Drought forecasting using feed-forward neural network

Salim DJERBOUAI, Doctoral student, s.djerbouai@yahoo.fr
Doudja SOUAG , university lecturer ,Souagd@yahoo.fr
Sarah HELLASSA, Doctoral student, Sarahzizou07@live.fr
LEGHYD Laboratory
University of Science and Technology Houari Boumediene, BP 32 Bab-Ezzouar Algiers, Algeria.

Abstract : Drought forecasting is a major component of a drought preparedness and mitigation plan. This study compares linear stochastic models (ARIMA/SARIMA) and Feed-forward neural network (FFNN) models for drought forecasting in the Algerois catchment in Algeria, using standardized precipitation index (SPI) as a drought quantifying parameter. The results obtained from two models are presented in this paper.

Keywords : Drought; ARIMA;SARIMA; FFNN; Forecasting.

Introduction

Drought is an insidious natural hazard that results from a deficiency of precipitation from expected or "normal" such that when it is extended over a season or longer period of time, the amount of precipitation is insufficient to meet the demands of human activities and the environment.

The Mediterranean is one of the regions where the impacts of drought events have shown an exponential increase during the last 20 years. Like other Mediterranean countries, Algeria has witnessed during the last twenty years an intense and persistent drought. This drought, which is characterized by an important rainfall deficit has affected the whole of Algeria particularly the north - west part. Similar droughts in amplitude and intensity have yet been noticed at the beginning of the century, between 1910 and 1940 (Medejerab and Henia, 2011). Forecasting future dry events in a region is very important for finding sustainable solutions to water management and risk assessment of drought occurrences.

Traditionally, stochastic models introduced by Box and jenkins, have been used for drought forecasting using time series of a drought quantifying parameter (Mishra and Desai, 2005 ; Durdu, 2010 ; Modarres, 2007 ; Han et *al*., 2010).

In recent decades, artificial neural networks (ANN) 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 applications of ANN models in drought forecasting include: (Morid et *al.*, 2007; Mishra and Desai, 2006).

The main objective of the present study is to develop time series of standardized precipitation index (SPI) and to compare neural networks models with linear stochastic models to forecast drought using SPI as drought index in the Algerois catchment in Algeria.

1. Study area and database

The case study area comprises of the western part of the The Algerois catchment(fig. 1), it has an area of 5225.3 km^2, a Mediterranean climate with an average annual rainfall between 600 and 800 mm in the coastal regions and between 500 and 1000 mm in the interior regions.

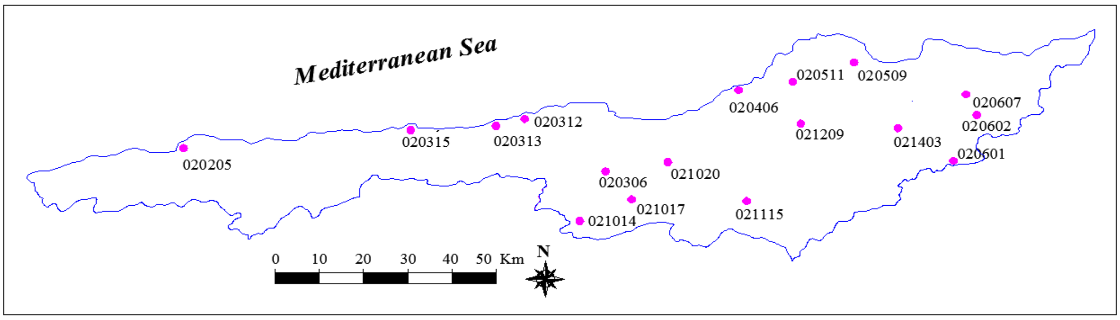

Fig. 1. Location of rain gauge stations used in the study

2. Methodology

2.1. Development of SPI series in Algerois catchment

McKee et al. (1993) developed the Standardized Precipitation Index (SPI) for the purpose of defining and monitoring drought. The SPI drought index was chosen to forecast drought in this study due to its simplicity, versatility, consistency and its independence from geographical position.

2.2. Artificial Neural Networks

Neural networks are a class of flexible nonlinear models that can discover patterns adaptively from the data. Theoretically, it has been shown that given an appropriate number of nonlinear processing units, neural networks can learn from experience and estimate any complex functional relationship with high accuracy.

Fig. 2 shows the proposed three-layer feed-forward model. The input nodes SPI_t, SPI_{t-1},.., SPI_{t-n} are the previous lagged observations while the output SPI_{t+L} provides the forecast for the future value where L is the lead time. In this study L was varied from 1 to 6 months lead time. For ANN model development , available data from 1936 to 2008 was divided into three subsets. First 60% of the samples are assigned to the training set, the next 20% to the validation set, and the remaining 20% of the data to the testing set.

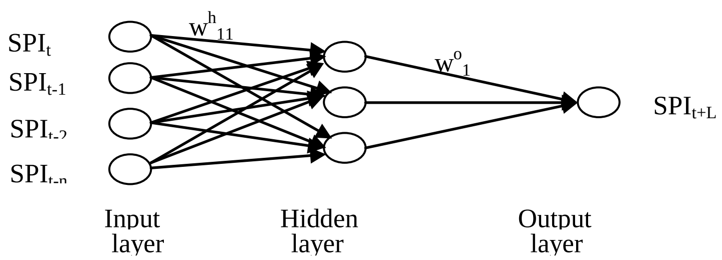

Fig. 2. The proposed Feed-forward Neural network

Optimal network architecture for different SPI series was found by trial and error approach. Performance criteria (NSE, RMSE, MAE), for each combination taking different number of input and hidden neurons, between observed and forecasted data were calculated. The combination having the best performance measures was chosen as optimal network.

2.3. Stochastic models

The general form of Stochastic models SARIMA(p,d,q)(P,D,Q)s model is given by:

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D x_t = \theta_q(B)\Theta_Q(B^s)a_t$$

The time series model development consists of three stages, i.e. identification, estimation and diagnostic check (Box and Jenkins, 1976).

For all developed SPIs time scales series, data set from 1936 to 1994 is used to estimate the model parameters and the data from 1995 to 2008 is used to check the forecast accuracy. Diagnostic checks are then applied to verify that models are adequate.

3. Results and discussions

The regional times series of SPI for 3,6 and 12 months time scale is calculated based on areal average precipitation estimated using Thiessen polygon method. SPI time series for 12 months time scale are shown Fig.3 .

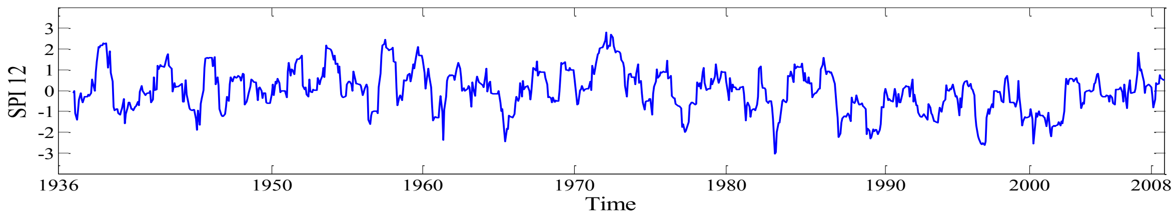

Fig. 3. SPI series for 12 time months scales based on average rainfall

3.1. Drought forecasting from selected models

Using the optimal structures of SARIMA and ANN models, the SPI time series for the three time scales were forecasted for 1 to 6-months lead times. The performance criteria obtained are shown in tables 1 and 2.

Table 1. The best SARIMA models forecasts

Lead- time (month)	NSE	RMSE	MAE	model Structure
(a) SPI 3 over different lead time				
1	0.4403	0.7318	0.5735	ARIMA(4,0,0)
2	0.0676	0.9392	0.7277	
3	-0.1936	1.065	0.7946	
4	-0.2173	1.0733	0.8128	

5	-0.2319	1.0813	0.813	
6	-0.2323	1.0847	0.8159	
(a) SPI 6 over different lead time				
1	0.6783	0.5343	0.4104	
2	0.4334	0.7091	0.5616	
3	0.2705	0.8092	0.6530	
4	0.0616	0.9205	0.7207	SARIMA$(5,1,2)(4,1,1)_6$
5	-0.0894	0.9936	0.7501	
6	-0.2200	1.0529	0.7869	
(a) SPI 12 over different lead time				
1	0.8576	0.329	0.2269	
2	0.6883	0.4877	0.3795	
3	0.5455	0.5906	0.4657	SARIMA $(0,1,0)(2,1,1)_{12}$
4	0.4178	0.6704	0.5301	
5	0.2952	0.7398	0.5847	
6	0.1971	0.7903	0.6243	

Table 2. The best ANN models forecasts

Lead- time (month)	NSE	RMSE	MAE	ANN Structure
(a) SPI 3 over different lead time				
1	0.4499	0.7255	0.579	4-10-1
2	0.1619	0.8905	0.6959	4-11-1
3	-0.0028	0.9754	0.7548	4-13-1
4	-0.008	0.9787	0.76	4-15-1
5	-0.0094	0.9818	0.7657	4-18-1
6	-0.0343	0.9906	0.7766	4-20-1
(b) SPI 6 over different lead				
1	0.7056	0.5111	0.4043	7-4-1
2	0.4706	0.6873	0.5479	7-5-1
3	0.2831	0.7976	0.6642	7-7-1
4	0.1241	0.8817	0.7048	7-8-1
5	0.0240	0.9307	0.7209	7-10-1
6	-0.0959	0.9888	0.7578	7-11-1
(c) SPI 12 over different lead time				
1	0.8608	0.3252	0.2422	13-4-1
2	0.7040	0.4743	0.3590	13-4-1
3	0.5805	0.5646	0.4112	13-5-1
4	0.4200	0.6639	0.5007	13-6-1
5	0.3106	0.7238	0.5749	13-6-1
6	0.2095	0.7751	0.6118	13-7-1

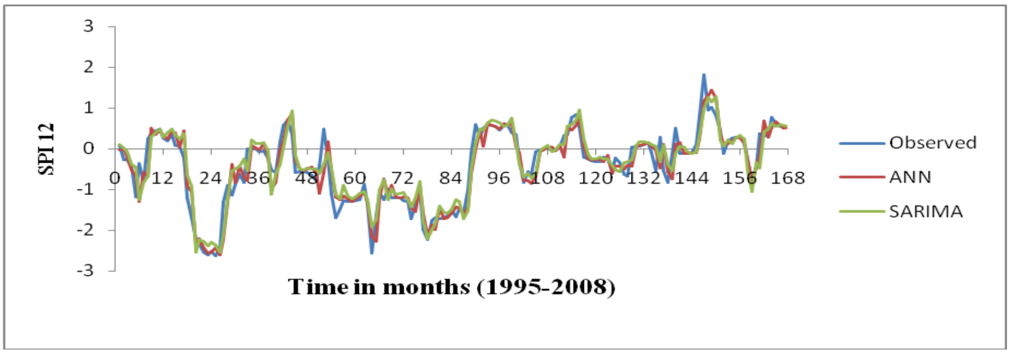

Fig. 4. A comparison between observed data and forecasts over 1-month lead time using SARIMA and ANN for SPI12

Analyzing the obtained results, it is possible to see that the SARIMA and ANN models present the same precision for all SPI time scales (example of SPI 12 shown in fig.4), and for all lead time forecasts. It may be seen that the forecast effectiveness is lower in SPI-3, the best result is obtained in the forecast of SPI-12. This is due to the high temporal variability in precipitation in SPI-3, while for the other scales; this variability is attenuated because more monthly data could be collected. Thus, while the SPI time scale increases, the SPI forecast is improved. It is also possible to see that, when lead time forecast is increased, a significant increase in the mean square error and a significant decrease in the NSE occur for all models, which is very common in hydrological forecasting models. For 1-month lead forecast, SPI3, SPI6 and SPI12 results were satisfactory. For 2 and 3 months lead times, only SPI12 presents consistent results. Over three months lead forecasts, neither SARIMA nor ANN are able to represent the drought tendency as it can be seen in tables 1 and 2.

Conclusions

The main objective of present study was to develop time series of standardized precipitation index (SPI) and to compare neural networks model with linear stochastic models to forecast drought using SPI as drought index in the Algerois catchment in Algeria. SPI was used as drought index, due to its several advantages as mentioned earlier.

The forecast was done for 1 to 6-month lead-time using stochastic and FFNN models using SPI time series for 3,6 and 12 months time scale. The results indicate that models developed to forecast drought found to give reasonably good results, and that stochastic and FFNN models present the same forecast precision.

It should be noted that unlike to stochastic models, forecast results from FFNN models can be improved.

References

Box, G.E.P., Jenkins, G.M., 1976. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

Durdu, O.F., 2010. Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey. Stoch. Environ. Res. Risk Assess. 24, 1145–1162.

Han, P., Wang, P.X., Zhang, S.Y., Zhu, D.H., 2010. Drought forecasting based on the remote sensing data using. ARIMA Models 51 (11–12), 1398–1403.

Mishra, A.K., Desai, V.R., 2005. Drought forecasting using stochastic models. J.Stoch. Environ. Res. Risk Assess. 19, 326–339.