Vol.12 No.3 Advances in Atmospheric Sciences August 1995

Neuroid BP–type Model Applied to the Study of Monthly Rainfall Forecasting①

Yan Shaojin (严绍瑾), *Peng Yongqing* (彭永清) and *Guo guang* (郭 光)
Nanjing Institute of Meteorology, Nanjing 210044

Received November 28, 1994; revised March 20, 1995

ABSTRACT

A neuroid BP–type three–layer mapping model is used for monthly rainfall forecasting in terms of 1946–1985 Nanjing monthly precipitation records as basic sequences and the model has the form $i \times j = 8 \times 3$, $K = 1$; by steadily modifying the weighing coefficient, long–range monthly forecasts for January to December, 1986 are constructed and 1986 month–to–month predictions are made based on, say, the January measurement for February rainfall and so on, with mean absolute error reaching 6,07 and 5,73 mm, respectively. Also, with a different monthly initial value for June through September, 1994, neuroid forecasting is done,indicating the same result of the drought in Nanjing during the summer, an outcome that is in sharp agreement with the observation.

Key words: Neuroid, BP–type three–layer mapping model, Monthly rainfall forecasting

I. INTRODUCTION

Flood and drought are the major threat to mankind. It is thus a compelling duty for meteorologists to investigate their regular pattern and prepare their forecasts. There is a range of forecasting techniques for such disasters. Strong inhomogeneity of rainfall distribution, however, exists both temporally and spatially, leading to greater difficulties. Nevertheless, it is because of their importance that efforts are always directed to the research from different aspects. In the past decade or so rapid progress has been made in nonlinear sciences due to the establishment of theories on dissipative struoture and choas dynamics, the example being artificial neuroid that has spread widely since the 1980s. The neuroid is a large– scale highly nonlinear dynamic system whose major features are to reflect the overall effects and to make large–scale parallel distribution treatment of a bulk of complicated information. The authors have published a paper on applying a neuroid BP multilevel mapping model for monthly mean temperature (Yan et al., 1995).

Flood and drought, being nonlinear phenomena, have a complicated evolution in their regular pattern. For this purpose, the precipitation for the last 50 years is taken under study. The year is defined as that of flood (drought) if the actual rainfall is above (below) the long–term mean by 30 mm and denoted *F*(*D*), the others being normal years and denoted R so that we have the following sequence:

R R D D F R D D F R D F D D D D F F D D D D R F F R R R F D F D D R FF F R D R D ...

One can see therefrom that R may be followed by another R, or one or

---

①This study is supported by the National Natural Science Foundation of China.

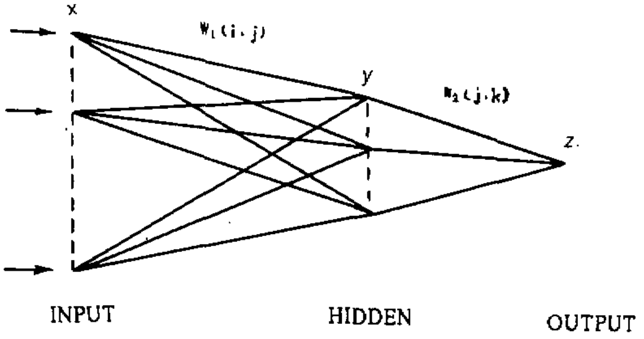

Fig.1. The three–layer neuroid BP model used.

three D'_s; two R'_s by as many D'_s; one F by two or three or even four D'_s, etc. Evidently, we are unable to expound the development only in virtue of a single linear relationship. For this reason, an attempt is made to apply the BP–type model capable of describing nonlinear evolution for predicting monthly rainfall.

II. ARTIFICIAL NEUROID BP MODEL

A neuroid consists of so–called units corresponding to cells that are connected by neuroid joints, each of which is responsible for multiplying the output from a unit by a weighing coefficient equivalent to the connecting intensity at the synapse and then sending the result to another, which, in turn, combines all such results from the neighboring units before output. For actual neural cells, as the intersynapse intensity changes, so does the neuroid behavior while for an artifical neuroid, when the connection weighing coefficient is altered, so is the gross behavior. The BP neuroid method which is also called the inverse transmission model consists in using the least mean variance with which to send a result to a hidden unit at the intermediate level so as to change the weighing coefficients of the matrix and thereby to reach the goal of making a forecast based on the expected input information. As nonlinear prediction, such a scheme remains useful. Fig.1 illustrates a simple three–layer neuroid BP model, where x denotes the input layer, x_i being the i–th component; y the hidden layer, joints totalling in $j (j \leqslant i)$; z the output layer, joints totalling in k (k<j). For convenience we often set $k = 1$ and let $w_1(i,j)$ and $w_2(j,k)$ be the weighing coefficient matrices between the input and hidden levels, and between the hidden and output ones, respectively. The input signal has to be transferred forward to a hidden joint to experience nonlinear effect (usually expressed in terms of the squashing function (SF) before sending the output signal from this point where the signal undergoes such effect again before the result is exported. The SF taken is normally of the s–type, e.g., $f(x) = 1/(1 + e - x)$ which is marked by continuous differentiability, monotonously progressive increase and saturation.

The operation is presented in Fig.1. One can see therefrom that the hidden layer Y_j comes from the input–layer X_i multiplied by the weighing coefficient w_{ij} so that we have

$$Y_j = \sum_i X_i W_{ij} \qquad (1)$$

with the SF of Y_j as

$$\tilde{Y} = \frac{1}{(1+e^{-Y_j})} \ . \tag{2}$$

The signal Z_k from the hidden to the output layer will be obtained via the SF, $-\tilde{y}_i$, multiplied by the weighing coefficient W_{jk}

$$Z_k = \sum_j W_{jk} \tilde{Y}_j \tag{3}$$

and for Z_k its SF is

$$\tilde{Z}_k = \frac{1}{1+e^{-zk}} \ . \tag{4}$$

Now with the expected output set to be d_k, we proceed to correct the weighing coefficient in terms of the backward transfer of squared error. And we let the squared error function be of the form

$$E_k = \frac{1}{2}(\tilde{Z}_k - d_k)^2 \ . \tag{5}$$

We now assume $\triangle w_1$ and $\triangle w_2$ to be the revised values of the coefficients w_{ij} and w_{jk}, respectively, viz.,

$$\triangle W_1 = -\frac{\partial E_k}{\partial W_1(i,j)} \ , \tag{6}$$

$$\triangle W_2 = -\frac{\partial E_k}{\partial W_2(k,j)} \ , \tag{7}$$

and we can have the revised coefficients at the i–th step for the time sequence, i.e.,

$$\Delta W_1(i,j)_n = \sum_{n=1}^{n} \Delta W_1(i,j)_n \ , \tag{8}$$

$$\Delta W_2(j,k)_n = \sum_{n=1}^{n} \Delta W_2(j,k)_n \ , \tag{9}$$

where $n = 1, 2, 3, \cdots N-d$, with N as the length of the series and d as the number of the independent coordinates for describing the system. After $(N-d)$ operations, the N components have been calculated once, a treatment that is called one cycling, and the weighing coefficient adjustment will go on till the requirement is reached, which calls generally for hundreds, thousands or more of such cyclings.

For the $(l+1)$ cyclings, the coefficients have the form

$$W_1(i,j)_{l+1} = W_1(i,j)_l + \eta\Big(\sum_{n=1}^{N-d} -\frac{\partial E_k}{\partial W_{ij}}\Big)_n \ , \tag{10}$$

$$W_2(j,k)_{l+1} = W_2(j,k)_l + \eta\Big(\sum_{n=1}^{N-d} -\frac{\partial E_k}{\partial W_{kj}}\Big)_n \ , \tag{11}$$

where η is referred to as the learning coefficient.

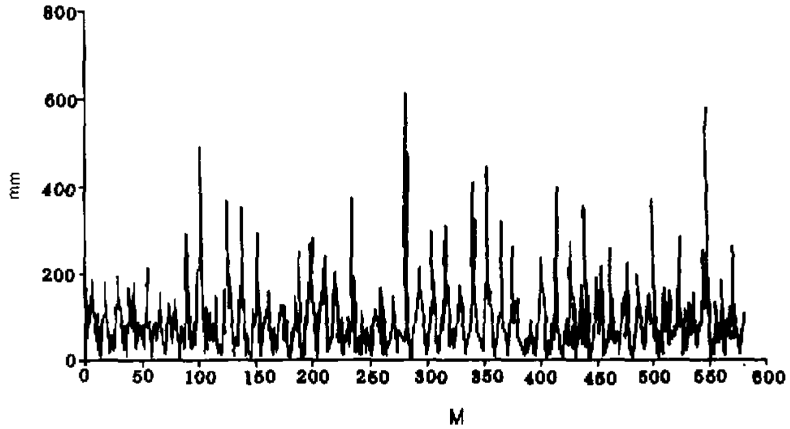

Fig. 2. Temporal sequence of monthly rainfall at Nanjing with O on the abscissa denoting January 1946.

III. NEUROID FORECASTING MONTHLY RAINFALL FOR NANJING

The January 1946–September 1994 data are used in this work, part coming from the compilation of the annual element records (January 1946–December 1988) by China Meteorological Administration and part provided by Jiangsu Institute of Meteorology. Fig. 2 gives the time series of monthly rainfall for the study city.

As shown in Fig.2, strong nonlinearity exists in the series, with the maximum arriving at more than 600 mm (July 1969) and the minimum at O mm (December 1980), a pattern that is so complicated that before establishing the model, we have to process the data to be

$O < x(t) < 1.$

1. *1986 Monthly Rainfall Prediction for Nanjing*

(1) The January 1946–December 1985 monthly rainfalls are taken as the study sample with the length N = 480 serving as the basic data for establishing the model and revising the weighing coefficient to prepare the 1986 rainfall on a monthly basis. Following Yan and Peng (1993), we find the fractal dimension $D_o = 4.8$, with which we get the top and bottom limits, $v_1 = |D_o + 1|$ and $v_2 = |2D_o + 1|$, for the number of degrees of freedom, v, used to describe the dynamic system under investigation. The values are taken at $v = 5$ and $v = 10$ as the first guess of the input component i values, followed by defining j value ($j < i$). As a result, $i \times j = 8 \times 3$ is found to be best based on multiple screenings.

Table 1. Predicted Monthly Rainfall for 1986

	J	F	M	A	M	J	J	A	S	O	N	D
Obs.(mm)	14.9	3.5	44.8	41.3	51.7	199.2	116.1	80.1	34.9	39.4	60.4	36.2
Pred.(mm)	14.0	0.7	44.5	53.9	54.0	205.0	115.5	81.3	8.0	54.6	58.1	34.3
Num.cyc.	225	147	147	145	2	241	145	3	151	2	2	246

(2) The initial values of W_1 and W_2 are given with the aid of a random generator, with

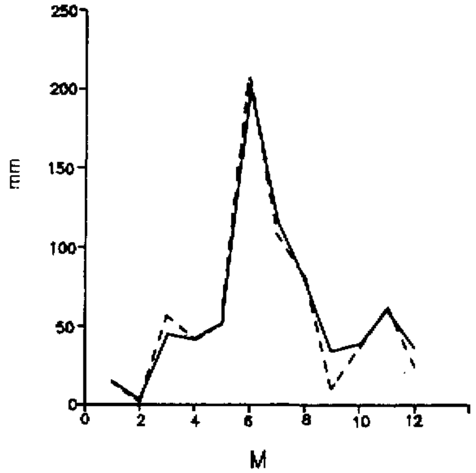

Fig. 3. 1986 monthly rainfall for Nanjing with observations (solid) and model predictions (broken line), based on Table 1.

$(\triangle W_1)n=1=0$, $(\triangle W_2)n=1=0$ and $\eta=0.15$ selected for use, and after hundreds of the coefficient modification it is required that the absolute error be 0.01 between the 480–th output and the real observation. Only with the weighing coefficient do we prepare a forecast for the next monthly rainfall, i.e., for January 1986.

(3) The 481st value is added to the previous 480 values to form a new sequence, i.e., N=481 in this case. The procedure is repeated for the 482nd, 493rd, ···492nd predictions. Thus we have the monthly rainfall for the whole year of 1986 tabulated in the following.

For comparison these results are plotted in Fig.3.

(4) To verify the predictions or to provide corrections for them we proceed to adopt the same BP algorithm to deal with the actual measurement of the previous month for the value of the following one. For example, we use the observation of, say, January 1986 for the prediction of February, etc. Such a procedure is repeated until all the predictions are obtained for all the months of the year, as shown in Table 2.

Table 2. 1986 Monthly Rainfall Predictions for Nanjing, Based on the Measurement of the Preceding Month for the Value of the Following

	J	F	M	A	M	J	J	A	S	O	N	D
Dobs.(mm)	14.9	3.5	44.8	41.3	51.7	199.2	116.1	80.1	34.9	39.4	60.4	36.2
Pred.(mm)	140	2.6	56.2	42.1	53.0	206.7	108.3	80.6	11.3	36.7	60.9	25.2
Num.cyc.	225	306	144	239	2	241	148	3	153	150	2	151

The magnitudes of the above table are plotted in Fig.4.

From Table 2 and Fig.4 one can see that, with the observation of the previous month known, the precipitation can be ferecasted quite accurately for the next month. The latter

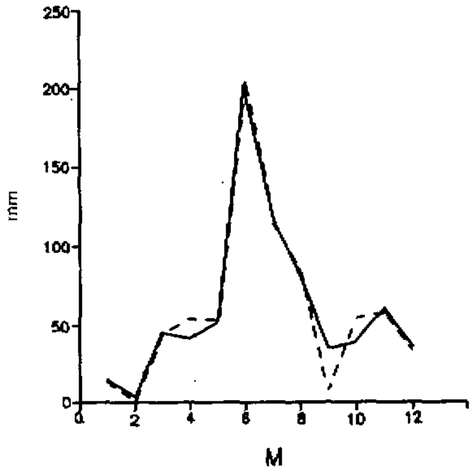

Fig. 4. Nanjing monthly rainfall predictions for 1986 with the measurements (solid) and prediction (broken line) given.

treatment can serve as a verification and provide a correction for the prediction made as well, thereby producing a more successful forecast.

2. *1994 Monthly Rainfall Projecting for Nanjing*

To better demonstrate the usefulness of BP algorithm in forecasting the rainfall, an attempt is made to predict the 1994 anomalous drought event occurring in Nanjing.

It is noted that the usable data ends in September 1994. To predict the drought in June–September we use a range of the sample sizes, with the results shown in the following table:

(1) N = 576 for January 1946 – December 1993;
(2) N = 578 for January 1946 – February 1994;
(3) N = 580 for January 1946 – Aril 1944;
(4) N = 582 for January 1946 – June 1994,

Table 3. Monthly Rainfall Predictions (mm) in January–September 1994 for Nanjing

	J	F	M	A	M	J	J	A	S
Obs.	15.2	32.0	45.2	59.3	87.9	106.1	46.0	83.0	16.0
Ft 1*	18.3	29.2	48.6	73.2	78.8	100.5	25.0	81.0	8.9
Ft 2			43.0	72.3	77.8	101.2	32.3	80.9	3.9
Ft 3					77.4	102.2	30.4	80.6	15.3
Ft 4							34.0	81.1	4.0
69MM**	35.1	48.1	71.3	92.4	93.5	155.9	178.8	111.7	88.5

* Ft 1 denotes forecast No.1 and Ft 2 forecast No.2, etc.

** 69MM denotes 1969 monthly rainfall mean.

One can see from Table 3 that the rainfall predicted for June and August is no more than *2 / 3 of that in normal years and that for July and September is even short of 20% the precipitation* of normal years. This shows that the 1994 summer drought event ferecasting is quite possible and it can be projected prior to the end of 1993.

IV. RESULTS AND DISCUSSION

1. The neuroid BP–type model applied in this work is found useful in calculating the 1986 rainfall on a monthly basis and the summer 1994 drought event. The BP algorithm for monthly precipitation is effective in forecasting the rainfall having time–varying nonlinear feature, quantitatively and qualitatively. And for most of the months the prediction is very close to the observation.

2. The Writers have reported the monthly mean temperature prediction using the neuroid BP model with the correlation coefficient of 0.98, where $i \times j = 3 \times 2$ insomuch as the temperature changes smoothly and the annual cycle is more obvious (Yan et al., 1995). However, change in monthly rainfall is marked by highly strong nonlinearity so that an optimal model of $i \times j = 8 \times 3$ is constructed based on screening, suggesting that intense nonlinearity calls for more state variables. This is not an final conclusion but our own view on thistype of problems.

3. Error analysis is presented of the two types of monthly rainfall predictions for 1986.

Table 4. Error Analysis of 1986 Monthly Rainfall Forecasts

	mean absolute error	mean relative error	mean variance	correlation coefficient
Type 1 *	6.0750	0.0796	9.8823	0.9832
Type 2	5.7333	0.0905	8.8345	0.9879

* Type 1 denotes the long–term prediction based on the data used for 1986.

Type 2 denotes the monthly prediction based on the value of the previous month for the next one of 1986.

4. Averages over the results from Forecasts 1–4 of the monthly rainfall in June to September 1994 are compared to the normal values of 1969 for the city.

Table 5. Comparison of Summer 1994 Rainfall Predictions to 1969 Normal Values (See text)

	June	July	August	Sept.
1969 monthly mean $\bar{R}$	155.9	178.8	117.7	88.5
observation R	106.1	46.0	83.0	16.0
mean of predictions from 4 models F	101.3	30.4	80.9	8.0
$R-\bar{R}$	−49.8	−132.8	−34.7	−72.5
$F-\bar{R}$	−54.6	−148.8	−37.7	−80.5

One can see therefrom that a severe drought event occurred in Nanjing in the summer of 1994.

5. Needed in making flood / drought predictions are the monthly rainfall categories, i.e.,

severe flood, slight wetness, normal, light dryness and terrible drought. Evidently, the results from the BP algorithm are quite good. It is suggested that the 1995 monthly rainfall be predicted in terms of the scheme presented here.

REFERENCES

Jiao Licheng (1990), *Neuroid Theory*, Xi'an Electronic Sci. / Tech. Univ. Press, pp. 34–36 (in Chinese).

Yan Shaojin and Peng Yongqing (1993), *Chaos Theory and Atmospheric Sciences*, China Meteorological Press, pp. 111–116 (in Chinese).

Yan Shaojin, Peng Yongqing and Guo Guang (1995), Monthly Mean Temperature Prediction Based on a Multi–Level Mapping Model of Neural Network BP Type, *Advances in Atmospheric Sciences*, **12**: 225–232.