ORIGINAL ARTICLE

Comparison of neural network configurations in the long-range forecast of southwest monsoon rainfall over India

Snehashish Chakraverty · Pallavi Gupta

Abstract The accurate long-range forecast of southwest rainfall can have manifold benefits for the country, from disaster mitigation and town planning to crop planning and power generation. In this paper, the rainfall has been modeled using artificial neural network (ANN) with different network configurations. Performance of these networks are compared with some results found in the literature. The networks have also been tested for the data outside the range of the trained data and compared with known results. The present network is found to be better in term of predictions than the previous results by others. Southwest monsoon rainfall over India for 6 years in advance has been predicted.

1 Introduction

The prediction of monsoon rainfall has been one of the major concentration in the world of sciences. Long-range forecast of rainfall certainly requires continuous effort and long planning using different approaches. Most of these approaches use either a dynamical prediction model or statistical models. A dynamical model being deterministic, does not require information about a particular situation beyond initial and boundary conditions. However, in spite of considerable

progress made in dynamical modeling and numerical weather prediction of precipitational patterns by dynamical method is still much below its desired level of accuracy.

In statistical method, neither one has a dynamical relation between the cause and effect of the system nor does one possess a definite notion about the relative roles of various processes that govern a phenomenon. However, performance of statistical prediction is often poor, especially when the irregular behavior of the observed data is a result of low dimensional [1]. However, statistical prediction is an active area of research, and use of non-linear techniques promises new developments. There are lot of studies for long-range prediction of monsoon which are aimed at finding suitable predictor parameters. Thus, these use limited data sets and treat monsoonal rainfall time series essentially as a stochastic process, the resulting forecast equations are regression relations with one or more predictor. The year 2002 turned out to be an all-India severe drought year, and the entire monsoon season had many intriguing features [2], specially the fact that in July the country received only half of the month's normal rainfall. This situation could not be foreseen by any operational statistical or dynamical model [2].

In recent years, another modeling technique, known as artificial neural networks (ANN), has become popular as an alternative technique for modeling and prediction of complicated time series of weather and climate variables. In this paper ANN with different architectures has been used to model the said problem and various results, comparisons and predictions as mentioned are discussed. The powerfulness and reliability of the present model may also be proved

S. Chakraverty (✉) · P. Gupta
B.P.P.P. Division, Central Building Research Institute, Roorkee 247 667, India
e-mail: sne_chak@yahoo.com

from the comparisons. Moreover, proposed methodology successfully predicted the drought in the year 2002, which was not predicted by other investigators.

2 Neural network and design of the model

A neural network is a parallel, distributed information processing structure consists of processing elements called neurons, which are interconnected and unidirectional signal channels called connections. Each processing element branches into as many output connection as desired and carry signals known as neuron output signal. The neuron output signal can be of any mathematical type desired. In other words, it is a mathematical representation of a biological neural network. In term of implementation, it is basically a coupled input–output map constructed through an iterative procedure. Recently, Goswami and Srividya [3] and Goswami and Kumar [4] used composite-neuron (CN) network for predicting rainfall patterns. They used a generalised structure viz. composite network and compared the performance of a CN with that of conventional NN. In the present study, the network has been trained using error back propagation algorithm [7] taking different number of hidden layers. Here, four network configurations designated as $N1$, $N2$, $N3$ and $N4$ with different parameters of the architectures have been considered. The descriptions of these architectures are given in Table 1. After exhaustive simulation the authors have concluded the best architecture which give comparatively better training and prediction against earlier predicted results by other researchers.

Before starting the training of the network, we must recognize the need for a measure of how close the network has come to an established desired value. This measure is the distance between the actual and the desired response, which serves as an error measure or error tolerance and it is used to correct network parameters externally. Since, supervised training is being considered here, the desired value is known for the given training set. For the error back propagation training algorithm,

an error measure is known as the root mean square error (RMSE). Any continuously error function can be used, but the choice of another error function does add additional complexity and should be approached with a certain amount of caution. The RMSE is defined as follows:

$$E_p = \frac{1}{2} \sum_{j=1}^{N} (d_{pj} - o_{pj})^2$$

where E_p is the error for the pth training vector, d_{pj} is the desired value for the jth output neuron (i.e. the training set value) and o_{pj} is the calculated output of the jth output neuron.

3 Data set

The data set used in the four configurations for training and prediction is of 38 years (1958–1995) taken from Rajeevan et al. [5], which are the previous years' rainfall. This data set is used to predict the rainfall for 6 years in advance. The five categories of rainfall is considered here as defined in Rajeevan et al. [5], which is given in Table 2 for the sake of completeness.

4 Training and testing

Training is the way a neural network learns. Training may be supervised or unsupervised. In this study supervised training has been used, which provides the network with the desired response. In the training process, firstly the input data is presented to the network, and then the network modifies the weights of the neurons and adjusts them to achieve a prediction of the next point in the input data with desired accuracy. When this procedure is carried out with a large sample of the input data (comprising the training set), the neural network 'learns' the relationship between input and output data and then the trained network can be used to make a prediction for a point immediately outside the training set. This process of predicting a point outside the training set is called testing.

Table 1 Description of the four network configurations

Configuration	Inter layer transfer function	RMSE	Network architecture	
			Number of hidden Layers	Number of hidden Nodes
$N1$	Sigmoid	0.005	2	20
$N2$	Sigmoid	0.005	3	20
$N3$	Sigmoid	0.0005	2	20
$N4$	Sigmoid	0.0005	3	20

Table 2 Category of monsoon rainfall for forecast [5]

Category	Rainfall range (%)
Drought	<90
Below normal	90–97
Near normal	98–102
Above normal	103–110
Excess	>110

Table 3 Comparison of performance of $N1$, $N2$, $N3$ and $N4$ network models with Rajeevan et al. [5] (16 parameter model)

Year	Actual	N1 (%)	N2 (%)	N3 (%)	N4 (%)	Rajeevan et al. [5]
1958	110	110.00	110.02	110.00	109.98	110
1959	114	113.99	113.99	114.00	113.99	117
1960	101	101.00	101.01	101.00	100.98	104
1961	122	121.86	121.80	121.94	121.88	120
1962	97	97.01	96.92	97.00	97.01	97
1963	98	97.99	97.99	98.00	97.98	97
1964	110	110.0	109.97	110.00	109.98	111
1965	82	81.99	82.03	82.00	81.97	82
1966	87	86.99	86.80	87.00	86.92	90
1967	100	99.99	100.02	99.99	99.93	102
1968	90	90.06	90.30	90.00	90.02	85
1969	100	99.93	100.11	99.99	99.98	98
1970	112	112.03	111.89	112.00	111.96	110
1971	104	103.99	103.77	104.00	103.94	99
1972	76	76.35	76.19	76.09	76.16	76
1973	108	107.98	107.87	107.99	107.97	108
1974	88	88.00	88.06	88.00	88.00	92
1975	115	114.99	115.05	115.00	114.99	118
1976	103	103.05	103.02	103.00	102.99	102
1977	104	103.99	103.97	104.00	104.04	102
1978	109	108.97	109.00	108.99	108.90	107
1979	81	81.00	81.01	81.00	81.016	82
1980	104	104.01	104.04	104.00	104.05	108
1981	100	99.99	99.97	100.00	100.03	104
1982	86	86.02	86.03	85.99	85.99	91
1983	113	112.84	113.04	112.95	112.96	113
1984	96	96.51	96.28	95.93	95.92	100
1985	93	91.97	91.76	93.05	93.09	99
1986	87	87.69	88.11	87.00	87.03	88
1987	81	80.99	80.89	80.94	81.00	79
1988	119	118.81	118.89	119.02	119.00	113
1989	101	101.17	101.25	101.05	101.01	102
1990	106	105.32	105.03	106.09	105.95	101
1991	91	92.82	93.09	91.20	91.01	94
1992	93	90.58	90.87	92.91	92.97	92
1993	100	101.28	100.73	99.99	100.04	103
1994	110	109.34	109.73	109.94	109.98	92
1995	100	100.18	100.07	100.22	99.80	97
r		0.999	0.999	0.999	0.999	0.928

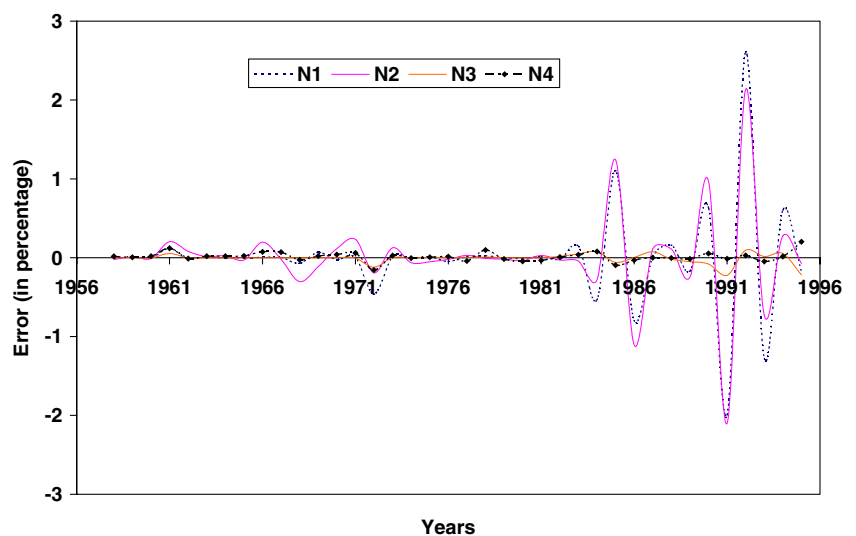**Fig. 1** Comparison of performance in learning of the four configurations $N1$–$N4$

Table 4 Comparison of the errors for the four (N1, N2, N3, N4) network configurations with Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4] and Report [6]

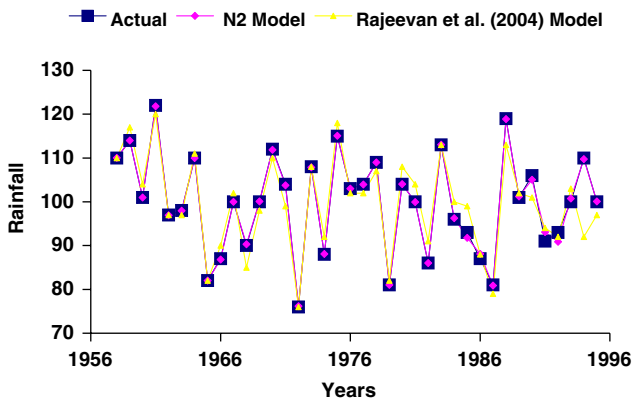| Year	Actual (%)	N1	N2	N3	N4	Rajeevan et al. [5]	Goswami and Srividya [3], Goswami and Kumar [4] and Report [6]
1996	103	4.77	4.74	5.30	7.21	3	8
1997	102	13.01	5.53	10.50	15.16	–2	4
1998	105	–6.19	–6.00	–6.73	–10.07	–6	–2
1999	96	0.22	0.45	0.71	5.03	–8	–2
2000	92	0.12	0.37	–1.12	–0.32	–6	3
2001	92	4.84	3.99	5.86	5.44	–7	–4
2002	82	–0.24	0.39	–8.87	0.80	–15	–17
r		0.776	0.904	0.629	0.705	0.746	0.344

Fig. 2 Performance of N2 configuration and its comparison with the Rajeevan et al. [5] and actual results

The learning parameter (β) and momentum (α) used in the four models are 0.1 and 0.4, respectively, which have been fixed after number of simulations. In practice, momentum is added as an aid to have rapid convergence. The momentum constant, α, has the effect of smoothing the error surface in weight space by filtering out high frequency variations.

5 Percentage error and evaluation parameter

In the present models, percentage error E_r in prediction and an evaluation parameter denoted by σ_r is

evaluated. For good prediction $\sigma_r < 1$, i.e the error in prediction should be less than the standard deviation of the data to be trained. In all the four models, σ_r is coming out to be less than 1. The percentage error is defined as follows [3].

$$E_r = v_o - v_p$$

and the evaluation parameter is defined as in Goswami and Srividya [3]

$$\sigma_r = \frac{|v_o - v_p|}{\sigma_o},$$

where v_p and v_o are, respectively, predicted and actual values and σ_o is the standard deviation of the training data set.

6 Results and conclusions

The data set is trained using a variety of network configurations and only four of them viz., N1, N2, N3, N4 and the corresponding results after the training are given in Table 3. The four panels in Fig. 1 show plots between the rainfall for the period 1958–1995 and the percentage error in rainfall during learning of the four configurations N1–N4. The second column in Table 3 gives the actual percentage rainfall. Last column shows

Table 5 Comparison of evaluation parameter (σ_r) for the four (N1, N2, N3, N4) network configurations with Rajeevan et al. [5]

Year	Actual (%)	N1 model	σ_r	N2 model	σ_r	N3 model	σ_r	N4 model	σ_r	Rajeevan et al. [5] Prediction	σ_r
1996	103	98.23	0.42	98.26	0.42	92.93	0.89	95.79	0.64	100	0.27
1997	102	88.99	1.15	96.47	0.48	78.49	2.08	86.84	1.34	104	0.18
1998	105	111.19	0.55	111	0.53	117.92	1.14	115.07	0.89	111	0.53
1999	96	95.78	0.02	95.55	0.04	95.07	0.08	90.97	0.44	104	0.71
2000	92	91.88	0.01	91.63	0.03	93	0.08	92.32	0.02	98	0.53
2001	92	87.16	0.43	88.01	0.35	81.3	0.95	86.56	0.48	99	0.62
2002	82	82.24	0.02	81.61	0.03	91.11	0.81	81.2	0.07	97	1.33

the results given in Rajeevan et al. [5], from statistical model viz. 16-parameter model for 38 years (1958–1995) where they also compared two models viz. 8 and 10 parameter models. In the last row of Table 3 correlation coefficients (*r*) between actual and predicted are given for the four networks along with the results given in Rajeevan et al. [5].

The training of the data set using all the four network configurations are ceased as soon as the network satisfies the given RMSE value. Then the weights are stored and used for prediction. The predicted results from all the four network configurations are compared with the desired results. It was seen when comparing with desired and predicted results outside the training data, the best result is given by the network configuration *N2*. Table 4 shows the comparison between the actual results and the errors with the forecast done by the four network configurations, Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4] (it was published as a scientific correspondence) and Report [6]. The last row of the Table 4 gives the correlation coefficient (*r*) between actual and predicted for the four models and compared it with the Rajeevan et al. [5] model, Goswami and Srividya [3], Goswami and Kumar [4] models. Among the four models, it can be seen that the best result is given by *N2* model. The forecast results of the four network configurations, Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4], are compared in term of error percentage (error percentage = actual rainfall % – forecast %). It is interesting to note that actual rainfall in 2002 was 82% (drought). However, Rajeevan et al. [5] predicted the rainfall for this year as 97% (near normal) and Goswami and Srividya [3], Goswami and Kumar [4] predicted as 99% (near normal). Present study successfully predicted the rainfall for 2002 as 81.61% that is very close to actual (82%). Another reliability of the present model may be seen for the year 2000. In this year actual rainfall was 92% but Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4], predicted as 98 and 89%, respectively. Again the proposed model forecast the rainfall for 2000 as around 92%. Moreover, the Report [6] predicted rainfall for 1995 as 98% but the result as obtained by the present model is very close to the actual one. The corresponding results comparing actual, *N2* model and the results of the Rajeevan et al. [5] are depicted in Fig. 2. The above discussions and various results given show the efficiency and reliability of the present model for the good prediction of the monsoon rainfall. As discussed that the evaluation parameter (σ_r) is a significant term for good prediction if it is less than 1. Therefore corresponding evaluation parameter
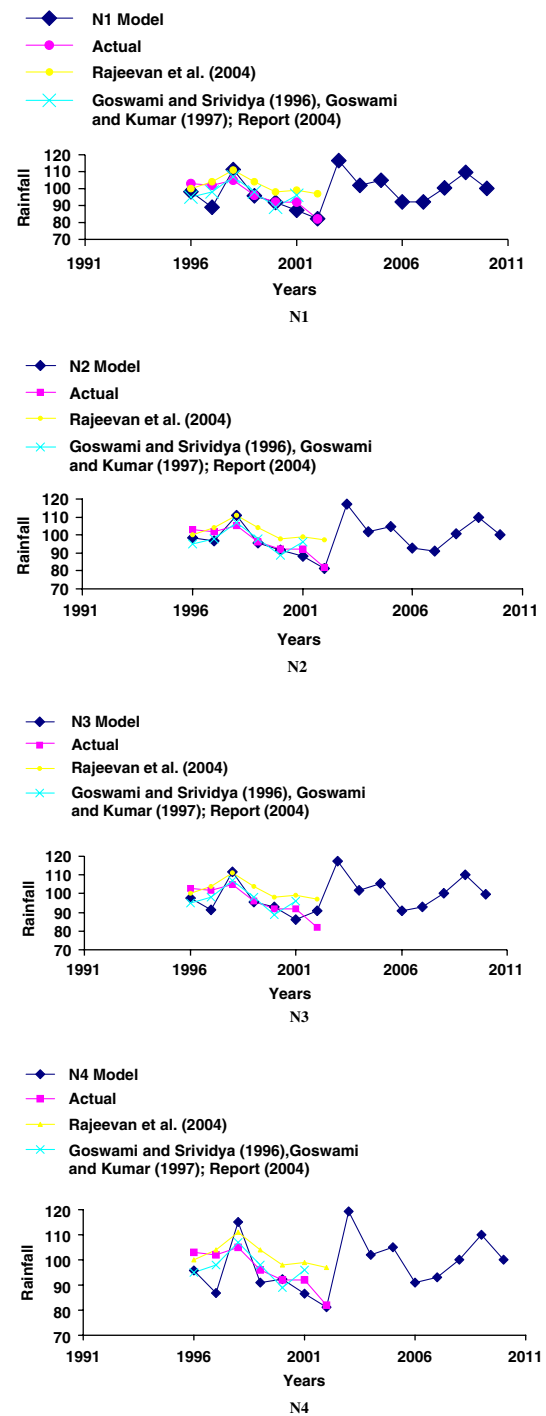

Fig. 3 Comparison of performance of *N*1, *N*2, *N*3 and *N*4 Models with actual, Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4] and Report [6]

for the years 1996–2002 (outside the training set) are computed and given in Table 5 where the standard deviation of the training set is 11.3004. This table also gives evaluation parameter for Rajeevan et al. [5]. It is worth mentioning by looking into the evaluation parameters that the *N2* model is better than the other

Table 6 Six years prediction in advance

Year	Predicted rainfall (%)	Category of rainfall
2005	104.65	Above normal
2006	92.65	Below normal
2007	91.24	Below normal
2008	100.66	Normal
2009	109.74	Above normal
2010	100.05	Normal

models. The predicted results obtained from the four models are depicted and compared with the actual, Rajeevan et al. [5], Goswami and Srividya [3], Goswami and Kumar [4], in Fig. 3. Table 6 presents prediction of the rainfall for 6 years in advance and this prediction is done using $N2$ model as it gives good results than other three models. The proposed prediction from the present ANN model, may however be checked by the future years' observations, which will certainly be interesting to see for the comparisons.

References

1. Farmer JD, Sidorowich JJ (1988) ''Exporting chaos to predict the future and reduce noise'', LA-UR-88-901, theoretical division and centre for nonlinear studies. Los Alamos National Lab, Los Alamos
2. Gadgil S, Srinivasan J, Nanjundiah R, Krishna Kumar K, Munot AA, Rupa Kumar K (2002) On forecasting the Indian summer monsoon: the intriguing season of 2002. Curr Sci 83:394–403
3. Goswami P, Srividya (1996) A novel neural network design for long range prediction of rainfall pattern. Curr Sci 70:447–457
4. Goswami P, Kumar P (1997) Experimental annual forecast of all-India mean summer monsoon rainfall for 1997 using a neural network model. Curr Sci 72:781–782
5. Rajeevan M, Pai DS, Dikshit SK, Kelkar RR (2004) IMD's new operational models for long-range forecast of southwest monsoon rainfall over India and their verification for 2003. Curr Sci 86:422–431
6. Report : Newindpress.com, features—health and science, April 8, 2004. CSIR computer model forecast rainfall 2 yrs in advance
7. Rumelhart DE, McClelland JL (1986) Parallel distributed processing. MIT Press, Cambridge