Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 74(2) 2020 149–157

Drought Modelling and Forecasting using ARIMA and Neural Networks for Ballari District, Karnataka

Rahul Patil[1], B.S. Polisgowdar[1], Santosha Rathod[2], U. Satish kumar[1], G.V. Srinivasa Reddy[1], Vijaya Wali[1] and Satyanarayana Rao

[1]*University of Agricultural Sciences, Raichur, Karnataka*
[2]*ICAR- Indian Institute of Rice Research, Hyderabad*

Received 21 December 2019; Revised 09 May 2020; Accepted 14 July 2020

SUMMARY

In the present study, standardized precipitation index (SPI) series was analysed for different timescales of 1, 3, 6, 9, 12 and 24 months has been used to assess the vulnerability of meteorological drought in the Ballari district of Karnataka. SPI values showed that the occurrences of droughts in the study period varied from moderately to extremely condition. Suitable Autoregressive Integrated moving average (ARIMA) model and neural network Artificial neural network (ANN) models were developed to predict drought at different 1, 3, 6, 9 and 12 month timescale and lead time of up to 6 months ahead. The best model was selected based on minimum Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). An overview of the result shows that both ARIMA and ANN models have a better ability to forecast drought at different scales and also up to 2 month lead time. The evaluation of model performance was carried out using root mean square error (RMSE) and mean absolute error (MAE) Furthermore, the ANN model performed well for all stations compared to ARIMA models. ARIMA was observed to forecast well at higher timescale.

Keywords: ARIMA, ANN SPI, AIC, BIC, RMSE and MAE.

1. INTRODUCTION

Among the extreme events, droughts are the most widespread and slowly developing atmospheric hazards which remain for a long duration, affecting natural resources, environment, and people. Furthermore, it corresponds to the failure of spatial and temporal precipitation and water availability and therefore consequent impact on agriculture, ecosystem and socioeconomic activities of human beings. The global land surface in extreme drought is predicted to increase from 1-3 per cent for the present day to 30 per cent by the 2090s (IPCC, 2012). More intense droughts and increased precipitation variability lead to increased stressed condition, agriculture and economic activities. The frequency of severe and widespread multi-year droughts has increased in India during the recent decades due to the erratic summer monsoon and increase in air temperature and thereby creating huge damage to crops and society (Mishra *et al.*, 2014).

There are several methods that have been used in the past as drought assessment tools such as measurement of lack of rainfall, shortage of stream flow, reduced levels of water storage, and drought Indices (DIs). Of these, DIs were widely used for drought assessment. Out of few indicessuch as standardized precipitation index (SPI), China Z index (CZI), Deciles, Percent Normal (PN) and Rainfall anamoly index (RAI), the Standardized precipitation index (SPI) is considered as one of the most used drought indices around the globe.

Historically little attention has been given to drought forecasting aspect which is very important from the point of view of drought preparedness and early warning as mentioned earlier. In addition, in drought-prone regions, another drought event is likely to occur before the region fully recovers from the previous event. However, early indication of drought conditions could reduce future impacts and lessen the need for government intervention in the future. Therefore, the

Corresponding author: Rahul Patil
E-mail address: rahul1235110@gmail.com

utilization of drought forecasting tool like ARIMA and ANN can be used for drought preparedness.

2. MATERIALS AND METHODS

The monthly rainfall data of Ballari has been collected from Dept. of economics and statistics, Multi storey building, Bangalore. The study area is situated at an elevation of 580 msl (15°15' N latitude and 76°93' E longitude) this region falls in the southern part of karnataka, which is the 9th largest state in India, covering an area of 191976 sq.km, but has the 2nd largest arid zone after the state of Rajasthan in India. (Alam *et al.*, 2016).

3. STANDARDIZED PRECIPITATION INDEX

The Standardized Precipitation Index (SPI) is one of the most widely used drought index (Hayes *et al.,* 1999; Deo, 2011) developed by McKee *et al.* (1993). To calculate the SPI values, first the long-term precipitation record is fitted to a probability distribution. Sonmez *et al.* (2005) used the gamma distribution to rainfall data as it fits well to rainfall data, because of variety of reasons. The first advantage of gamma distribution is that it is bounded on the left at zero. Secondly, the gamma distribution is positively skewed. The current study also used the gamma distribution to fit the long-term rainfall record; gamma distribution is defined by its probability density function of Equation (1)

$$f(x,\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)} \quad \text{for } x, \alpha, \beta>0 \qquad (1)$$

where, α and β are the shape and scale parameters respectively;

x is the rainfall amount; and $\Gamma(\alpha)$ is the gamma function.

The maximum likelihood method was used to estimate the optimal values of α and β parameters using Equations (2) and (3) respectively.

$$\alpha = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right) \qquad (2)$$

$$\beta = \frac{\bar{X}}{\alpha} \qquad (3)$$

$$A = \ln(\bar{x}) - \frac{\sum_{i=1}^{n} \ln(x)}{n} \qquad (4)$$

$\bar{X}$ = Mean rainfall

ln = Natural log

n = total months

The resulting parameters were then used to derive the cumulative probability for non-zero rainfalls using Equation (1).

$$F(x,\alpha,\beta) = \int_0^x f(x,\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^x x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)} dx \qquad (5)$$

Which can be expressed by Equation (6)

$$F(x,\alpha,\beta) = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-t} dt \qquad (6)$$

where, $t = x/\beta$

Since the gamma function is undefined for $x = 0$ and the rainfall time series data may contain zero values, the cumulative probability of zero and non-zero values, $H(x)$ was calculated using Equation (7).

$$H(x) = q + (1-q)\,F(x;\alpha,\beta) \qquad (7)$$

Where, q is the probability of zero rainfall. If m is the number of zeros present in a rainfall time series, then q is estimated by m/n.

The cumulative probability was then transformed into a standardized normal distribution so that the SPI mean and variance were zero and one respectively with the help of equation 8 and 9. Following Mishra and Desai (2006), the current study employed the approximate transformations provided by Abramowitz and Stegun (1965) to transform the cumulative probability distribution into a standardized normal distribution, which are given in Equations (8) and (9):

$$SPI = -\left(K - \frac{c_0 + c_1K + c_2K^2}{1 + d_1 + d_2k^2 + d_3K^3}\right),$$

$$\text{When } K = \sqrt{\ln\left(\frac{1}{(H(x))^2}\right)} \qquad (8)$$

for $0 < H(x) \leq 0.5$

$$SPI = +\left(K - \frac{c_0 + c_1K + c_2K^2}{1 + d_1 + d_2k^2 + d_3K^3}\right),$$

$$\text{When } K = \sqrt{\ln\left(\frac{1}{1-(H(x))^2}\right)} \qquad (9)$$

for $0.5 < H(x) \leq 1$

Where, $C_o = 2.515517$, $C_1 = 0.802853$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$ and $d_3 = 0.001308$.

In this study, SPI was calculated on a different time scales. The SPI threshold ranges that are used to define drought conditions are presented in Table 1 (McKee *et al.*, 1993). For example a 3-month timescale (SPI3 February) SPI at the end of February compares the December–January–February precipitation total in that particular year with the December–February precipitation totals of all the years on record for that location.

Table 1. Drought classification based on SPI (McKee *et al.*, 1993)

Drought classes	SPI
≥ 2.0	Extremely wet (EW)
1.50 to 1.99	Severe wet (SW)
1.0 to 1.49	Moderately wet (MW)
0.99 to -0.99	Near normal (N)
-1.0 to -1.49	Moderate drought (MD)
-1.50 to -1.99	Severe drought (SD)
≤-2.0	Extreme drought (ED)

4. MAGNITUDE AND DURATION OF STRONGEST, LONGEST AND INTENSE DROUGHT

Yevjevich (1967) proposed the use of 'run theory' to define the drought characteristics, as shown in Fig. 1. A run is defined as a portion of time series of drought parameter X_t, in which all values are either below or above the selected truncation level of X_o; accordingly it is called either a negative run or a positive run. The truncation level has been defined as mean of the series over a long period of time. The knowledge of the components of a drought event is very important for the mathematical analysis of drought.

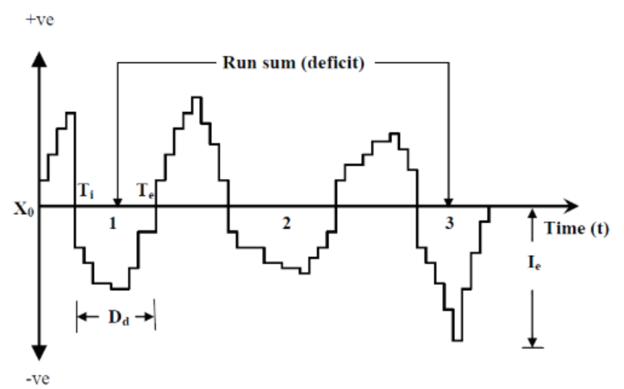
(a) Most intense drought (I_e)

(b) Drought initiation time (T_i).

(c) Drought termination time (T_e)

(d) Drought duration (D_d)

The magnitude and duration of drought events were detected using a runs theory, which will help to detect the strongest (S), longest (D) and intense (I) droughts. The strongest drought event is the one which is having a higher drought magnitude when its event values are summed up. The longest drought is the one in which duration of drought is high and the intense drought is one which is having a least index value compared to other drought events over the study period.



1. Drought with the highest severity.

2. Drought with the longest duration.

3. Drought with the highest intensity.

Fig. 1. Drought parameters using run theory for a given threshold level, X0

5. PREDICTION OF DROUGHT USING ARIMA AND ANN MODELS

Drought is a global phenomenon that occurs virtually in all landscapes. Due to the random nature of contributing factors, occurrence and severity of droughts can be treated as stochastic in nature. Early indication of possible drought can help to set out drought mitigation strategies and measures in advance. Therefore drought forecasting plays an important role in the planning and management of water resource systems. One of the basic deficiencies in mitigating the effects of drought is the inability to forecast drought conditions reasonably well in advance by either few months or seasons (Mishra and Desai, 2006).

5.1 ARIMA models

Autoregressive (AR) models can be effectively coupled with moving average (MA) models to produce a general and useful class of time series models named Auto Regressive Moving Average (ARMA) models. In an ARMA model the current value of the time series is expressed as a linear aggregate of '*p*' previous values and a weighted sum of *q* previous deviations (original value minus fitted value of previous data) plus a random component.

However, an ARIMA model can be used when the data are stationary. This class of models can be extended to non-stationary series by allowing differencing of data series. These models are called Auto Regressive Integrated Moving Average (ARIMA) models. Box and Jenkins (1976) provides a new generation of forecasting tools, known as the ARIMA methodology, which emphasizes on analyzing the stochastic properties of time series on their own rather than constructing single or simultaneous equation models. ARIMA models allow each variable to be stated by its own lagged values and stochastic error terms. The general non-seasonal ARIMA model is AR to order '*p*' and MA to order '*q*' and operates on d^{th} difference of the time series Z_t; thus a model of the ARIMA family is classified by three parameters (p, d, q) that can have zero or positive integral values (Mishra and Desai, 2006)

The general ARIMA model may be written as

$$\emptyset(B)\nabla^d Z_t = \theta(B)a_t \tag{10}$$

Where θ (B) are polynomials of order p and q, respectively. Non-seasonal AR operator of order p is written as

$$\emptyset(B) = \left(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots \emptyset_p B^p\right) \tag{11}$$

and non-seasonal MA operator of order q is written as

$$\theta(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots \theta_q B^q\right) \tag{12}$$

The development and application of ARIMA model majorly includes model identification, parameter estimation, diagnosis check and forecasting using selected models. The meteorological drought at different timescale were forecasted over different lead times (1, 2, 3, 4, 5 and 6 months) using optimal networks. For instance, a 1-month lead time prediction means that during January 2017, the prediction for February 2017 is computed. The RMSE and MAE were estimated between the observed and forecasted. The quantitative evaluation of the different model performance was carried out using root mean square error (RMSE) and mean absolute error (MAE) over different lead time for all SPI series

5.2 Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) is an information processing system that resembles the structure and operation of the brain (Maier *et al.*, 2010). The ANN modeling approach was developed in the 1940s by McCulloch and Pitts (1943) and gradually progressed with advances in calibration methodologies. Given sufficient data and complexity, ANN can be designed to model any relationship between a series of independent and dependent variables – inputs and outputs to the network respectively (Hornik *et al.*, 1990). One of the advantages of the ANN technique is that there is no need for the modeller to fully define the intermediate relationships (i.e., physical processes) between inputs and outputs (Morid *et al.,* 2002). This feature makes ANNs particularly suitable for the analysis of complex processes, like drought forecasting, where the relationships of a large number of input variables with the output need to be explored (Morid *et al.*, 2007). Because of this advantage, in recent years, the ANN modeling approach has been used in many research fields including drought forecasting (Morid *et al.*, 2007).

An ANN model was fitted to the SPI time series at different timescales. Although many types of neural network models have been proposed, the most popular one for time series forecasting is the feed forward network model. Fig. 2 shows a typical three-layer feed forward model used for forecasting. The input nodes are the previously lagged observations, while the output provides the forecast of a future value. Hidden nodes with appropriate nonlinear transfer functions are used to process the information received by the input nodes. In this study Direct multistep neural network approach was applied for forecasting the SPI series for 6 months ahead (Mishra *et al.* 2007).

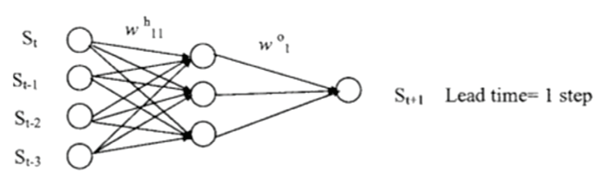


Fig. 2. Feed forward neural network for univariate time series

Direct multistep neural network approach (DMSNN): This model is based on multiple outputs, when several nodes are included in the output layer, and each output node represents one time step to be forecasted Fig. 3. This study includes six output nodes, indicating a one-to-six-month lead time. When an ANN is used for forecasting time series, input nodes are reconnected to a number of past observed values to identify the processes at future time steps. The activation function determines the relation between input and outputs of a node and a network. In the present work, a popularly used sigmoid function was employed.

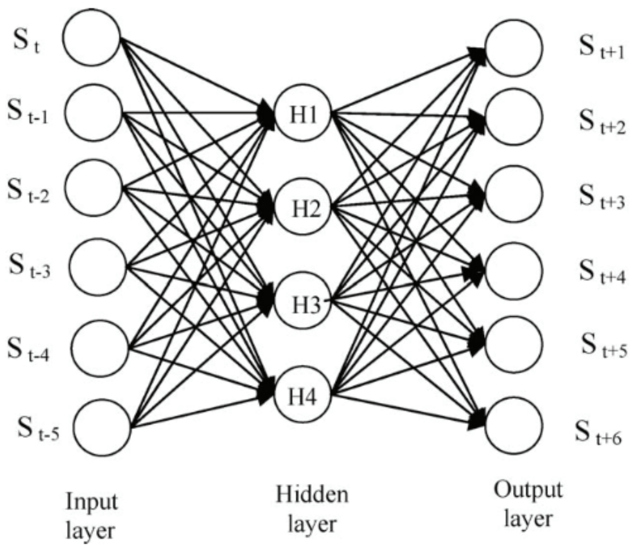


Fig. 3. Direct multi-step Neural Network approach

5.3 Comparison of ANN and ARIMA

The results of ANN and ARIMA models were compared. The quantitative evaluation of the different model performance was performed using root mean square error (RMSE) and mean absolute error (MAE) over different lead time and timescale.

6. RESULTS AND DISCUSSION

6.1 Historical drought events at Ballari station

The historical drought events were observed at different timescales for Ballari station during study period 1965-2017 and are presented in Table 1. The results show that the most severe drought captured by SPI during study period were 1985 with SPI values of -1.9, -1.95, -2.68, -1.99, -2.12 and -2.54 for SPI1_Aug, SPI3_Sept, SPI6_Oct, SPI9_Nov, SPI12_Dec and SPI24_Dec respectively. In addition, few more historical drought years observed were during 1971 (-1.62 SPI1_June, -1.81 SPI1_July, -1.33 SPI3_Sept, -1.15 SPI9_Nov, -1.27 SPI12_Dec and -1.1 SPI24_Dec), 1972 (-2.11 SPI1_Aug, -1.0 SPI9_Nov, -1.06 SPI12_Dec and -1.68 SPI24_Dec), 1976 (-1.2 SPI1_July, -1.68 SPI3_Sept, -2.97 SPI6_Oct, -2.88 SPI9_Nov, -3.03 SPI12_Dec), 2003 (-1.73 SPI1_June, -2.31 SPI3_Sept, -1.8 SPI6_Oct, -2 SPI9_Nov, -2.14 SPI12_Dec and -2.24 SPI24_Dec) and 2004 (-1.28 SPI1_Aug, -1.19 SPI3_Sept, -1.10 SPI6_Oct, -1.11 SPI12_Dec and -2.33 SPI24_Dec).

6.2 Magnitude and duration of longest, strongest and intense drought events at Ballari using SPI at different timescale

The drought magnitude and severity was analysed and estimated at different timescales of 1, 3, 6, 9, 12 and 24 months for Ballari station and are presented in a Table 2. The results show that longest drought months under SPI_1 was found to be during 1985 (Aug-Oct) and 2016 (Aug-Oct) with a duration of 3 months and severe event with a magnitude of -4.59 for 1985 (Aug-Oct) was recorded. Furthermore, the most intense drought month was observed during 2003 (Sept) with an intensity of -2.21. All along, for the other timescales of 3, 6, 9, 12 and 24 months the longest drought event recorded were during 1976 (May-Dec), 1976-77 (May-Mar), 1985-86 (Apr-May), 1984-86 (Oct-Aug) and 2003-05 (Jun-Jun) with a duration of 8, 11, 14, 23 and 25. On the other hand, the most intense months were 2003 (Jun), 2003 (Sep), 2003 (Jul), 1985 (Sep) and 1985 (Dec) with an intensity of -4.66, -3.72, -3.73, -3.33 and -2.54 for 3, 6, 9, 12 and 24 months respectively.

7. DROUGHT FORECASTING USING ARIMA AND ANN AT DIFFERENT LEAD TIME FOR BALLARI STATION

The development, forecasting, validation and comparison of the ARIMA and ANN model was carried out for Ballari station at different timescale

7.1 Development of ARIMA model

The autocorrelation test was carried out in order to know whether the data set is autocorrelated or not and the results are presented in Table 4. The view over the results found that for 3, 6, 9, 12 and 24 timescale the Chi-square and probability for Box test were 254, <0.001; 392.87, <0.001); 504.64, <0.001; 558.52, <0.001; and 632.28, <0.001. The results revels that the dataset at different timescale were observed to be Autocorrelated. Once the series is stationary and autocorrelated the next step is the development of model.

The Models selected for Ballari station at 3, 6, 9, 12 and 24 timescale based on the lower AIC and BIC values are presented in Table 5 are (0, 0, 2), (0, 0, 5), (3, 0, 4), (1, 0, 0) and (2, 0, 0) with an AIC and BIC values of (1387.56 & 1401.06), (1229.79 & 1256.42), (934.76 & 970.69), (714.76 & 723.85) and (78.40 & 91.88) for 3,

Table 2. Drought events according to several times steps for Ballari station

Years	SPI1 (Jun)	SPI1 (July)	SPI1 (Aug)	SPI3 (Sept)	SPI6 (Oct)	SPI9 (Nov)	SPI12 (Dec)	SPI24 (Dec)
1961	-1.81	-	-	-	-	-	-	-
1962	-	-	-	-	-1.22	-1.26	-	-
1963	-1.57	-1.55	-	-	-	-	-	-
1965	-	-1.81	-	-	-	-1.17	-	-
1967	-	-	-1.59	-	-	-	-	-
1968	-	-	-2.52	-	-	-	-	-
1971	**-1.62**	**-1.81**	**-**	**-1.33**	**-**	**-1.15**	**-1.27**	**-1.1**
1972	-	-	**-2.11**	**-**	**-**	**-1**	**-1.06**	**-1.68**
1973	-	-1.19	-	-	-	-	-	-
1975	-1.37	-	-	-	-	-	-	-
1976	-	**-1.2**	**-**	**-1.68**	**-2.97**	**-2.88**	**-3.03**	**-**
1977	-	-	-	-	-	-	-	-1.51
1979	-	-	-1.55	-	-	-	-	-
1980	-	-	-1.43	-1.11	-	-	-	-
1984	-1.81	-	-	-	-1	-1.27	-1.39	-
1985	-	-	-1.9	-1.95	-2.68	-1.99	-2.12	-2.54
1986	-	-	-	-	-	-	-	-1.45
1987	-	-1.74	-	-	-	-	-	-
1988	-1.32	-	-	-	-	-	-	-
1991	-	-1.33	-	-	-	-	-	-
1994	-1.06	-	-	-1.58	-	-1.23	-1.2	-
1995	-	-	-	-	-	-	-	-1.24
1996	-	-1.26	-	-	-	-	-	-
1997	-	-1.47	-	-2.8	-2.23	-2.02	-1.77	-
2002	-	-	-	-1.63	-	-	-1	-
2003	**-1.73**	**-**	**-**	**-2.31**	**-1.8**	**-2**	**-2.14**	**-2.24**
2004	**-**	**-**	**-1.28**	**-1.19**	**-1.1**	**-**	**-1.11**	**-2.33**
2006	-	-1.04	-1.9	-	-	-	-	-
2012	-1.39	-	-	-	-	-	-	-
2013	-	-	-1.24	-	-	-	-	-
2016	-	-	-1.68	-	-	-	-	-

Table 3. Magnitude and duration of longest, strongest and intense drought events at Ballari region using SPI at different timescale

Station	Longest		Strongest		Highest	
	Year	D	Year	S	Year	I
Ballari (SPI_1)	1985(Aug-Oct) and 2016 (Aug-Oct)	3	1985(Aug-Oct)	-4.59	2003(Sept)	-2.21
Ballari(SPI_3)	1976 (May-Dec)	8	1976(May-Dec)	-14.55	2003(Jun)	-4.66
Ballari(SPI_6)	1976-77(May-Mar)	11	1976-77 (May-Mar)	-25.25	2003(Sep)	-3.72
Bellary(SPI_9)	1985-86(Apr-May)	14	1985-86 (Apr-May)	-25.9	2003(Jul)	-3.73
Ballari(SPI_12)	1984-86(Oct-Aug)	23	1984-86 (Oct-Aug)	-39.58	1985(Sep)	-3.33
Ballari(SPI_24)	2003-05(Jun-Jun)	25	2003-05 (Jun-Jun)	-44.54	1985(Dec)	-2.54

6, 9, 12 and 24 respectively. Furthermore the maximum likelihood values for selected models were -690.79, -608.73, -459.38, -355.43 and -36.20 respectively. The parameters estimate for different models are presented in Table 6. In addition, the residuals were obtained by differencing original series with the fitted series and residuals were found to be white noise as presented in Table 7 with probability values of 0.63, 0.56, 0.96, 0.25 and 0.88 for 3, 6, 9, 12 and 24 months timescale respectively. The performance of the model during development is presented in Table 9 and the results show that the RMSE and MAE were observed to be low with a value of (0.68 & 0.52), (0.61 & 0.43), (0.48 & 0.39), (0.41 & 0.20) and (0.25 & 0.17) respectively.

7.2 Drought forecasting using ARIMA model

Soon after the development of the models forecasting was done at 1-6 month lead time period and the results are presented in Table 10. The results revel that the performance of the models was found to good at 1-2 leads time. Furthermore, for higher lead time the results were observed to more erroneous due to accumulation on error over increasing lead time.

7.3 Drought forecast using ANN model

The best fit models at different timescales were selected based on the least RMSE value in the training stage and the selected models are (7-4-6), (13-7-6), (19-10-6), (25-13-6) and (25-13-6) for 3, 6, 9, 12 and 24 months timescale. The performance of the model during development is presented in Table 9 and the results show that the RMSE and MAE were observed to be low with a value of (0.68 & 0.52), (0.61 & 0.43), (0.48 & 0.39), (0.41 & 0.20) and (0.25 & 0.17) respectively. Soon after model selection drought forecasting was carried out at 3, 6, 9, 12 and 24 timescale for Ballari station at 1-6 lead time and the results are presented in Table 10. The results reveal that the RMSE and MAE for SPI_3 at 1 month lead time were 0.48 and 0.37 respectively and for 6 month lead time the values were 1.69 and 1.32 respectively. A glance over the results clearly explains that as the lead time increases the RMSE and MAE value increases leading to an addition of errors. The forecasted values of SPI6 time series for one month lead time are presented in Table 11.

7.4 Comparison of forecast results

The forecasting ability of different ARIMA and ANN models was carried out at 3, 6, 9, 12 and 24 months timescale and the results are evaluated in terms of RMSE and MAE and are presented in Table 10. The results reveal that for 1 month ahead forecast the RMSE and MAE for different timescale were (0.91 & 0.73), (0.69 & 0.60), (0.65 & 0.50), (0.59 & 0.45) and (0.13 & 0.04) for ARIMA model and (0.48 & 0.37), (0.53 & 0.41), (0.50 & 0.390), (0.35 & 0.28) and (0.19 and 0.16) for 3, 6, 9, 12 and 24 respectively. An inspection

Table 4. Autocorrelation test for different time scales of Ballari station

Time scales	Chi-Square	Lag order	P-value
SPI_3	254.41	1	<0.001
SPI_6	392.87	1	<0.001
SPI_9	504.64	1	<0.001
SPI_12	558.52	1	<0.001
SPI_24	632.28	1	<0.001

Table 5. Log likelihood AIC and BIC values of ARIMA model for different time scales of Ballari station

Time Scales	Model	Log-Likelihood	AIC	BIC
SPI_3	(0, 0, 2)	-690.79	1387.56	1401.06
SPI_6	(0, 0, 5)	-608.73	1229.79	1256.42
SPI_9	(3, 0, 4)	-459.38	934.76	970.69
SPI_12	(1, 0, 0)	-355.43	714.87	723.85
SPI_24	(2, 0, 0)	-36.20	78.40	91.88

Table 6. Parameter estimation of ARIMA by maximum likelihood method for different time scales of Ballari station

Time scales	Model	Parameters	Estimate	S.E.	Z value	P-value
SPI_3	(0, 0, 2)	MA1	0.71	0.03	21.14	<0.001
		MA2	0.56	0.03	16.80	<0.001
SPI_6	(0, 0, 5)	MA1	0.74	0.03	20.61	<0.001
		MA2	0.63	0.04	15.08	<0.001
		MA3	0.55	0.04	12.12	<0.001
		MA4	0.49	0.04	11.54	<0.001
		MA5	0.43	0.03	11.67	<0.001
SPI_9	(3, 0, 4)	AR1	-0.74	0.07	-10.39	<0.001
		AR2	0.67	0.04	14.40	<0.001
		AR3	0.58	0.06	9.05	<0.001
		MA1	0.97	0.07	20.37	<0.001
		MA2	0.69	0.11	6.198	<0.001
		MA3	0.09	0.09	0.99	0.31
		MA4	0.141	0.05	2.76	0.005
SPI_12	(1, 0, 0)	AR1	0.90	0.016	56.22	<0.001
SPI_24	(2, 0, 0)	AR1	0.88	0.03	22.91	<0.001
		AR2	0.07	0.03	1.96	0.04

Table 7. Auto correlation check for residuals of ARIMA model at different timescales of Ballari station

Time scales	Chi-Square	Lag order	P-value
SPI_3	0.22	1	0.63
SPI_6	0.32	1	0.56
SPI_9	0.0019	1	0.96
SPI_12	1.27	1	0.25
SPI_24	0.02	1	0.88

Table 8. ANN models specifications for Ballari station at different time scales

Time Scales	Models	Parameters
SPI_3	7-4-6	37e
SPI_6	13-7-6	106
SPI_9	19-10-6	211
SPI_12	25-13-6	352
SPI_24	25-13-6	352

over the results reveals that the ANN performs better than ANN for all timescale except SPI_24. In SPI_24 the ARIMA performed better compared to ANN it may be due to the linearity in the dataset. ARIMA models perform better in the linear dataset compared to nonlinear one. Furthermore, for both model the forecasting ability is restricted to 2 lead times beyond which results found to be inaccurate.

8. CONCLUSION

Analysis of the computed SPI series shows that the study area has experienced the severe droughts during 1971, 1972, 1976, 1985, 2003 and 2004 followed by a moderate drought at regular intervals throughout the study area. This study summarises that the region is more prone to drought at a regular intervals. The most intense drought months were 2003 (Jun), 2003 (Sep), 2003

Table 9. Performance of different models in training data set different time scales of Ballari station

Criteria	SPI_3		SPI_6		SPI_9		SPI_12		SPI_24	
	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN
RMSE	0.68	0.62	0.61	0.42	0.48	0.22	0.41	0.11	0.25	0.09
MAE	0.52	0.47	0.43	0.3	0.39	0.15	0.2	0.08	0.17	0.06

Table 10. Comparison of different lead time forecast for ARIMA and ANN at different timescale interms of RMSE and MAE for Ballari station

Time scale	Model	Performance measures	Ballari					
			Lead time					
			1	2	3	4	5	6
SPI3	ARIMA (0, 0, 2)	RMSE	0.91	1.28	0.48	-	-	-
		MAE	0.73	1.07	0.34	-	-	-
	ANN (7-4-6)	RMSE	0.48	1.34	1.6	1.61	1.62	1.69
		MAE	0.37	1.12	1.31	1.32	1.32	1.32
SPI6	ARIMA (0, 0, 5)	RMSE	0.69	1.05	1.28	1.41	1.52	-
		MAE	0.6	0.88	1.09	1.18	1.27	-
	ANN (13-7-6)	RMSE	0.53	0.99	1.29	1.45	1.53	1.63
		MAE	0.41	0.82	1.11	1.25	1.37	1.45
SPI9	ARIMA (3, 0, 4)	RMSE	0.65	1.12	1.41	1.58	1.66	1.71
		MAE	0.5	0.88	1.15	1.35	1.44	1.49
	ANN (19-10-6)	RMSE	0.5	0.98	1.3	1.47	1.53	1.5
		MAE	0.39	0.79	1.11	1.2	1.36	1.36
SPI12	ARIMA (1, 0, 0)	RMSE	0.59	0.99	1.26	1.36	1.34	1.25
		MAE	0.45	0.77	0.97	1.08	1.08	1.08
	ANN (25-13-6)	RMSE	0.35	0.62	0.82	0.97	1.01	1.04
		MAE	0.28	0.47	0.67	0.77	0.8	0.84
SPI24	ARIMA (2, 0, 0)	RMSE	0.13	0.37	0.47	0.52	0.55	0.55
		MAE	0.04	0.3	0.38	0.43	0.44	0.44
	ANN (25-13-6)	RMSE	0.19	0.42	0.53	0.63	0.73	0.8
		MAE	0.16	0.36	0.46	0.54	0.61	0.69

(blank cell indicates forecast values were zero)

Table 11. The forecasted value of SPI6 series at one step ahead using ARIMA and ANN for Ballari region

Sl. No.	observed	ARIMA	ANN
1	0.57	0.36	0.55
2	0.74	0.64	0.80
3	0.13	0.46	0.18
4	-0.28	0.16	-0.06
5	0.39	-0.21	-0.33
6	1.08	0.21	0.38
7	1.82	0.74	0.61
8	1.12	1.40	1.51
9	0.27	1.11	1.02
10	-0.27	0.51	0.75
11	-0.95	-0.37	-0.64
12	-1.84	-1.06	-1.33
13	-3.51	-1.92	-2.19
14	-3.03	-2.78	-2.22
15	-1.59	-2.40	-1.77
16	-0.3	-1.16	-0.47
17	-0.64	-0.17	-0.46
18	0.33	-0.25	0.11
19	0.55	0.86	0.61
20	1	0.61	0.89
21	1.09	0.57	0.87
22	1.64	0.56	1.19
23	1.78	1.48	1.30
24	1.67	1.28	0.91

(Jul), 1985 (Sep) and 1985 (Dec) with an intensity of -4.66, -3.72, -3.73, -3.33 and -2.54 for 3, 6, 9, 12 and 24 months respectively. An overview of the result shows that both ARIMA and ANN models have a better ability to forecast drought at different scales and also up to 2 month lead time. Furthermore, the ANN model performed well for all stations compared to ARIMA models. ARIMA was observed to forecast well at higher timescale

REFERENCES

Abramowitz, M. and Stegun, I. A. (1965). Handbook of mathematical functions with formulas, graphs and mathematical table, Courier Dover Publications, New York, USA.

Alam, N. M., Susheel, K. S., Jana, C., Raizada, A., Mandal, D., Kaushal, R., Sharma, N. K., Mishra, P. K. and Sharma, G. C. (2016). Forecasting meteorological drought for a typical drought affected area in India using stochastic models. *Journal of the Indian society of agricultural statistics*, **70(1)**, 71-81

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis Forecasting and Control*. Perntice Hall, Englewood Cliffs, NJ, USA.

Deo, R.C. (2011). On meteorological droughts in tropical Pacific Islands: time-series analysis of observed rainfall using Fiji as a case study. *Meteorological Applications*, **18(2)**, 171-180.

Hayes, M. J, Svoboda, M. D, Wilhite, D. A. and Vanyarkho, O. V. (1999). Monitoring the 1996 drought using the standardized precipitation index. *Bull. Am. Meteorol. Soc*., **80(3)**, 429-438.

Hornik, K., Stinchcombe, M. and White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feed forward networks. *Neural Networks* **3(1)**, 551-560.

IPCC (2012) Managing the Risks of Extreme Events and Disasters to Advance Climate ChangeAdaptation.*A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change,* Cambridge University Press, Cambridge, UK and New York, NY, USA, pp.1-19.

Maier, H. R., Jain, A., Dandy, G. C. and Sudheer, K. P. (2010) Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software,* **25(8)**, 891-909.

McCulloch, W. S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biology. **5(4)**, 115-133.

McKee, T. B., Doesken, N. J. and Kleist, J. (1993). The relationship of drought frequency and duration to time scales, in Proceedings of the 8th Conference on Applied Climatology, Anaheim, California, USA, 179-184.

Mishra, A. K. and Desai, V. R. (2006). Drought forecasting using feed-forward recursive neural network, Ecological Modelling, **198(1-2)**, 127-138.

Mishra, A., Desai, V. R. and Singh, V. P. (2007) Drought forecasting using a hybrid stochastic and neural network model. *J. Hydrol. Engg.,* **12(6)**, 326-368.

Mishra, V., Shah, R. and Thrasher, B. (2014). Soil moisture droughts under the retrospective and projected climate in India. *Journal of Hydrometeorology*, 141006071153004. https://doi.org/10.1175/JHM-D-13-0177.

Morid, S., Gosain, A.K. and Keshari, A.K. (2002). Solar radiation estimation using temperature-based, stochastic and artificial neural networks approaches. *Nordic Hydrology* **33(4)**, 291-304.

Morid, S., Smakhtin, V. and Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Applied Climatology* **27(15)**, 2103-2111.

Sonmez, F.K., Komuscu, A.U., Erkan, A. and Turgu, E. (2005). An analysis of spatial and temporal dimension of drought vulnerability in Turkey using the Standardized Precipitation Index, *Natural Hazards* **35(2)**, 243-264.

Yevjevich, V.(1967) Objective approach to definitions and investigations of continental hydrologic droughts. *Hydrology Paper 23*. Colorado State University, Fort Collins, Colorado, USA.