MAUSAM, **53**, 2 (April 2002), 225-232

551.509

Short range SW monsoon rainfall forecasting over India using neural networks

PANKAJ JAIN

Indian Institute Technology, Kanpur, India

and

ASHOK KUMAR

PARVINDER MAINI, S. V. SINGH

Department of Science and Technology, NCMRWF, New Delhi

(*Received 12 December 1997, Modified 1 August 2001*)

सार - दैनिक वर्षा के पूर्वानुमान के लिए संपूर्ण भारत के विभिन्न परीक्षण केन्द्रों पर फीड फावर्ड न्यूरल संजालों का उपयोग किया गया है। इस कार्य के लिए 1985 से 1988 तक के चार वर्षों के आँकड़ों वाले ट्रेनिंग सेट तथा 1989-1990 के आँकड़ों वाले प्रमाणिक सेट सहित यूरोपीय मध्यम अवधि मौसम पूर्वानुमान केन्द्र (ई.सी.एम.डब्ल्यू.एफ.) के छः वर्षों के आँकड़ों का उपयोग किया गया है। वर्षा और अन्य वायुमंडलीय परिवर्तनों के मध्य अनुकूल संबंध को विकसित करने के लिए न्यूरल संजालों का उपयोग किया गया है। इस अध्ययन के लिए अधिकतम परिवर्तनों के चयन हेतू कोई प्रयास नहीं किया गया है। चुने गए परिवर्तन राष्ट्रीय मध्यम अवधि मौसम पूर्वानुमान केन्द्र (ई.सी.एम.डब्ल्यू.एफ.) में रैखिक समाश्रयण निदर्श को विकसित करने के लिए पहले से प्राप्त किए गए परिवर्तनों के समान ही हैं। न्यूरल संजालों से प्राप्त हुए परिणाम रैखिक समाश्रयण के समान ही बेहतर पाए गए हैं। अनेक मामलों में इनमें 10 से 20 प्रतिशत तक सुधार आया है। यह निश्चय ही सराहनीय तथ्य है कि चुने गए परिवर्तन रैखिक समाश्रयण के लिए अभी तक सर्वोतम पाए गए हैं।

ABSTRACT. Feedforward Neural Networks are used for daily precipitation forecast using several test stations all over India. The six year European Centre of Medium Range Weather Forecasting (ECMWF) data is used with the training set consisting of the four year data from 1985-1988 and validation set consisting of the data from 1989-1990. Neural networks are used to develop a concurrent relationship between precipitation and other atmospheric variables. No attempt is made to select optimal variables for this study and the inputs are chosen to be same as the ones obtained earlier at National Center for Medium Range Weather Forecasting (NCMRWF) in developing a linear regression model. Neural networks are found to yield results which are atleast as good as linear regression and in several cases yield 10 - 20 % improvement. This is encouraging since the variable selection has so far been optimized for linear regression.

Key words – Neural network, Rainfall, Forecasting, Southwest monsoon.

1. Introduction

Feedforward artificial neural networks (Masters, 1993; Müller and Reinhardt, 1991) have proven to be very useful in solving a wide range of pattern recognition problems. In weather forecasting also they have been used to predict tornado formation (Marzban and Stumpf, 1995), thunderstorm forecasting (McCann, 1992), monsoon rainfall forecasting (Navone and Ceccatto, 1994), cloud type classification (Bankert, 1994) *etc*. In the present paper we will apply neural network to study the problem of short term precipitation forecasting.

The procedure currently in use at National Center for Medium Range Weather Forecasting (NCMRWF) for daily forecasting uses a man-machine combination where a meteorologist considers several different inputs along with human judgement to produce the final weather forecast. The inputs are the forecast produced by NCMRWF T-80 model, statistical interpretation of the T-80 model output and the satellite imagery. Previous studies (Kumar and Maini, 1996) have shown that statistical interpretation is more reliable than the direct model output (DMO). The perfect prog method (PPM), currently being used for building the statistical model, uses direct observations of atmospheric variables for fixing the regression model parameters. The data used for development of PPM equations is the analysis data obtained from European Centre of Medium Range Weather Forecasting (ECMWF). This data is at 2.5 × 2.5

degree grid resolution and is refered as ECMWF/TOGA basic level-III data set. In the absence of sufficient data needed for development of PPM equations ECMWF analysis data was used and it has shown a good relation with the observed rainfall at different Indian stations, which is also clear from the skill of the forecast obtained from these PPM equations (Kumar and Maini, 1996).

The statistical model essentially provides a relationship between observed variables for a particular day with the observed rainfall on the same day. The atmospheric model itself plays no role at this stage. Once the parameters of the regression equation are fixed, it is used in conjunction with atmospheric model, such as T-80, to provide real time forecast. The T-80 model is necessary for the purpose of time evolving the variables to the day when the forecast is desired. An alternative approach, Model Output Statistics (MOS), where the regression model is trained on the atmospheric model output is expected to be superior but could not be implemented so far since the NCMRWF T-80 model became operational only in 1993. In time, after sufficient data is accumulated, it might be better to replace the PPM method with the MOS technique.

In the current study we will determine if replacing the statistical regression procedure with neural networks leads to better results. The neural network simulations will be performed by using the same predictors that are being used at NCMRWF for the forecast based on linear regression analysis. The output of the model consists of either the Probability of Precipitation (PoP) or the Quantity of Precipitation. In the case of Quantity of Precipitation it was found that it is better to consider the Cube Root of Precipitation (CRP) as the output variable due to the very long tail of the observed rainfall distribution. The PoP is interpreted as predicting a rainy day if its value is large than or equal to 0.5 and a dry day if its value is less than 0.5. In the present paper we shall concentrate on 24 hour forecast.

Neural networks have a massively parallel, layered structure with each layer consisting of several nodes called neurons. They provide a mapping from the input vector x_i, i=1,2,...,n, to the output vector y_j, j=1,2,...,m. Besides the input and output layers the network may also contain one or more hidden layers. Each neuron produces an output O = $f(Z)$, where $Z = \sum_i z_i w_i + b$, z_i (i=1,2,...,n) are the inputs to the given neuron, f(Z) is called the activation function and is usually taken to be the sigmoidal function $1/[1+\exp(-Z)]$, w_i are the weights associated with the network and b is the bias of the neuron. The weights w_i and the bias b represent the parameters of the network which are to be determined by using some known data sample, called the training set, of the pattern to be learned. Neural networks have been found to have remarkable ability for pattern recognition. It has been shown that a two hidden layer network can learn most functions with compact domain. More details on neural networks and their applications can be found in several books written on the subject such as (Masters, 1993; Müller and Reinhardt, 1991).

2. Data

The study was conducted by using the six year data compiled by the European Center for Medium Range Weather Forecasting (ECMWF) for several stations such as Delhi, Ludhiana, Bangalore, Srinagar, Hyderabad *etc* in India for the monsoon season during the years 1985-1990. The monsoon season comprises of the months June, July and August in North-West India and of months June – September in the rest of India. The data for each station consists of several variables such as temperature, relative humidity, wind velocity *etc* measured at ground level and at high altitudes. These variables are available twice every day (0000 UTC and 1200 UTC) at grid locations all over India. Besides these directly measured variables several derived variables such as precipitable water, saturation deficit, temperature gradient *etc* are also included in the data. At each grid location a total of 47 variables are available. ECMWF data is used since this is the only complete compilation available at the time of this study. In future, sufficient more reliable data may become available at NCMRWF after the T-80 model has been used for several years.

The Probability of Precipitation (PoP) and the Quantity of Precipitation (QP) are predicted independently using separate data sets with independent variable selection. In the case of QP, the cube root of precipitation (CRP) is used as the dependent variable. In order to determine the values of each of these 47 variables, which are to be considered as predictors at the station, the variable values of nine grid points surrounding the station are considered. The value of each of these predictors at the station are obtained from the nine grid point values by using canonical correlations. The predictors for both of the output variables, *i.e*. CRP and PoP, have been selected at NCMRWF (Kumar and Maini, 1996) on the basis of percentage of variance explained. We have taken these variables as inputs for the neural network studies and concentrate on the comparison of the final fits of neural networks and the linear regression. The number of predictors vary with station as well as depend on whether the output is the rainfall or the probability of precipitation. For example, for cube root of precipitation (CRP) the data consisted of nine inputs and one output for the case of Delhi and Ludhiana and ten inputs and one output for the

TABLE 1

The two most significant parameters used for the cube root of precipitation (CRP) for all the stations considered in this paper

Station	Cube Root of Precipitation (CRP)		
	Parameter	Level (hPa)	Reference time
Delhi	Mean relative humidity	1000-500	Average
	Vertical velocity	700	0000 UTC
Ludhiana	Mean relative humidity	1000-500	Average
	Vertical velocity	500	Average
Hissar	Relative humidity	500	1200 UTC
	Geopotential height	500	Average
Anand	Relative humidity	700	Average
	Vorticity	1000	Average
Kanpur	Mean relative humidity	1000-500	1200 UTC
	Vertical velocity	500	0000 UTC
Pune	Precipitable water	1000-500	Average
	Temperature gradient	850-700	Average
Trichur	Geopotential height	1000	Average
	Relative humidity	850	Average
Solan	Geopotential height	850	Average
	Vertical velocity	500	0000 UTC
Raipur	Relative humidity	500	0000 UTC
	Vertical velocity	500	1200 UTC
Jabalpur	Mean relative humidity	1000-500	Average
	Vertical velocity	500	Average
Hyderabad	Mean relative humidity	1000-500	Average
	Vertical velocity	500	Average
Bangalore	Relative humidity	1000	Average
	Relative humidity	700	Average
Srinagar	Vertical velocity	500	Average
	Vorticity	1000	1200 UTC

TABLE 2

The two most significant parameters used for the Probability of Precipitation (PoP) for all the stations considered in this paper

Station	Probability of Precipitation (PoP)		
	Parameter	Level (hPa)	Reference time
Delhi	Mean relative humidity	1000-500	Average
	Vertical velocity	850	Average
Ludhiana	Mean relative humidity	1000-500	Average
	Geopotential height	850	Average
Hissar	Relative humidity	500	1200 UTC
	Geopotential height	500	Average
Anand	Mean relative humidity	1000-500	Average
	Vorticity	1000	Average
Kanpur	Mean relative humidity	1000-500	1200 UTC
	Vertical velocity	500	0000 UTC
Pune	Mean relative humidity	1000-500	Average
	Temerature gradient	850-700	Average
Trichur	Temperature	850	Average
	Saturation deficit	1000-500	Average
Solan	Geopotential height	850	Average
	Relative humidity	700	0000 UTC
Raipur	Mean relative humidity	1000-500	Average
	Geopotential height	850	Average
Jabalpur	Mean relative humidity	1000-500	1200 UTC
	Vorticity	1000	Average
Hyderabad	Mean relative humidity	1000-500	Average
	Vertical velocity	500	Average
Bangalore	Geopotential height	700	0000 UTC
	Relative humidity	1000	Average
Srinagar	Vertical velocity	500	Average
	Zonal wind component	1000	1200 UTC

case of Hissar. For PoP instead, the data had five inputs for Delhi, six for Ludhiana and seven for Hissar. The two most significant predictors for the case of CRP and PoP are given in Tables 1 and 2 respectively for all the stations. The complete list of predictors considered is given in Kumar *et al.* (1999). We divided the data into two sets corresponding to the training and validation set. The training set for all the cases is taken to be the four year data from 1985 to 1988. The 1989 and 1990 data is used for validation. For example, for the case of Delhi the total number of samples used for training and validation were 362 and 182 respectively.

3. Procedure

The training algorithm used in the neural network for minimization of error is the conjugate gradients procedure complemented by simulated annealing to evade local minima. The conjugate gradients method is expected to be more efficient than the more commonly used back propagation algorithm and hence the network is expected to learn faster (Masters, 1993). The simulated annealing method is necessary to escape from local minima which are usually present abundantly in the error function. We tried a series of networks with increasing complexity in order to determine which one performs the best. The error measure was taken to be the usual mean squared sum of errors. The success of the network has to be gauged in terms of several different measures, besides the mean squared sum of errors, which may be better indicators of its skill in terms of weather prediction. We use the ratio measure, H.K. index, B.S. as well as the mean square error calculated using the validation data set as the indicators to determine the performance of the network.

TABLES 3-15

Comparison of the performance of feed forward neural network with the regression analysis for (a) Probability of Precipitation (PoP) and (b) Cube Root of Precipitation (CRP) for various stations spread all over India. The results are given for the validation set using several different measures of the skill of the network which include B. S., Ratio measure and H. K. index, as defined in the text. The total number of data points for the stations Delhi, Ludhiana, Hissar, Anand, Kanpur, Solan, Raipur, Jabalpur and Srinagar used for training and validation are 362 and 182 respectively. For the remaining stations the corresponding number of data points are 482 and 242 respectively. The nature of the neural network model used, *i.e.* the number of layers and the number of neurons in each layer, is specified by the title of each model. For example, in the case of the neural network model 8-2-1, there are three layers. The first, *i.e.* the input layer, has eight neurons corresponding to the eight inputs. The next (hidden) layer has two neurons and the third (output) layer has one neuron

3a. PoP (Delhi)

Model	B.S.	Ratio	H.K.
Linear Regression	0.17	0.753	0.46
N N 5 - 1	0.17	0.77	0.504
N N 5 - 1 - 1	0.17	0.764	0.496
N N 5 - 2 - 1	0.17	0.764	0.50

3b. CRP (Delhi)

Model	Ratio	H.K.	RMSE
Linear Regression	0.670	0.45	12.03
N N 9 - 1	0.692	0.453	12.01
N N 9 - 1 - 1	0.731	0.492	12.00
N N 9 - 2 - 1	0.731	0.492	12.01

4a. PoP (Ludhiana)

Model	B.S.	Ratio	H.K.
Linear Regression	0.14	0.79	0.44
N N 7 - 1	0.14	0.81	0.504
N N 7 - 1 - 1	0.14	0.81	0.504
N N 7 - 2 - 1	0.14	0.80	0.54

4b. CRP (Ludhiana)

Model	Ratio	H.K.	RMSE
Linear Regression	0.676	0.50	13.41
N N 9 - 1	0.70	0.51	14.6
N N 9 - 1 - 1	0.703	0.512	13.9
N N 9 - 2 - 1	0.71	0.50	14.6

5a. PoP (Hissar)

Model	B.S.	Ratio	H.K.
Linear Regression	0.17	0.758	0.27
N N 6 - 1	0.16	0.758	0.292
N N 6 - 1 - 1	0.16	0.758	0.281
N N 6 - 2 - 1	0.17	0.753	0.27

5b. CRP (Hissar)

Model	Ratio	H.K.	RMSE
Linear Regression	0.665	0.42	7.45
N N 10 - 1	0.69	0.40	7.02
N N 10 - 1 - 1	0.764	0.427	6.46
N N 10 - 2 - 1	0.75	0.30	6.84

6a. PoP (Anand)

Model	B.S.	Ratio	H.K.
Linear Regression	0.16	0.758	0.52
N N 4 - 1	0.17	0.758	0.52
N N 4 - 1 - 1	0.17	0.742	0.488
N N 4 - 2 - 1	0.17	0.742	0.49

6b. CRP (Anand)

Model	Ratio	H.K.	RMSE
Linear Regression	0.714	0.44	21.67
N N 6 - 1	0.725	0.46	20.98
N N 6 - 1 - 1	0.747	0.50	22.70
N N 6 - 2 - 1	0.753	0.51	20.97

7a. PoP (Kanpur)

Model	B.S.	Ratio	H.K.
Linear Regression	0.15	0.786	0.56
N N 5 - 1	0.15	0.791	0.57
N N 5 - 1 - 1	0.15	0.791	0.57
N N 5 - 2 - 1	0.15	0.791	0.57

7b. CRP (Kanpur)

Model	Ratio	H.K.	RMSE
Linear Regression	0.626	0.37	15.71
N N 6 - 1	0.650	0.40	15.44
N N 6 - 1 - 1	0.731	0.49	16.38
N N 6 - 2 - 1	0.731	0.49	16.37

8a. PoP (Pune)

Model	B.S.	Ratio	H.K.
Linear Regression	0.17	0.769	0.37
N N 4 - 1	0.16	0.773	0.383
N N 4 - 1 - 1	0.16	0.773	0.383
N N 4 - 2 - 1	0.16	0.773	0.389

8b. CRP (Pune)

Model	Ratio	H.K.	RMSE
Linear Regression	0.773	0.39	9.24
N N 10 - 1	0.773	0.40	9.93
N N 10 - 1 - 1	0.773	0.41	10.09
N N 10 - 2 - 1	0.769	0.39	9.74

9a. PoP (Trichur)

Model	B.S.	Ratio	H.K.
Linear Regression	0.09	0.864	0.50
N N 4 - 1	.087	0.868	0.553
N N 4 - 1 - 1	.086	0.868	0.538
N N 4 - 2 - 1	.081	0.876	0.637

9b. CRP (Trichur)

Model	Ratio	H.K.	RMSE
Linear Regression	0.855	0.42	14.44
N N 7 - 1	0.839	0.31	14.6
N N 7 - 1 - 1	0.855	0.42	14.53
N N 7 - 2 - 1	0.839	0.37	15.30

10a. PoP (Solan)

Model	B.S.	Ratio	H.K.
Linear Regression	0.16	0.775	0.53
N N 6 - 1	0.16	0.780	0.545
N N 6 - 1 - 1	0.16	0.769	0.530
N N 6 - 2 - 1	0.16	0.769	0.530

10b. CRP (Solan)

Model	Ratio	H.K.	RMSE
Linear Regression	0.665	0.38	18.77
N N 6 - 1	0.665	0.38	20.12
N N 6 - 1 - 1	0.687	0.42	17.79
N N 6 - 2 - 1	0.681	0.391	20.42

11a. PoP (Raipur)

Model	B.S.	Ratio	H.K.
Linear Regression	0.15	0.791	0.40
N N 4 - 1	0.15	0.791	0.412
N N 4 - 1 - 1	0.15	0.775	0.40
N N 4 - 2 - 1	0.15	0.791	0.443

11b. CRP (Raipur)

Model	Ratio	H.K.	RMSE
Linear Regression	0.714	0.15	16.31
N N 11 - 1	0.714	0.174	16.97
N N 11 - 1 - 1	0.703	0.15	16.65
N N 11 - 2 - 1	0.703	0.15	16.65

12a. PoP (Jabalpur)

Model	B.S.	Ratio	H.K.
Linear Regression	0.17	0.769	0.50
N N 4 - 1	0.17	0.764	0.50
N N 4 - 1 - 1	0.17	0.753	0.48
N N 4 - 2 - 1	0.17	0.769	0.51

12b. CRP (Jabalpur)

Model	Ratio	H.K.	RMSE
Linear Regression	0.714	0.36	14.93
N N 8 - 1	0.703	0.33	14.89
N N 8 - 1 - 1	0.725	0.39	14.69
N N 8 - 2 - 1	0.709	0.35	16.02

13a. PoP (Hyderabad)

Model	B.S.	Ratio	H.K.
Linear Regression	0.18	0.731	0.44
N N 3 - 1	0.18	0.731	0.44
N N 3 - 1 - 1	0.18	0.731	0.43
N N 3 - 2 - 1	0.18	0.736	0.44

13b. CRP (Hyderabad)

Model	Ratio	H.K.	RMSE
Linear Regression	0.603	0.32	10.31
N N 6 - 1	0.624	0.35	10.04
N N 6 - 1 - 1	0.657	0.39	11.59
N N 6 - 2 - 1	0.645	0.37	10.23

14a. PoP (Bangalore)

Model	B.S.	Ratio	H.K.
Linear Regression	0.20	0.678	0.32
N N 4 - 1	0.20	0.682	0.33
N N 4 - 1 - 1	0.20	0.682	0.33
N N 4 - 2 - 1	0.20	0.674	0.31

14b. CRP (Bangalore)

Model	Ratio	H.K.	RMSE
Linear Regression	0.587	0.25	10.11
N N 8 - 1	0.63	0.29	9.96
N N 8 - 1 - 1	0.63	0.28	9.96
N N 8 - 2 - 1	0.62	0.28	10.00

15a. PoP (Srinagar)

Model	B.S.	Ratio	H.K.
Linear Regression	0.12	0.852	0.42
N N 8 - 1	0.12	0.835	0.40
N N 8 - 1 - 1	0.12	0.835	0.40
N N 8 - 2 - 1	0.13	0.819	0.36

15b. CRP (Srinagar)

Model	Ratio	H.K.	RMSE
Linear Regression	0.736	0.55	4.23
N N 6 - 1	0.731	0.55	4.26
N N 6 - 1 - 1	0.835	0.55	4.63
N N 6 - 2 - 1	0.824	0.55	4.41

The ratio measure is equal to the number of correctly predicted days divided by the total number of days, B. S. is equal to the sum of square errors divided by the total number of days and the H. K. index is defined as,

$$\text{H.K.} = \frac{\text{N(dry)} \times \text{N(wet)} - \text{M(dry)} \times \text{M(wet)}}{[\text{N(dry)} + \text{M(dry)}] \times [\text{N(wet)} + \text{M(wet)}]}$$

where N(dry) and M(dry) are the total number of dry days correctly and incorrectly predicted respectively with analogous definitions for N(wet) and M(wet). The results of these networks are compared with the linear regression procedure currently being used by NCMRWF.

It is important to understand that in the present context neural network is essentially being used as a nonlinear regression model. The ability of neural networks to learn a large set of functions with relatively few neurons makes them the preferred choice in comparison to other nonlinear models. It is essential that neural network does not over fit the data since then it would loose its power of generalization. This requires that we use as small a neural network as possible with the total number of weights much smaller in comparison to the available data points for training. In our simulations we start with the simplest possible network and then start increasing its complexity till the training error does not show any significant decrease. We simultaneously test it on the validation set to check if the network has not over fitted the data. We do not make any attempt to learn data sample with zero error as is sometimes done in literature. This is an acceptable procedure if one is interested in only storing some known patterns with which other patterns can later be compared. However it is not applicable in the present case since the data is likely to be very noisy and does not necessarily contain the complete information which determines the output.

4. Results and discussion

The validation results of our simulations both for probability of precipitation and quantity of precipitation for the independent period of 1989-1990 are given in Tables 3-15. The data for years 1985-1988 was used for training. In these tables the results of the neural network are compared with those of the regression model for a twenty four hour forecast. Several different neural network models are compared as listed in the first column of each table. The neural network model specifies the number of neurons in each layer starting from the first layer. For example the neural network model 5-2-1 specifies that there are three layers in the model with the input layer containing five neurons, the next layer containing two neurons and the third (output) layer containing one neuron.

We find that overall the results are slightly better with neural networks in comparison to linear regression. In the case of PoP the B.S. and the ratio does not show much change for almost all the stations. However the H. K. index shows improvement in several cases which include Delhi, Ludhiana, Trichur and Raipur. In none of the cases does the H. K. index go down significantly. The fact that the H. K. index gives better results for these stations implies that the overall skill of the network in identifying the wet days is better than that of the regression model. In the case of CRP the ratio shows improvement in many stations which include Delhi, Ludhiana, Hissar, Anand, Kanpur, Hyderabad, Bangalore and Srinagar. The H. K. index also shows improvement for Delhi, Anand, Kanpur and Hyderabad. In none of the cases do we find worse results in comparison to linear regression. The RMSE does not show significant change in any of the stations. Hence in the case of CRP we find that although the RMSE does not show any improvement, the overall skill of the neural networks in identifying a wet day is significantly better in comparison to the linear regression. To summarize, we find that in many cases the neural network leads to 10-20% improvement in the validation results and rarely does it perform worse than the regression model. This small improvement is encouraging since the predictor variables have so far been selected to give optimal results with linear regression. Our results also show that the hidden layers have so far not been found to be of much use and in most cases the results with no hidden layers are not much different with the inclusion of hidden layers. This indicates that perhaps strong nonlinearities are not present in the data. It is important to understand the source of this behavior since a priori we may expect that the amount of rainfall as well as the probability of precipitation should be a highly nonlinear function of other atmospheric variables.

One possible reason for the absence of nonlinearities is that the ECMWF data, used in this study, is available only on a rather coarse grid of 2.5° latitude × 2.5° longitude. Precipitation amount at a given station shows strong spatial dependence even on the scale of a few km. It is clearly not possible to predict this small scale variation on the basis of the coarse grid data available. One way this problem might be solved in future is through use of satellite imagery, which can provide information on cloud cover and atmospheric moisture content at considerably better spatial resolution. However considerable research work is required in order to extract relevant information from satellite images before it can be used for this purpose.

TABLE 16

Results of the neural network simulations for (a) Delhi (CRP) (b) Ludhiana (CRP) (c) Hissar (CRP) (d) Kanpur (CRP) (e) Srinagar (PoP) after systematically reducing the number of inputs. Here CRP and PoP stand for Cube Root of Precipitation and Probability of Precipitation respectively. The top row gives the neural network model used. The number of inputs in each case are specified by the first digit of the model (For example in the case of model 5-1, there are 5 inputs). The validation results are given in terms of Ratio, H.K. index and RMSE for CRP and in terms of Ratio, H. K. index and B. S. for PoP. We clearly see that although the training error increases slightly, the validation results remain almost unchanged or may even show an improvement in some cases

(a) Delhi (CRP)

NN model	9-1	5-1	3-1	2-1	2-1-1
Ratio	0.69	0.69	0.68	0.67	0.70
H.K.	0.45	0.45	0.43	0.43	0.46
RMSE	12.01	12.06	11.86	11.89	11.87
Training Error	5.34	5.73	5.94	6.01	5.89

(b) Ludhiana (CRP)

NN model	9-1	5-1	3-1	2-1	1-1
Ratio	0.70	0.70	0.71	0.71	0.71
H.K.	0.51	0.50	0.52	0.52	0.54
RMSE	14.6	13.54	13.91	13.67	13.87
Training Error	7.00	7.59	7.97	8.27	8.99

(c) Hissar (CRP)

NN model	10-1	5-1	3-1	2-1	1-1
Ratio	0.69	0.72	0.68	0.69	0.71
H.K.	0.40	0.42	0.37	0.42	0.46
RMSE	7.02	7.32	7.37	7.35	7.72
Training Error	3.77	3.96	4.15	4.35	4.78

(d) Kanpur (CRP)

NN model	6-1	5-1	2-1	2-1-1	2-2-1
Ratio	0.65	0.70	0.66	0.72	0.72
H.K.	0.40	0.48	0.43	0.53	0.53
RMSE	15.44	15.69	16.09	16.36	16.77
Training Error	6.33	6.38	6.71	6.59	6.54

(e) Srinagar (PoP)

NN model	8-1	5-1	4-1	4-2-1
Ratio	0.84	0.85	0.86	0.86
H.K.	0.40	0.46	0.48	0.51
B.S.	0.12	0.11	0.12	0.12
Training Error	30.74	32.13	32.5	32.55

Another potential problem is that the variable selection procedure used so far might bias the data towards linear regression. For example, the choice of variables at the test station from their values at the surrounding grid locations have been selected to give best linear correlation with the output rainfall. Furthermore the selection of variables for regression model from the large set of potential predictors has again been done on the basis of linear correlation. We have not used neural networks for this purpose so far, since this would be computationally very time intensive. It is important to realize that the entire six year data set was utilized for the purpose of this selection. Therefore certain amount of optimization for the validation set has already been done in the predictor selection procedure. Given this bias towards linear regression, we find it encouraging that neural networks are able to do as well as the regression model and in many cases better.

We also carried out some further simulations to determine the relative significance of the selected variables. We show some representative results obtained after systematically reducing the input variables in Table 16. The variables are deleted in the order of their increasing correlation with the output. Hence the variable with least correlation gets deleted first. We expect that the results will be best if all the variables are included and the error will increase monotonously as the input variables are reduced. This is indeed found to be the case for the training error. However for the case of validation error we found that in several cases it does not change appreciably even as we reduce the input variables to a very small number. In some cases the error even goes down as the number of variables are reduced. This suggests that for these cases most of these variables might be redundant and might simply be introducing noise. However considerable more effort is required in order to determine if this is indeed the case. The present results are only suggestive since it is based only on a few representative stations. Furthermore our main objective is to compare the results of neural network with regression analysis given a particular set of predictors. Hence in the present paper we do not attempt to make a new selection of variables.

We should point out that in our study we have restricted ourselves to relatively small networks. Our basic philosophy was to increase the complexity of the network until training and validation error does not show significant decrease. We have also avoided using excessively large number of different network topologies. If one continues to play with network topology sufficient number of times it is quite likely that one may hit upon a fit which performs well both in training and validation set. However, in that case it is not clear whether we have trained the network properly or merely used the validation set also implicitly as a part of the training set. We also point out that in most of our simulations the stopping criteria has been to attain the minimum of the training error. In other words the iterations are stopped after the training error does not decrease any further (or decreases less than some pre-determined amount with successive iterations). One may also try to use an alternative approach where the iterations are stopped after the validation error has reached a minimum. We find that this also leads to only small improvement in validation results.

The present study suggests several directions for future work. It shows that proper selection of input variables is very important. It is, therefore, important to develop an independent variable selection procedure that can be used with neural networks and can potentially extract any nonlinearities present in the data. This is, however, highly computation intensive and will be pursued in future. We intend to carry out our further studies using the NCMRWF T-80 model output, which is now feasible since the model has been operational for several years and sufficient data sample is available.

5. Conclusions

In this paper we have analyzed the atmospheric data from several stations in India in order to determine whether neural networks can be useful in predicting precipitation. We find that in most cases neural network gives small improvement in results in comparison to regression analysis. This small improvement in the neural network results are encouraging, given the fact that a considerable preprocessing had already been done on the data in order to select combinations of predictors which are likely to give optimal results with linear regression. In most cases, however, we did not find much use for hidden layers which are usually required if the output displays a very nonlinear dependence on the inputs. This is somewhat surprising since we would have expected that the observed rainfall would show a very nonlinear dependence on other atmospheric variables. A complete neural network based study, which directly tries to correlate the observed raw atmospheric variables with precipitation may be much more useful in modelling the nonlinearities that are likely to be present in any relationship between these variables. We hope to complete such a study in near future.

References

Bankert, R. L., 1994, "Cloud Classification of a AVHRR Imagery in Maritime Regions Using a Probabilisitic Neural Network", *J. Appl. Meteor.*, **33**, 909-918.

Frankel, D., Schiller, I., Draper, J. S. and Barnes, A. A., 1990, "Investigation of the Prediction of Lightning Strikes Using Neural Networks, Preprints", 16th Conference on Severe Local Storms, Kananaskis Provincial Park, Alberta, Canada, *Amer. Meteor. Soc.*, 7-11.

Kumar, A. and Maini, P., 1996, "Statistical Interpretation of General Circulation Model: A prospect for Automation of Medium Range Local Weather Forecast in India", *Mausam*, **47**, 3, 229-235.

Kumar, A., Maini, P. and Singh, S. V., 1999, "An Operational Model for Forecasting Probability of Precipitation and Yes/No Forecast", *Weather and Forecasting*, **14**, 38-48.

Marzban, C. and Stumpf, G. J., 1995, "A Neural Network for Tornado Prediction Based on Doppler Radar-derived Attributes", *J. Appl. Meteor.*, **35**, 617-626.

Masters, T., 1993, "Practical Neural Network Recipes in C++", Academic Press, p493.

McCann, D. W., 1992, "A Neural Network Short-Term Forecast of Significant Thunderstorms", *Weather and Forecasting*, **7**, 525-534.

Müller, B. and Reinhardt, J., 1991, "Neural Networks: An Introduction", Springer-Verlag: The Physics of Neural Networks Series, p266.

Navone, H. D. and Ceccatto, H. A., 1994, "Predicting Indian monsoon rainfall : a neural network approach", *Climate Dynamics*, **10**, 305-312.