Experimental forecasts of all-India monthly and summer monsoon rainfall using neural network

P. Goswami†, Gration Crasta*, V. Sreekanth*, K. V. Shobha* and M. Kavitha Naik*

CSIR Centre for Mathematical Modelling and Computer Simulation, NAL Belur Campus, Bangalore 560 037, India

*Department of Computer Science, Mangalore University, Mangalore 574 199, India

The cognitive network for rainfall forecast was developed as an alternative forecast tool for rainfall pattern. Based on the hindcast skill and success in experimental forecast for last four years of all-India summer monsoon rainfall (ISMR), we record in the present study our cognitive forecast of ISMR for the years 1999 and 2000. While these forecasts are purely experimental, they provide an objective evaluation of the method. In addition, the present study also reports an enhanced scope of cognitive forecast of rainfall pattern, by providing experimental forecasts of all-India monthly mean precipitation for the years 1999 and 2000. These monthly forecasts also provide more stringent evaluation of the method.

ACCURATE and long-range forecast of rainfall patterns continues to be a major challenge for the scientific community. In view of the tremendous impact of rainfall patterns in areas like agricultural planning, accurate and long-range forecast of rainfall certainly requires persistent effort and a multi-pronged strategy using different techniques and methodologies[1]. One such tool is neural network (NN)[2–6], which we have been exploring and developing for several years. In particular, a generalized NN, termed cognitive network (CN), was developed and tested for hindcast skill in long-range forecast of all-India summer monsoon rainfall (ISMR)[7].

Since the development of the CN, we have followed use-and-probe strategy with the NN forecast tool. In particular, we believe that while many of the shortcomings of NN in general and CN in particular, need to be addressed at a conceptual and theoretical level, the capabilities and the scope of applicability of CN as a forecast tool also need to be explored. These two efforts, through a feedback, can, and should, lead to better understanding and further development. This is particularly relevant for a discipline like NN, which has not yet been explored and organized like other branches of classical physics, e.g. classical dynamics.

With this philosophy at the background, we have been generating experimental forecasts of ISMR for the last four years[8]. It is noteworthy that all the three forecasts for 1996, 1997 and 1998 were generated well ahead of the monsoon season, and were found to be of good quality. The CN forecast for 1998 ISMR was, for example, 945 mm (or 107% of long-term mean) while the observed value was about 106%. The success for 1998 assumes further significance since it was made in 1996, two years in advance. While it is still too early to celebrate, these successes along with the hindcast skill achieved during our investigation suggest it to be worthwhile to pursue this approach.

The purpose of this note is two-fold. First is to record our NN forecast of ISMR for 1999 and 2000. These forecasts should be interpreted as ensemble forecasts where, in the present case, an ensemble is formed by generating forecasts through a number of NN configurations with comparable hindcast skill. The principle and design of a CN have been discussed extensively in our earlier work[7]. Figure 1 represents schematically the architecture of a 2-layer CN, with structure function f and g (ref. 7). In each of the experiments below, the networks have been trained using error back propagation algorithm as in our earlier studies[7,8]. The inputs to the networks consist of past rainfall data of appropriate time scale. Tables 1 and 2 show two typical sub-ensembles, one generated from a 1-layer CN and the other from a 2-layer NN.

Table 3 gives the ensemble average forecasts for 1999 and 2000. The ensemble standard deviations in this case provide a measure of the uncertainty in the forecasts. A similar methodology was adopted for generating the forecasts for 1997 and 1998. Thus, according to these forecasts, 2000 is likely to be a deficit monsoon year. However, the forecast for 2000 will be updated if our forecast for 1999 turns out to be appreciably different from the observed value.

The second purpose of this article is to report and record an enhanced scope of our forecast system. In particular, we have explored the capability of the CN forecast system to generate all-India monthly rainfall (AIMR). Since the methodology is essentially the same as adopted and reported earlier for ISMR, we omit these details here. Essentially the same configuration (a 2-layer CN with eight neurons in each layer) was employed with a typical set of parameters as shown in Table 4.

We have used structure functions $f(x)$ and $g(x)$ defined by

$$F(x) = \tanh(\alpha_1 + \alpha_2^* x) \text{ and } g(x) = 1/(1 + e^{(\theta^* x + \beta)}),$$

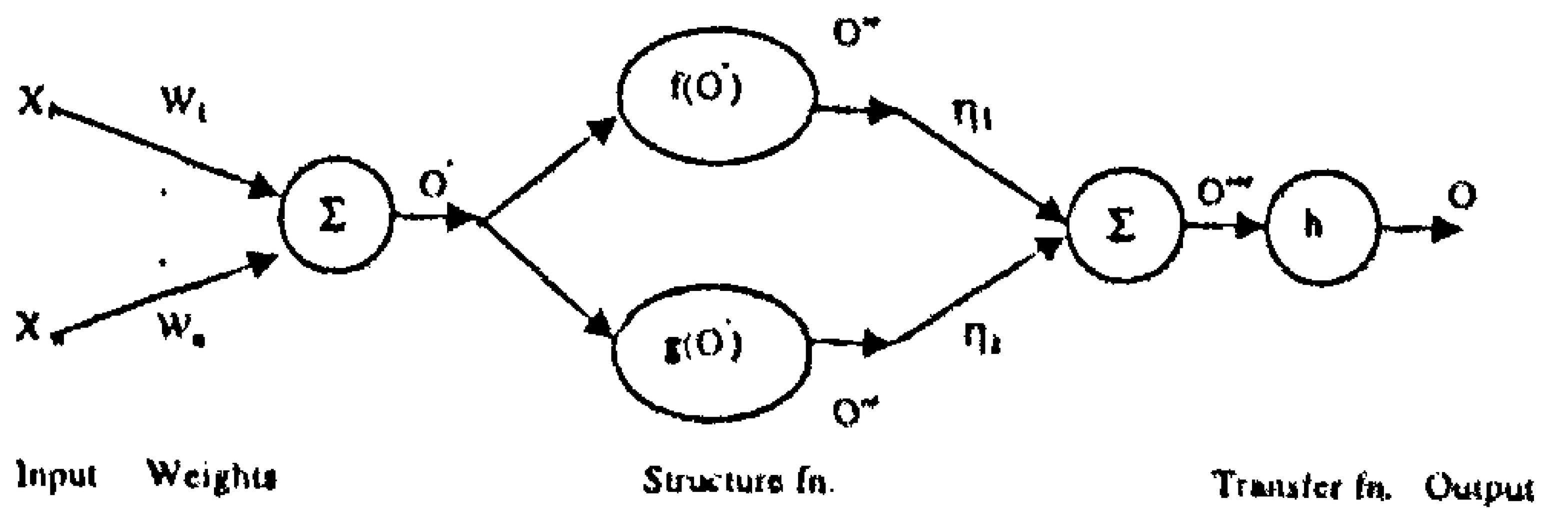

Figure 1. Schematic diagram of a 2-layer cognitive network with structure functions f and g.

†For correspondence. (e-mail: goswami@cmmacs.ernet.in)

Table 1. Sub-ensemble of ISMR forecast generated by 1-layer NN

	Experiments R (mm)					
	1	2	3	4	5	
1999	791	802	804	856	878	826
2000	738	813	830	853	800	806

Table 2. Sub-ensemble forecast generated by 2-layer NN

	Experiments R (mm)					
	1	2	3	4	5	
1999	896	896	896	897	896	896
2000	779	779	779	784	779	780

Table 3. Ensemble average forecasts for 1999 and 2000

Year	Ensemble mean (mm)	Ensemble SD (mm)
1999	861	43
2000	793	30

where typically $\beta = -0.7$ and $\theta = 0.4$. Similarly, α_1 and α_2 were taken to be 0.6 and -0.5 respectively. Here η_1 and η_2 are the cognitive weights.

The CN was then employed to train on 50 points of a 123-year data set on AIMR[9]. The following parameters were used to evaluate the performance of hindcast skill.

$$\text{Success rate,} \quad \mu = \frac{n - n^1}{n} \times 100,$$

where n is the total number of predictions and n^1 is the number of predictions out of phase. Thus for $n^1 = 0$, the success rate is 100%, while for $n = n^1$ the success rate is 0%.

$$\text{Average absolute error,} \quad < e > = \frac{1}{N} \sum_{i=1}^{N} e_i,$$

$$e_i = | X_{pi} - X_{ti} |,$$

where X_p is the predicted value and X_t is the target value. The other evaluation parameters are:

$$\gamma_1 = \frac{M_p}{M_o}, \; \gamma_2 = \frac{\sigma_p}{\sigma_o} \text{ and } \text{Bias} = M_o - M_p,$$

where M_p and M_o are the mean of the predicted and

Table 4. Summary of design of 2-layer NN forecast of AIMR

Evaluation parameters	Month											
	Jan.	Feb.	March	April	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Learning rate($\times 10^{-4}$)	3	5	1	1	1	5	5	5	5	1	5	1
Eta1η	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.4	0.7
Eta2η	0.7	0.7	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.6	0.3

Table 5. Summary of performance for AIMR hindcasts for 74 years (1921–1994)

	Month											
	Jan.	Feb.	March	April	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Avg. abs. error $<e>$ mm	9.9	9.8	7.7	7.3	16.2	54	65	53	44	37	24	9.7
Success rate (%) μ	50	48	36	60	43	46	48	43	48	38	45	46
γ_1	0.8	0.9	1	1	1	1	1	1	1	1	0.8	1
γ_2	0.7	0.9	1	1	1	0.7	0.6	0.9	0.8	0.8	0.8	1
Bias	– 2.5	– 1	2.6	1.7	5	19.2	6.1	13.3	10.7	– 1.2	– 8.1	1.3

Table 6. Forecasts (in mm) of AIMR generated by a 2-layer NN for 1995–2000

	Month											
	Jan.	Feb.	March	April	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
1995	6.8	6.5	13.8	14.6	49.8	87.4	308.6	247	220.7	26.4	41.3	20.7
1996	4.7	2.9	15.3	31.9	51.0	109.8	284	239	158.9	14.9	21.9	8.7
1997	2.2	21.0	13.8	24.1	45.0	69.7	286.4	205	196.5	36.2	31.0	11.4
1998	7.0	11.2	14.3	29.7	36.4	240.2	238.6	231	197.4	30.7	23.6	12.3
1999	2.4	6.5	14.0	16.1	51.9	286.5	245.6	235	156.0	36.7	25.1	12
2000	4.9	8.0	21.2	24.0	49.7	98.9	224.6	230	102.8	45.3	12.0	8.97

Table 7. Comparison between observed forecast ISMR and sum of forecast AIMR

Year	Observed forecast ISMR A (mm)	Sum of forecast AIMR (June–Sept.) B (mm)	A/B
1995	944	864	1.1
1996	991	792	1.3
1997	886	756	1.2
1998	889	808	1.1
1999	861	1044	0.8
2000	793	656	1.2

observed values; and σ_p and σ_o are the standard deviation (SD) of the predicted and observed values respectively.

Table 5 summarizes the hindcast skill of a 2-layer CN for 12 months for 74 hindcasts.

While the hindcast skill for AIMR is poor when compared to ISMR, it is encouraging to note that average error is still smaller than the SD of the data in most cases. We record here, therefore, our experimental forecast for 12 months for the year 1999 and 2000 (Table 6).

We would like to emphasize that these forecasts are purely experimental, and not meant for any operational use. The success of the previous years is to be accepted with usual and necessary caution. In case of AIMR, there are greater uncertainties as the forecast (or hindcast) skill is somewhat inversely proportional to the SD of the data. More importantly, the present forecasts of AIMR are like a pattern-forecast for the entire period from 1995 to 2000, since the observed data in this case was employed only until 1994. These monthly forecasts, on the other hand, will put much more stringent tests on the NN forecasts. For example, the sum of AIMR for June to September should also show a good agreement with forecast for ISMR, so that an organized hierarchy of NN forecast systems can be developed. In the present case, this condition was tested and verified. As an example, Table 7 shows a comparison between observed ISMR along with forecast ISMR for 1999 and 2000 and the sum of forecast AIMR for June to September for 1995 to 2000.

Even moderate success in AIMR can have enormous impact in areas like agricultural planning. For example, both for 1998 and 1999, and especially for 2000, the low value of forecast ISMR is due to significant deficit in the AIMR for September. Indeed, the ISMR for 2000 is seen to be peaked in essentially two months, July and August. These features have practical and significant implications for issues like crop choice, irrigation planning and sowing schedule. For the same reason, however, a very thorough and objective evaluation of the forecast skill is necessary. We hope that the results published in this article will provide a definite step in this direction.

1. Das, P. K., *Monsoons* (eds Fein, J. S. and Stephens, P. L.), John Wiley, New York, 1987, pp. 549–578.
2. Farmer, J. D. and Sidorowich, J. J., *Exploiting Chaos to Predict the Future and Reduce Noise*, LA-UR-88-901, Theoretical Division and Centre for Nonlinear Studies, Los Alamos National Lab., 1988.
3. Tong, H., *Threshold Models in Nonlinear Time Series Analysis*, Springer, 1983, p. 323.
4. Elser, J. B. and Tsonis, A. A., *Bull. Am. Meteorol. Soc.*, 1992, **73**, 49–59.
5. Navone, H. D. and Ceccatto, H. A., *Climate Dyn.*, 1994, **10**, 305–312.
6. Grieger, B. and Latif, M., *Climate Dyn.*, 1994, **10**, 267–276.
7. Goswami, P. and Srividya, *Curr. Sci.*, 1996, **70**, 447–457.
8. Goswami, P. and Pradeep Kumar, *Curr. Sci.*, 1997, **72**, 781–782.
9. Parthasarathy, B., Munot, A. A. and Kothawala, D. R., IITM Research Report No. RR-065, 1995.

ACKNOWLEDGEMENTS. G.C., V.S., K.V.S. and M.K.N. thank Dr R. N. Singh, Scientist-in-Charge, C-MMACS for permission and encouraging support to carry out this project work at C-MMACS.

Received 21 November 1998; revised accepted 8 March 1999

Use of silicon carbide fibers for *Agrobacterium*-mediated transformation in wheat

Neeta Singh and H. S. Chawla*

Department of Genetics and Plant Breeding, G.B. Pant University of Agriculture and Technology, Pantnagar 263 145, India

Silicon carbide fibers have been used for causing wounds in the immature wheat embryos for *Agrobacterium*-mediated genetic transformation. Immature embryos were put in a 5% silicon carbide fiber (SCF) suspension and vortexed for 2 or 3 min followed by co-cultivation with *Agrobacterium tumefaciens* strain LBA4404. The strain harboured the binary vectors pBI121 and pTOK233 which contained selectable marker genes and *gus* as the reporter gene. *Agrobacterium*-infected explants were stained for GUS activity. Without wounding the GUS expression was observed in 2.4% of the embryo, while 33.3% of the embryos showed GUS expression after wounding with SCF for 2 min. The binary vector pBI121 showed better response than pTOK233 for GUS expression. We propose that SCFs can be used for wounding to improve frequency of transformation by disarmed *Agrobacterium* strains.

TRANSFORMATION studies were conducted on wheat with the *Agrobacterium* strain LBA4404 which carried the binary vectors pBI121 (Clonetech, USA) and pTOK233 (Yuko Hiei, Japan). The T-DNA of pBI121 contained the *gus* reporter gene controlled by 35S promoter and *nptII* selectable marker gene controlled by the nopaline synthase (NOS) promoter. The binary vector pTOK233 contained the *gus* reporter gene controlled by CaMV35S promoter, which had an intron in the N-terminal region of

*For correspondence. (e-mail: ag@gbpuat.ernet.in)