Comparison of Long–Term Forecasting of June–August Rainfall over Changjiang–Huaihe Valley①

Jin Long (金 龙)
Jiangsu Institute of Meteorology, Nanjing 210008
Luo Ying (罗 莹)
Jiangsu Institute of Applied Climatology, Nanjing 210009
Lin Zhenshan (林振山)
Department of Atmospheric Sciences, Nanjing University, Nanjing 210093
Received November 29, 1995; revised July 10, 1996

ABSTRACT

In terms of an Artificial Neural Network (ANN) established is a long–term prediction model for June–August flood / drought in the Changjiang–Huaihe Basins and a regression forecasting expression is formulated with the aid of the same factors and sample size for comparison. Results show that the ANN is superior in predictions and fittings due to its higher self–adaptive learning recognition and nonlinear mapping especially in the years of severe flood and drought. This shows great promise in using ANN in the research of flood / drought prediction on a long–range basis.

Key words: Artificial neural network (ANN), stepwise regression, Long–term prediction

I. INTRODUCTION

At present, statistical forecasting remains a technique of much importance to the research of flood / drought prediction on a long–range basis, into which are being introduced new theories and methods. The ANN, as a burgeoning cross–discipline, has in recent years been one of the actively pursued frontiers. The ANN, marked by self–adaptive learning, memory and association, etc, has drawn much attention of more than one field of sciences and achieved noteworthy findings, as, for instance, intelligence control, signals processing and model recognition (Fernado and Luis, 1990). Some research on ANN has recently been carried out in the context of atmospheric sciences. The National Severe Storm Forecast Center of U. S. A has put effort into ANN application to thunderstorms in its operational short–term prediction system (Mccann, 1992). Also, good results have been reached in ANN prediction of drought and monthly mean temperature by means of phase–space continuation of one–dimensional time series of a given predictand (Jin, 1996; Yan, 1995). However, is there any difference in the results from the ANN and traditional statistical forecasting? Does the ANN model have higher usefullness in long–range forecasting of flood / drought? These problems need to be studied.

II. BP–TYPE ANN MODEL

The reasons that the version of error Back Propagation (BP) is chosen out of a multitude

①The work is supported by the Natural Science Foundation of Jiangsu Province.

of the ANN models lie largely in that a myriad of parameters are gained through the ANN's progressive learning of the input during its learning / training in addition to factors of learning and momentum selected in advance—these are added to the advantages of the BP type algorithm: sound on a theoretical basis, and good in widespread utilization. This means that rather than follow an artificial pattern, the BP–ANN approximates a certain pattern by extracting information out of fed raw data with the aid of learning.

The BP–ANN fulfils its self–learning through repeated iteration, as is briefly described below (Nielsen, 1989).

Set the learning sample input and expectation output to be denoted by A_k and C_k ($k = 1,2, \ldots$ m), respectively, followed by randomly giving initial association weighing functions V_{hi} from the input to hidden layers of the model used and also such functions W_{ij} from the hidden to output layers, along with thresholds of the hidden– and output–layer elements given as θ_i and γ_i, respectively. With these, calculation is performed of A_k and C_k in the following manner.

1. Based on the association weighing matrix at initial instant that is represented by a group of stochastically given small quantities and the input, we proceed to figure out new excitation values of the hidden–layer elements in terms of the expression

$$b_i = f(\sum_{h=1}^{n} a_h v_{hi} + \theta_i) \ , \tag{1}$$

where $i = 1,2, \ldots, P$ and the excitation function has the form of Sigmoid function:

$$f(x) = 1 / (1 + e^{-x}) \ . \tag{2}$$

2. Find the excitation of the output layer elements using

$$C_i = f(\sum_{i=1}^{p} W_{ij} b_i + \gamma_i) \ , \tag{3}$$

in which $j = 1,2, \ldots, q$ and W_{ij} at initial instant is denoted by a given group of random small quantities.

3. Calculate the generalized error of the output layer elements by means of

$$d_j = c_j(1 - c_j)(c_j^k - c_j) \ , \tag{4}$$

where $j = 1,2, \ldots q$ and C_j^k is expected output of the output element j.

4. Compute the error of hidden elements with respect to each of d_j with the aid of

$$e_i = b_i(1 - b_i) \sum_{j=1}^{q} w_{ij} d_j \ . \tag{5}$$

5. Regulate the association weighing for elements of the hidden layer changed to those of the output layer:

$$\Delta W_{ij} = \alpha b_i d_j \ , \tag{6}$$

in which $i = 1,2, \ldots p$, $j = 1,2, \ldots q$ and α is the learning factor ($0 < \alpha < 1$).

6. Adjust the thresholds of the output layer element through

$$\Delta \gamma_j = \alpha d_j \ , \tag{7}$$

where $j = 1,2, \ldots q$.

7. Regulate the association weighing for elements of the input changed to those in the

hidden layer by

$$\Delta V_{hi} = \beta a_h e_i \quad , \tag{8}$$

where $h = 1,2, \dots n$, $i = 1,2, \dots, p$ and β is the factor of momentum ($0 < \beta < 1$).

8. Adjust the thresholds of hidden elements

$$\Delta\theta_i = \beta e_i \quad . \tag{9}$$

9. Repeat the procedure from (1) through (8) until $j = 1,2, \dots, q$ and $k = 1,2, \dots, m$ with the aim to getting the difference between real and expected output, and training is ended when the output error of the whole sample size is smaller than the set convergence error. In the training course, the BP type has its weighing stepwise-regulated, and gains information from the fed samples. It is easy to see that the ANN-stored knowledge is none other than the association weighing and corresponding critical values of the input and output layers connected by the hidden one. Such values are defined when ANN learning is over. In the light of these parameters and measured data to be fed, prediction can be accomplished only through simple addition and multiplication.

III. STEPWISE REGRESSION AND ANN MODELS FOR FORECASTING JUNE-AUGUST CHANG JINAG-HUAIHE BASIN FLOOD / DROUGHT ON A LONG-RANGE BASIS

The poor normality of single-station rainfall record is taken into account for the sake of comparability in the establishment of ANN and regression models. 25 stations inside the province are sorted out over the Changjiang-Huaihe basins with Huaiyin in the north and Suzhou in the south, with June-August mean precipitation as the predictand for both the schemes. General examination of the correlation between 1952-1992 June-August mean rainfall and previous monthly mean 500 hPa height, leads to 19 good correlation factors, of which 12, as predictors, arrive at >0.01 significance. With the aid of the 19 predictors and the predictand, predictors are singled out for the regression expression by means of a stepwise regression schems. For $F = 2.5$, seven predictors are automatically chosen therefrom. Hence we have

$$\begin{aligned} Y = 2279.4 + 1.8802x_3 + 3.1814x_4 + 1.9841x_6 + 1.6223x_{10} \\ - 2.5060x_{12} + 8.1471x_{14} - 9.7631x_{15} \end{aligned} \tag{10}$$

where the subscript of variable x denotes the serial number of the predictor. Eq.(10) has its multiple correlation coefficient $R = 0.8836$. Significance tests show that at $a = 0.01$ significance level, $F_c = 16.8$ is greater than $F_{a=0.01} = 3.3$, thereby indicating the high enough significance of Eq.(10) is based on 1952-1992 June-August mean rainfall with the 1993 corresponding data as the verification. Afterwards, these selected factors come as the BP network input into making a related ANN model. However, since, in constructing the ANN model, the BP excitation function (2) takes the form of Sigmoid function, a scheme is developed for normalizing predictors and predicatand to meet the needs of function transformation and in the light of actual data. The scheme is

$$K_i = \frac{z_i - s}{t - s} \quad , \tag{11}$$

where K_i is the normalized input into the ANN, z_i the primitive preditors and predictand, s and t the coefficients to be specified. In this study we set their normalizations to be over the range 0.2—0.8 so that s and t can be found by

$$a - s = 0.2(t - s) \ ,$$
$$b - t = 0.8(t - s) \ , \tag{12}$$

where a and b represent a maximum or minimum of the sequences of predictors and predictand. Note that (11) ad (12) are general formulations for normalizing all the sample series.

A table (not shown), summarizes 1952–1992 normalized input matrix into and expected output from the ANN of the seven predictors and predictand obtained by virtue of (11) and (12). All the normalizations are put into a three-layer BP network input terminal for the ANN to have learning / training. The BP possesses seven input nodes in the intermediate hidden layer. The learning factor α is taken as 0.7 and the momentum factor β as 0.9. We preset the convergence error = 0.0003 for training. Operation is repeated following the procedure from (1) to (8). The training is ended as the set convergence error is reached. According to the resulting association weighing functions, thresholds and predictors uninvolved in the training, it is convenient to get the prediction out of the tranined network.

IV. COMPARISON OF RESULTS FROM THE TWO METHODS

In term of the above schemes and the same predictors, a stepwise regression model and an ANN are constructed in the context of 1952–1993 samples (total of 42 samples) and used to make forecast of 1994 June–August rainfall for the project valleys in an attempt to investigate their usefulness and fitting accuracy. At first, comparison is made of the 1993 fittings from the 1952–1992 data (total of 41 samples). One can see from Fig.1 that though based on the identical factors, the ANN is better as regards the fittings because of its higher imitative learning and memorial recognition. Statistics shows that the stepwise regression yields mean relative error of 12.14 and maximum relative error of 45.87% between the fitting and measurement versus 7.23 and 28.83%, respectively, for the BP technique. Further, the regression fittings show 20 cases of relative error > 10% compared to 8 cases for the BP network. Also, Fig.1 shows the particular superiority of the BP type for severe flood and drought events in history and indicates that for the six years of most severe disasters (half for floods), the BP (regression) relative error is 8.67 (24.04%), a little (roughly 100%) higher than the mean of all the samples. In 1991 (1978) as the most severe flood (drought) year, the BP fitting error is 0.33 (6.57%) versus 19.18 (45.87%) for the regression and great difference in fittings for such extreme events occurs to the forecasting error. Table 1 gives the comparison of results from the two models for 1993 and 1994. To make comparison objectively, the predictors, originally used in the 1993 forecast are employed for the 1994 experiment, with the indetical nodes at the three layers and the same convergence error of 0.0003 for the BP model in the two experiments. Table 1 shows the merits of the ANN in both experiments. On the other hand, the regression gives big difference for the 1994 operation (the 3rd case of the most severe drought) though the same multiple correlation coefficient of 0.886 is used for both experiments. This means that such a coefficient of the regression affects the fitting, but it does not necessarily mean the poor quality of prediction. Moreover, we compared the 1993 / 1994 fitting curves for either of the models (figure not shown) and found that their respective trends are almost coincident. This suggests that either of the models can yield very consistent trend for the fitting to follow. Nevertheless, as illustrated in Table 1, great difference is seen in the regression forecasting accuracy whilst the ANN yields steady accuracy both for the anomalous (1993) and normal (1994) year, with much lower mean relative error and absolute error, and the sum of squared residual reduced by 68.8% and 69.6% for 1993 and 1994,

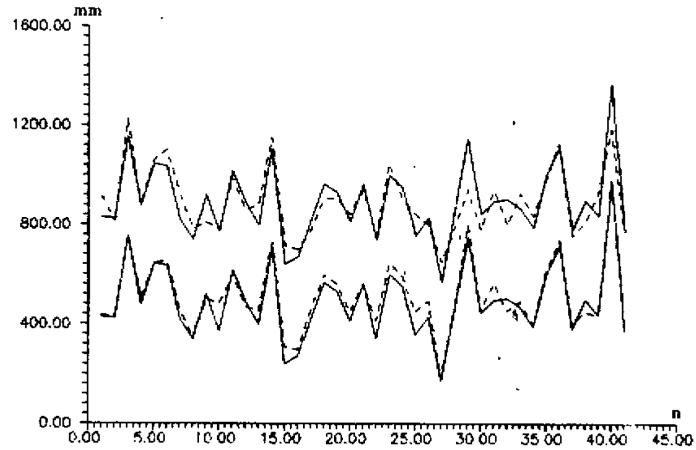

Fig. 1. June–August measured rainfall (full line), ANN fitting (lower broken) and regression fitting (upper broken, whose ordinate has moved 400 upward).

respectively, as compared to the regression. Evidently, such reduction suggests the potential of the ANN technique. The foregoing analysis shows that the ANN is superior in the fittings in the severe cases of flood / drought over the research basins, a concern that awaits further investigation.

For further investigation of the models, with the last introduced predictor (x_6) removed, the other six (cf. Sec. III) are put into the two models, separately, for the 1993 and 1994 experiments. It turns out that the ANN prediction is superior with mean relative error of 14.6% versus 25.8% for the regression, and that the regression forecasting is poor for 1994 severe drought, evidence that agrees with the above investigation.

Table 1. Comparison in the Accuracy of the Two Methods

Year	obser.	pred.	absolute error	relative error(%)	pred.	absolute error	relative error(%)
		ANN				regression	
1993	596.5	566.4	30.1	5.59	603.1	6.6	1.11
1994	270.0	254.9	15.1	5.04	429.6	159.6	59.1
average			22.6	5.32		83.0	30.1

V. CONCLUDING REMARKS

This article shows that ANN, being of higher ability of nonlinear mapping, is capable of recognizing through self-adaptive learning the complicated nonlinear relation between dependent and independent variables in samples and, in particular, it is advantageous scheme in the identification and extension of such nonlinear relation in extreme years, thus providing a new line for the research of long-term forecasting of flood and drought. However, this study is confined to a limited record and so it needs to be verified when more data are available. The authors are making efforts at comparison of the models on time series and ANN techniques in order to investigate the applicability of ANN to long-range forecasting.

REFERENCES

Fernado M. S. and B. A. Luis (1990), Acceleration Techniques for Back Propagation Algorithm, *Neural Network*, **1**: 351–370.

Jin Long, Luo Ying and Yuan Chenzhong (1996), An Important Method to defend and Relieve the Damage of Agrometeorological Drought, Journal of Nanjing University, 32 (Supplement): 115–120 (in Chinese).

Mccann, D. W. (1992), A Neural Network Short–term Forecast of Significant Thunderstorms, Weather and Forecasting, **7**: 525–534.

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

Nielsen, R. H. (1989), Theory of the Back Propagation Neural Network, Proc. of IEEE International Conf. on Netural Network, **1**: 593–605 (Washington D. C.).