast lead times varied from 1 to 12 months. Because it is the medium-range and the long-range forecasts that are critical for drought preparedness, we further discuss only the results of forecasting with a lead time of 3 months and longer. It is also virtually impossible to illustrate all results for all stations. We therefore primarily use the results from the worst drought affected station (Dharamgarh) as per the SPI analysis results to illustrate the main points. The statistics presented in table 1 indicate that for a lead time of 3 months, the models 1, 2, 5 and 6 have performed better than 3 and 4. It was also observed that the model 2 (Rainfall+potential evapotranspiration) performed better than model 1 with only rainfall as input.

Table 1. Results of SPI forecasting (three months lead time) at Dharamgarh station

Input model	Architecture	Training				Validation			
		R	R2	MAE	RMSE	R	R2	MAE	RMSE
1	6-15-1	0.97	0.91	0.19	0.24	0.93	0.81	0.17	0.65
2	7-15-1	0.97	0.91	0.18	0.23	0.97	0.94	0.15	0.17
3	5-13-1	0.89	0.71	0.32	0.40	0.73	0.22	0.34	0.50
4	4-10-1	0.82	0.38	0.41	0.52	0.87	0.21	0.33	0.42
5	5-13-1	0.95	0.87	0.22	0.28	0.87	0.68	0.22	0.32
6	6-11-1	0.98	0.94	0.15	0.20	0.99	0.96	0.11	0.13

*The three digits refer to numbers of neurons in input, hidden and output layers respectively. For example, an architechture 6-15-1 refers to six neurons in the input layer, fifteen neurons in the hidden layer and one neuron in the output layer.

Similarly, the model 6 (Rainfall+SPI+potential evapotranspiration) performed better than model5 (Rainfall+SPI). This emphasized the importance of potential evapotranspiration in drought forecasting. The forecasts significantly improved with model6, where the R^2 values for training and validation periods were 0.94 and 0.96 respectively. Models 3 and 4 with SPI as a single input and different combinations of other inputs had not resulted in accurate forecasts. In case of drought forecasting with 6 month and 9 month lead times, the model 4 and model 5 didn't work well at all like the other models. The R , R^2 values were low and the RMSE and MAE values were high in both the cases as compared to the 3 month and 12 month lead times.

The forecast with 12 month lead time indicated that all the six models worked and the best model architecture was 7-7-1 with RMSE and MAE of 0.1, 0.07 in training and 0.12, 0.1 in validation respectively. Figure 1 shows the variation of R, R^2, RMSE and MAE for forecasting of the SPI with lead times of 3, 6, 9 and 12 months during the training and validation period for all seven selected stations in the Tel river basin. The figures indicated that the variation of R^2, RMSE and MAE values were more in Dharamgarh and Bhawanipatna stations for drought forecasting at different lead times. Figure 2 displays the observed time series of the SPI values against the forecasted ones with the lead times of 3, 6, 9 and 12 months. In addition, the corresponding scatter plots are also presented. In all cases the significance level of R^2 is 1%. The results effectively illustrated the high accuracy of medium and long range forecasts of SPI at Dharamgarh station. The results at other stations were broadly similar. From the above maps, it was observed that SPI forecasts for 3 and 12 months were better than that for 6 and 9 months.

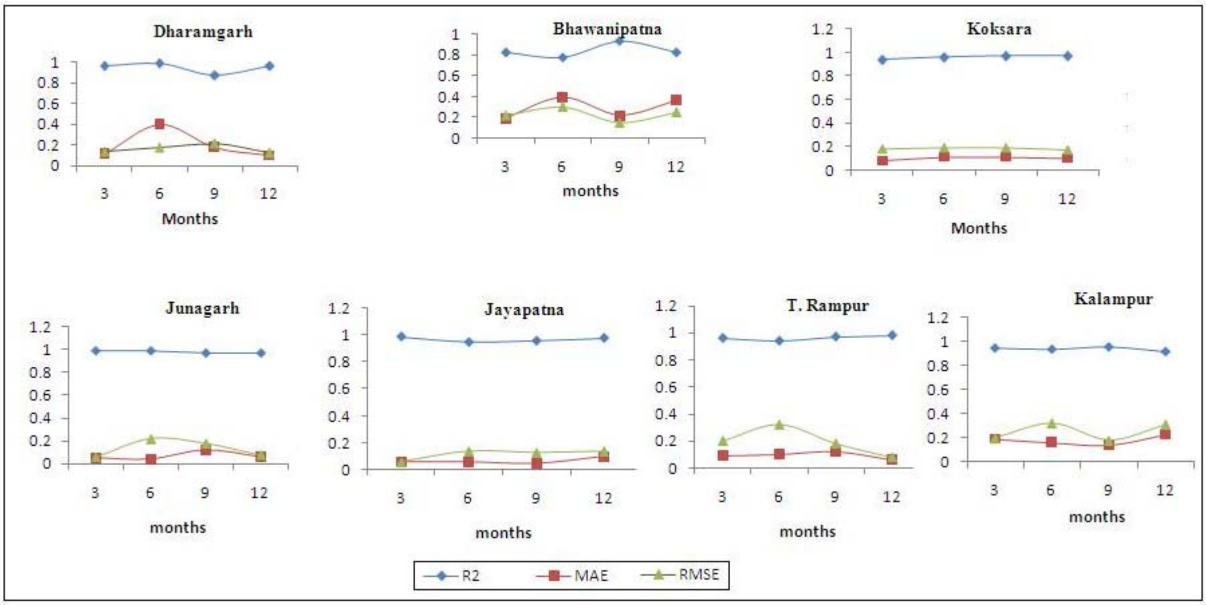

Fig. 1: Evaluation of the SPI forecasts using R^2, RMSE and MAE criteria for different future months during validation period at selected stations.

4. Conclusions

The paper described the procedure for drought forecasting using the ANN and its application in the Tel river basin of India. The Standardized Precipitation Index was used as the predictant, while different combinations of the past drought index, precipitation and potential evapotranspiration were used as predictors. 20 different models were tested for each drought index at seven rainfall stations in this region with lead times of 1 to 12 months. The SPI with 3 month and 12 months lead times forecasted better than other lead times at all rainfall stations. The best models had R^2 values of 0.93-0.99 for a lead time of 3 months. The models including potential evapotranspiration as an input performed better than the models considering rainfall and SPI alone. The structure of the model inputs (previous rain, drought index and potential evapotranspiration) does not vary with the lead time, which makes the models convenient for operational purposes like management of water resources.

5. References

[1] A.F. Marj, and A. M. J. Meijerink. Agricultural drought forecasting using satellite images, climate indices and artificial neural network. *International Journal of Remote sensing.* 2011, 32(24): 9707-9719.

[2] A.K. Mishra and V.R. Desai. Drought forecasting using feed-forward recursive neural network. *Ecological modeling.* 2006, 127-138

[3] A.K. Mishra, V.R. Desai, and V. P. Singh. Drought forecasting using a hybrid stochastic and neural network model. *Journal of Hydrologic Engineering.* 2007, 626-638.

[4] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, Artificial neural networks in

hydrology. I. Preliminary concepts. *Journal Hydro. Eng.* 2000a, 5(2): 115-137.

[5] Juan Camilo Ochoa-Rivera. Prospecting droughts with stochastic artificial neural networks. Journal of Hydrology. 2008, 352: 174-180.

[6] S. S. Mishra and R. Nagarajan. Spatio-temporal drought assessment in Tel river basin using Standardized Precipitation Index (SPI) and GIS. 2011, 2(1): 79-93.

[7] S. Morid, V. Smakhtin, and K. Bagherzadeh. Drought forecasting using artificial neural networks and time series of drought indices. *International journal of Climatology.* 2007, 27: 2013-2111.

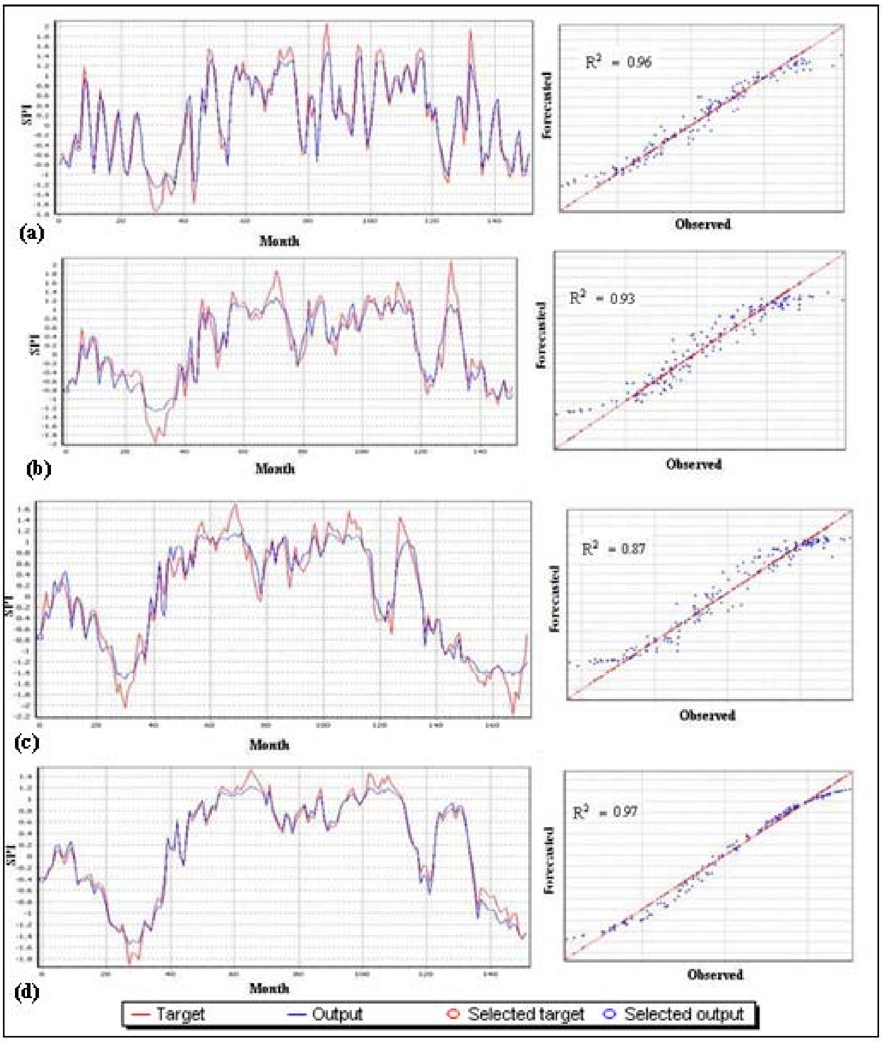

Fig. 2: Comparison of observed and forecasted SPI at the Dharamgarh station with lead times of 3 months (a), 6 months (b), 9 months (c) and 12 months (d) from 1965-2012.