# STUDY ON MIXED MODEL OF NEURAL NETWORK FOR FARMLAND FLOOD/DROUGHT PREDICTION*

*Jin Long* (金 龙),

Jiangsu Meteorological Institute, Nanjing 210008

*Luo Ying* (罗 莹),

Jiangsu Institute of Climate Application, Nanjing 210009

*Guo Guang* (郭 光)

Nanjing Institute of Meteorology, Nanjing 210044

and *Lin Zhenshan* (林振山)

Department of Atmospheric Sciences, Nanjing University, Nanjing 210093

Received July 14, 1996; revised October 17, 1996

ABSTRACT

The paper concerns a flood/drought prediction model involving the continuation of time series of a predictand and the physical factors influencing the change of predictand. Attempt is made to construct the model by the neural network scheme for the nonlinear mapping relation based on multi-input and single output. The model is found of steadily higher predictive accuracy by testing the output from one and multiple stepwise predictions against observations and comparing the results to those from a traditional statistical model.

**Key words**: flood/drought prediction, mixed model, nonlinear mapping, soil humidity, neural network

## I. INTRODUCTION

China suffers from monsoon climate-related disasters, especially flood /drought that occur frequently. Therefore, medium- and long-term predictions of wetness/dryness are important research subjects for disaster prevention and reduction in agriculture. It is common practice, however, that rainfall or its departure is used as a predictand in current research and operation at home and abroad (Zhou and Huang 1990). Attempt is made in the paper to construct a model with a predictand based on soil humidity as a comprehensive indicator of water regime of crops aiming at better application to fighting against floods/droughts and management of water resources.

* This study is supported by the Excellent Talent Foundation of the State Education Commission.

## II. PRINCIPLES OF THE MIXED MODEL

### 1. *Structure of the Model for Predicting Floods/Droughts*

At present, statistical techniques remain dominant in operational forecasting of the wetness/dryness on long- and medium-range basis in meteorology. Two methods in widespread use are multivariate analysis and time series analysis. The common model based on regression in multivariate analysis has the form

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon, \tag{1}$$

where $y$ is the predictand, $x_i$ the previous-period predictor, $\beta_j$ the regression coefficient, with $i=j=1, 2, \cdots, m$ and $\varepsilon$ white noise. In another kind of time series analysis, the analysis is made mainly on the predictand itself with reference to "time domain". The common form of the autoregression model AR (P) is given by

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p x_{t-p} + \varepsilon_t, \tag{2}$$

where $y_t$ is the predictand's series, $\beta_i$ the model's coefficient and $\varepsilon_t$ white noise. Model (1) is based on the idea that the future state comes from changes in previous-period pertinent factors while model (2) relates the future state to previous condition of the very predictand. In fact, for generalized series analysis the 1D time series has its future value dependent on the past through a recurrent relation, namely

$$Y_t = F(Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p}). \tag{3}$$

Likewise, generalized regression analysis can be expressed as

$$Y_t = F(X_1, X_2, \cdots, X_m). \tag{4}$$

Models (1) and (2) are now in widespread use for flood/drought prediction. However, analysis of medium-/long-term weather process associated with these calamities shows that since the affecting factors are many and combined in a complicated fashion, no fully deterministic equation can give quantitative description and prediction, and the model may not be good enough to depend only on pertinent exogenous factors and determine their occurrence in advance, which may be parasitic utterly on the previous state. It is hence believed that flood/ drought prediction may be given by a mixed model involving the previous regime and external variables which most likely is not necessarily linear. If noise contained in the new type of model presented is assumed to be inherent in the original series, then we have a model of generalized matrix form

$$Y_t = \Phi(X_i, Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p}), \tag{5}$$

where $X_i$, $Y_{t-1}$, $Y_{t-2}$, $\cdots$, $Y_{t-p}$ are all column vectors. Assuming $\Phi$ to be nonlinear mapping from input to output, we shall now deal with the problems how to approximate Eq. (5) and its possibility to forecast these disasters.

### 2. *Model Establishment and Nonlinear Mapping Realization*

Jin and Luo (1992) indicated that flood/drought affect crops through the content of soil humidity (SH). Study of 1D SH series shows that SH at an instant, $t$, is substantially influenced by the previous regime at $t-1$, $t-2$, etc. And the physical process is understandable, meaning that greater SH at $t-1$, $t-2$, $\cdots$, will produce stronger evaporation, other conditions being equal, resulting in more water lost, and vice versa.

Besides, analysis by the SH balance equation reveals that previous rainfall will directly affect subsequent SH; change in sunshine strength has effect, too, by altering the evaporation denoted by the thermal terms. It is clear from the arguments that change in SH essential to crop growth bears a relation to some meteorological elements as well as the previous SH, a fact that agrees with our considerations in making the mixed model (5) and serves as the basis. The next task is to make model (5) have the function of nonlinear mapping $\Phi$.

The 1980s saw the achievements in the research into information processing in human brain based on neurological progress and the science of computers, and considerable advances have been made in many fields of learninng in application of numerous artificial neural network (ANN) technology for processing large-scale nonlinear parallel distributed message (Lee et al. 1992; Yin and Xu 1991). And the feedforward, or BP, network (Zhang et al. 1993), though having some limitations, is still in wide use, representing a kind of nonlinear mapping from multiple input to single output or a set of output, and, in particular, the mapping is realized without the knowledge of the internal structure of the study system but such a structure can be obtained through the teacher's learning/training by simulating the observed sample with the aid of the BP technique, which provides a theoretically sound basis for preparing the mixed-type model.

## III. FLOOD/DROUGHT PREDICTION WITH VERIFICTAION

### 1. *Learning Algorithms of the Model Establishment for Prediction*

The flood/drought prediction model based on Eq. (5) is a 3-layer BP network, one layer being for input and another for output with the hidden one in between, which utilizes an algorithm of inverse propagation of error for the learning input matrix (Jin et al. 1996) as a means of the network training. The mathematic principle is illustrated as follows.

Let $E_k$ be the cost (i. e., error) function of the input of a training set with respect to $(A_k, C_k)$ at the output layer and we have, for the total model input set, the whole cost function in the form

$$E = \sum_{k=1}^{m} E_k. \tag{6}$$

Thus, for the $k$-th paired input, the weighing input of unit $j$ at the output level takes on the form

$$\mathrm{net}C_j = \sum_{i=1}^{p} w_{ij} b_i, \tag{7}$$

but the actual output of $j$ is given as

$$C_j = f(\mathrm{net}C_j) = f(\sum_{i=1}^{p} w_{ij} b_i). \tag{8}$$

On the other hand, the weighing input of unit $i$ at the hidden level is expressed formally as

$$\mathrm{net}\ b_i = \sum_{h=1}^{n} v_{hi} a_h, \tag{9}$$

with the real output of the unit is in the form

$$b_i = f(\text{net}b_i) = f(\sum_{h=1}^{n} v_{hi} a_h). \tag{10}$$

For unit $j$ of the output layer, error of general character is defined by

$$d_j = -\frac{\partial E_k}{\partial \, \text{net}C_j}, \tag{11}$$

which can be rewritten as

$$d_j = -\frac{\partial E_k}{\partial C_j} f'(\text{net}C_j), \tag{12}$$

and for $i$ at the hidden, we specify its general error occurrence in the form

$$e_i = -\frac{\partial E_k}{\partial \, \text{net}b_i} = -\frac{\partial E_k}{\partial b_i}\frac{\partial b_i}{\partial \, \text{net}b_i} = f'(\text{net } b_i)(-\frac{\partial E_k}{\partial b_i}) = f'(\text{net}b_i)\sum_{j=1}^{q} d_j w_{ij}, \tag{13}$$

where $d_j$ represents the error of the output level, which, as arriving at the hidden through inverse migration, has the expression of $e_i$.

With the aid of the connection weighings $w_{ij}$ and $v_{hi}$ for the input and hidden levels, we employ gradient descending (i. e., the weighings are a direct measure of negative gradients) to minimize the cost function.

To compute the variation in the weighing coefficients we make use of the following

$$\Delta w_{ij} = -\alpha \frac{\partial E_k}{\partial w_{ij}} = \alpha d_j \left[\frac{\partial(\sum_{j=1}^{p} w_{ij} b_i)}{\partial w_{ij}}\right] = \alpha d_j b_i, \tag{14}$$

$$\Delta v_{hi} = -\beta \frac{\partial E_k}{\partial v_{hi}} = -\beta \frac{\partial E_k}{\partial \text{net } b_i}\frac{\partial \text{net } b_i}{\partial v_{hi}} = \beta e_i a_h, \tag{15}$$

where $\alpha$ $(0<\alpha<1)$ and $\beta$ $(0<\beta<1)$ stand for a learning and a momentum factor, respectively.

In view of the fact that the total cost function $E$ is defined for the input training set of the model as a whole so that to realize its actual gradient descending across a curved surface it is necessary to find negative gradient for the weighings, viz.,

$$-\frac{\partial E}{\partial w_{ij}} = \sum_{k=1}^{m}(-\frac{\partial E_k}{\partial w_{ij}}), \tag{16}$$

$$-\frac{\partial E}{\partial v_{hi}} = \sum_{k=1}^{m}(-\frac{\partial E_k}{\partial v_{hi}}), \tag{17}$$

in which case the related expressions for the weighing variation are given as

$$\Delta w_{ij} = -\alpha \frac{\partial E}{\partial w_{ij}} = \sum_{k=1}^{m}(-\alpha \frac{\partial E_k}{\partial w_{ij}}), \tag{18}$$

$$\Delta v_{hi} = -\beta \frac{\partial E}{\partial v_{hi}} = \sum_{k=1}^{m}(-\beta \frac{\partial E_k}{\partial v_{hi}}). \tag{19}$$

And if $E$ is defined as the square sum of error, namely,

$$E = \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{q}(C_j^k - C_j)^2, \tag{20}$$

$$E_k = \frac{1}{2}\sum_{j=1}^{q}(C_j^k - C_j)^2, \tag{21}$$

with $f$ denoting Sigmoid function, leading to

$$y = f(x) = \frac{1}{1+e^{-x}}, \tag{22}$$

then we obtain

$$f'(x) = \frac{-1}{(1+e^{-x})^2}e^{-x}(-1) = \frac{1}{1+e^{-x}}(1-\frac{1}{1+e^{-x}}) = y(1-y). \tag{23}$$

From Eqs. (12) and (13) we come to the expressions of error calculation for both levels, respectively,

$$d_j = -\frac{\partial E_k}{\partial c_j}f'(\mathrm{net}c_j) = (c_j^k - c_j)f(\mathrm{net}c_j)[1 - f(\mathrm{net}c_j)] = c_j(1-c_j)(c_j^k - c_j), \tag{24}$$

$$e_i = f'(\mathrm{net}\ b_i)\sum_{j=1}^{q} d_j w_{ij} = b_i(1-b_i)\sum_{j=1}^{q} d_j w_{ij}. \tag{25}$$

And from Eqs. (14) and (15) we get, respectively, the equations for weighing's change,

$$\Delta w_{ij} = \alpha b_i d_j, \tag{26}$$

$$\Delta v_{hi} = \beta a_h e_i. \tag{27}$$

Hence, the procedures for the BP network used come down to the following:

Under the assumption that the input pair of the training sample set are $A_k$ and $B_k$ ($k = 1, 2, \cdots, m$), $v_{hi}$ and $w_{ij}$ are given randomly as the initial connection weighings, separately, for the input-hidden and the hidden-output levels, and $\theta_i$ and $\gamma_j$ as the threshold values of the latter two layers, respectively, we are allowed to deal with the calculation of $A_k$ and $B_k$, as shown below.

(1) Following a learning sample input and the connection weighing matrix, a new magnitude of the hidden activation is calculated through

$$b_i = f(\sum_{h=1}^{n} a_h v_{hi} + \theta_i), \tag{28}$$

where $i = 1, 2, \cdots, p$, $v_{hi}$ is a group of small stochastic quantities given initially and the Sigmoid function (22) used as the activation function.

(2) We have to compute the Sigmoid function for the output-level units

$$c_j = f(\sum_{i=1}^{p} b_i w_{ij} + \gamma_i), \tag{29}$$

with $j = 1, 2, \cdots, q$ and initial $w_{ij}$ denoted by a group of small stochastic quantities.

(3) We can have general errors of the output-level units from Eq. (24) where $j = 1, 2, \cdots, q$ and $C_j^k$ is the expected output value of $j$ in this layer.

(4) Errors are found of units at the hidden level with respect to those of each $d_j$ with $i = 1, 2, \cdots, p$, based on Eq. (25).

(5) Adjustment is made of the connection weighings of the units from the hidden to output layers in terms of Eq. (26) with $i = 1, 2, \cdots, p$; $j = 1, 2, \cdots, q$; $\alpha$ is the learning factor ($0 < \alpha < 1$).

(6) Adjustment of the threshold values of the output units

$$\Delta\gamma_j = \alpha d_j. \tag{30}$$

(7) The use of Eq. (27) allows to make adjustment of the weighings for units from the input to hidden layers, where $h = 1, 2, \cdots, n$; $i = 1, 2, \cdots, p$ and $\beta$ is the momentum factor ($0 < \beta < 1$).

(8) Then we adjust the threshold values of the hidden units in virtue of

$$\Delta\theta_i = \beta e_i. \tag{31}$$

The repetition of the procedures (1) — (8) is carried out through $j=1, 2, \cdots, q$ and $k=1,2,\cdots, m$ till a minimum $d_j$ is reached, suggestive of the end of learning. In this way, a self-adaptation learning system is established for the network, which has now the ability to remember and conceive the samples that are fed.

## 2. *Scheme for Establishing the Model Learning Matrix*

The decade SH is taken as the predictand for Eq. (5) in the context of SH averaged over 0—50 cm soil for January 1992 to March 1995 (total of 117 decades) from Xuzhou Agrometeorological Station, Jiangsu with the subsequent data used for test. Continuation is made of the 1D time series scalar to construct part of the data for preparing model (5). The observed SH 1D sequence has the form

$$Y(t_1),\ Y(t_2),\cdots,Y(t_n), \tag{32}$$

which is then extended into a multi-dimensional series by means of the time lag coefficient $\tau$. We thus get

$$\begin{cases} Y(t_1),\ Y(t_2),\cdots,\ Y(t_n), \\ Y(t_1+\tau),\ Y(t_2+\tau),\cdots,Y(t_n+\tau), \\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots, \\ Y(t_1+k\tau),\ Y(t_2+k\tau),\cdots,\ Y(t_n+k\tau), \end{cases} \tag{33}$$

the last of which is taken as the column vector on the left-hand side of Eq. (5), the others as those on its right-hand side. In establishing the model the 1D SH data are extended into 6D series, leading to the fed matrix-form sample size $N=112$. Further, following the considerations in Section II, the mid February 1992—mid March 1995 decade mean rainfall and sunshine duration ($N=112$, one decade ahead of the SH data) from the station are taken as another two column vectors on the right-hand side of Eq. (5). Thus we have a BP-type learning matrix with $X_1$, $X_2$, $Y_{t-5}$, $Y_{t-4}$, $\cdots$, $Y_{t-1}$ as input and $Y_t$ as expected output.

To meet the conditions of Sigmoid function, the constructed learning matrix after it has been normalized is put on the input points of the BP network (7 input nodes and 1 output node). After adjusting the model the last two decades (early to mid April 1995) of the sample size are predicted for SH with the analysis indicating that, with the learning (momentum) factor taken as 0. 7 (0. 9), the number of hidden nodes as 9 and convergence error of 0. 003, the forecasting error is kept steadily small. On this basis, further experiments are made with these parameter's values.

## 3. *Analysis of Predictions*

To objectively investigate the predictive ability of the established model, 20 decade SH predictions were made for spring to summer 1995 with the model and parameter values, comparison to observations followed. Error analysis from one-step prediction (one decade in advance) in Table 1 shows the effectiveness to be satisfactory. For the 20 forecasts, the maximum (minimum) relative error (RE) is 26. 4% (0. 5%), with the mean of 8. 03%. If the relative error (prediction/observation) $<10\%$ is taken as a criterion of successful forecast, the accuracy reaches 70% and for the corresponding climatic predictions (only the mean is given), the probability of prediction failure amounts to 70%. By use of the expression for residual sum of squares

$$SSE = \sum_{j=1}^{m}(y_j - \hat{y}_j)^2, \tag{34}$$

further calculation can be made of

$$C_1 = \frac{SSE_1 - SSE_2}{SSE_1}, \tag{35}$$

to find out the difference in accuracy between the climatic and model (5) predictions. In Eq. (35) $SSE_1$ and $SSE_2$ are the residual sum of squares for 20 climatic forecasts and 20 model (5) predictions, respectively so that

$$C_1 = \frac{438.50 - 104.74}{438.50} = 76.1\%, \tag{36}$$

which indicates that the accuracy of model (5) results is 76. 1% higher than that of its counterparts.

**Table 1.** Verification of ANN-Produced SH Forecasts

| No. | Observation | one-step | | two-step | | three-step | |
|---|---|---|---|---|---|---|---|
| | | Prediction | RE* | Prediction | RE | Prediction | RE |
| 1 | 19.74 | 19.37 | 0.019 | 18.98 | 0.039 | 19.04 | 0.36 |
| 2 | 18.63 | 19.76 | −0.061 | 19.63 | −0.054 | 19.58 | −0.052 |
| 3 | 15.42 | 19.49 | −0.264 | 19.53 | −0.267 | 19.38 | −0.257 |
| 4 | 16.00 | 16.61 | −0.038 | 16.91 | −0.057 | 16.85 | −0.053 |
| 5 | 15.98 | 16.10 | −0.008 | 16.26 | −0.018 | 16.27 | −0.018 |
| 6 | 14.98 | 14.89 | 0.006 | 14.48 | 0.034 | 14.30 | 0.045 |
| 7 | 16.62 | 13.66 | 0.178 | 13.76 | 0.172 | 14.32 | 0.139 |
| 8 | 16.64 | 16.19 | 0.027 | 15.78 | 0.052 | 15.76 | 0.053 |
| 9 | 15.70 | 15.82 | −0.008 | 15.69 | 0.000 | 15.42 | 0.018 |
| 10 | 19.84 | 16.30 | 0.179 | 16.24 | 0.181 | 16.06 | 0.190 |
| 11 | 16.02 | 18.51 | −0.155 | 19.32 | −0.206 | 19.12 | −0.194 |
| 12 | 20.82 | 19.45 | 0.066 | 18.43 | 0.115 | 17.33 | 0.168 |
| 13 | 20.62 | 22.40 | −0.086 | 22.75 | −0.103 | 22.92 | −0.112 |
| 14 | 25.72 | 19.68 | 0.235 | 18.87 | 0.266 | 21.35 | 0.170 |
| 15 | 26.50 | 27.91 | −0.053 | 27.58 | −0.041 | 27.91 | −0.053 |
| 16 | 29.62 | 27.56 | 0.070 | 27.72 | 0.064 | 24.77 | 0.164 |
| 17 | 27.54 | 27.68 | −0.005 | 27.56 | −0.001 | 27.72 | −0.007 |
| 18 | 26.36 | 25.84 | 0.019 | 25.80 | 0.021 | 25.32 | 0.040 |
| 19 | 26.20 | 24.54 | 0.063 | 24.73 | 0.056 | 24.73 | 0.056 |
| 20 | 21.50 | 24.29 | −0.130 | 23.13 | −0.076 | 22.73 | −0.057 |
| mean | | | 0.083 | | 0.091 | | 0.094 |

* RE means relation error

For the experimental predictions with tests, the 20 decade-to-decade forecasts were carried out in the identical conditions of the model structure, input, output, number of hidden nodes and convergence error (which are similar to those for operational forecasting) resulting in applicability of all the forecasts as suggested in Table 1 as regards the error variation. To make further examination of the prognostic model, we made predictions at 2 and 3 decades in advance (or two- and three-stepwise), respectively, for the 20 forecasts (see the corresponding parts of Table 1). Error analyses show that their accuracy is equally desirable and the mean relative error (MRE) is 9.1% (9.4%) for the two- (three-) stepwise predictions, very close to the one-stepwise result, a fact that offers basis for preparing better consensus forecasts.

In addition, since the presented model of mixed type is an attempt and so is the realization of nonlinear mapping from input to output with the aid of the BP network, to examine if the calculation scheme is superior to the traditional one becomes an important aspect in verifying the quality of the developed model. For this reason, for the input matrix of the BP network, the expected output is taken as a predictand and the other 7 columns as predictors to establish a prognostic model regressively, the first equation is then obtained

$$\hat{y}_t = 0.123 - 0.0413x_1 - 0.0111x_2 - 0.0373x_3 - 0.0996x_4 - 0.1479x_5 - 0.1237x_6 - 0.6871x_7, \tag{37}$$

with $N = 112$ and the complex correlation coefficient of 0.8284. On this basis the prediction is made for the next decade ($N=113$). To objectively compare, forecasts and tests were conducted in a similar way to that in Michaels and Gerzoff (1984). That is, after a forecast is made for the subsequent decade by the newly-formed prognostic equation, observations for this decade are put into the sample set to build up another subsequent decade and so on, altogether 20 equations are constructed, these are used for the period in spring—summer 1995, with the accuracy shown in Table 2. Error analyses indicate the steady variation in the complex correlation coefficient ($R$) of these

**Table 2.** Test of Regression-Derived SH Forecasts

| No. | Observation | Prediction | RE* | CC** | $n$ |
|---|---|---|---|---|---|
| 1 | 19.74 | 20.12 | −0.019 | 0.8284 | 110 |
| 2 | 18.62 | 20.17 | −0.083 | 0.8287 | 111 |
| 3 | 15.42 | 19.19 | −0.245 | 0.8288 | 112 |
| 4 | 16.00 | 16.74 | −0.046 | 0.8279 | 113 |
| 5 | 15.98 | 16.67 | −0.043 | 0.8310 | 114 |
| 6 | 14.98 | 16.33 | −0.090 | 0.8339 | 115 |
| 7 | 16.62 | 15.95 | 0.041 | 0.8374 | 116 |
| 8 | 16.64 | 17.16 | −0.031 | 0.8394 | 117 |
| 9 | 15.70 | 17.22 | −0.097 | 0.8413 | 118 |
| 10 | 19.84 | 16.85 | 0.151 | 0.8434 | 119 |
| 11 | 16.02 | 19.53 | −0.219 | 0.8408 | 120 |
| 12 | 20.82 | 17.15 | 0.176 | 0.8397 | 121 |
| 13 | 20.62 | 20.54 | 0.004 | 0.8358 | 122 |
| 14 | 25.72 | 20.49 | 0.203 | 0.8358 | 123 |
| 15 | 26.50 | 24.61 | 0.071 | 0.8303 | 124 |
| 16 | 29.62 | 25.79 | 0.129 | 0.8320 | 125 |
| 17 | 27.54 | 28.78 | −0.045 | 0.8346 | 126 |
| 18 | 26.36 | 28.09 | −0.065 | 0.8377 | 127 |
| 19 | 26.20 | 26.26 | −0.002 | 0.8391 | 128 |
| 20 | 21.50 | 25.16 | −0.170 | 0.8411 | 129 |
| mean | | | 0.097 | | |

* RE=relation error; ** CC=correlation coefficient

prognostic equations because of a long sample length for the regression formulae. Test of the regressions shows higher significance level. Inspection of Tables 1 and 2 as regards the corresponding errors reveals the considerable advantage of ANN-yielded accuracy over the regression-given accuracy (relative error of 8.3% versus 9.7%). Also, it is found that for all the experimental forecasts the accuracy for the predicted maxima and minima is a lot higher. Thus, our experiment provides a new way to raise accuracy, a hard nut to crack for traditional statistical forecasting.

**Table 3.** Verification of Issued SH Forecasts for the Sowing and Growth Stages of Winter Wheat

| Prediction for the sowing (October 1995) | | | | | |
|---|---|---|---|---|---|
| | | ANN | | Regression | |
| Time span | Observation | Prediction | RE* | Prediction | RE* |
| 1st dec. | 25.14 | 24.14 | 0.040 | 21.83 | 0.131 |
| 2nd dec. | 24.22 | 25.10 | −0.036 | 28.39 | 0.034 |
| mean | | | 0.038 | | 0.825 |
| **Prediction for the growth (January—May, 1996)** | | | | | |
| | | ANN | | Regression | |
| Time span | Observation | Prediction | RE* | Prediction | RE* |
| 1st dec. Jan. | 20.78 | 18.67 | −0.1015 | 19.84 | −0.0452 |
| 2nd dec. Jan. | 19.16 | 19.16 | 0.0000 | 20.49 | 0.0695 |
| 3rd dec. Jan. | 20.12 | 20.25 | 0.0067 | 20.05 | −0.0034 |
| 1st dec. Feb. | 18.02 | 19.09 | 0.0596 | 20.09 | 0.1148 |
| 2nd dec. Feb. | 17.94 | 17.94 | 0.0002 | 18.74 | 0.0446 |
| 3rd dec. Feb. | 17.38 | 17.58 | 0.0116 | 18.52 | 0.0655 |
| 1st dec. Mar. | 19.68 | 18.42 | −0.0640 | 17.73 | −0.0992 |
| 2nd dec. Mar. | 19.32 | 18.58 | −0.0381 | 19.14 | −0.0093 |
| 3rd dec. Mar. | 21.56 | 19.70 | −0.0864 | 19.68 | −0.0874 |
| 1st dec. Apr. | 15.58 | 22.14 | 0.4209 | 21.42 | 0.3749 |
| 2nd dec. Apr. | 14.58 | 18.23 | 0.2502 | 17.90 | 0.2276 |
| 3rd dec. Apr. | 16.10 | 14.36 | −0.1084 | 15.46 | −0.0398 |
| 1st dec. May | 16.52 | 16.23 | −0.0176 | 16.31 | −0.0128 |
| 2nd dec. May | 15.66 | 15.69 | 0.0021 | 16.95 | 0.0825 |
| 3rd dec. May | 14.92 | 16.27 | 0.0907 | 16.61 | 0.1132 |
| mean | | | 0.8386 | | 0.9266 |

* RE means relation error

To compare the accuracy from the two methods we use (35) to calculate

$$C_2 = \frac{SSE_3 - SSE_2}{SSE_3} = \frac{120.88 - 104.74}{120.88} = 13.3\%, \tag{38}$$

where $SSE_3$ denotes the residual sum of squares for the regression scheme. Results show that the accuracy given by the developed model is 13.3% higher than that from the regression forecasts due mainly to the fact that the study system is not definitely a linear entity. Considered in the mixed model is not just the evolution of the system itself but the role of external physical processes as well and these integrated effects are revealed through nonlinear mapping, which is able to better describe the substantial relation inside than the linear technique. With the mixed model a SH forecast was made for the decades of October

1995 (sowing period for winter wheat), and the forecast soil humidity for winter wheat growth (January through May, 1996) was good (Table 3), with the ANN prediction for sowing of 3. 8% for mean relative error (prediction minus measurement) compared to corresponding 8. 28% given by the regression for the first two decades of October, 1995. This provided a sound basis for farmers in their action of supplying water for the sowed crop. Also, the January—May 1996 (total of 15 decades) humidity predictions show that the ANN approach yielded, on average, relative error of 8. 38% versus 9. 26% from the regression, a result that was close to that from experiments with 20 independent samples (8. 3% vs 9. 7%). This demonstrates that the developed model can provide an effective service for drought prevention/combat and water resource management in agricultural operations, the model can also give more stable and advantageous dryness/wetness prognosis as well. Nonetheless, the present study is confined to the comparison of results between ANN and regression techniques, and further exploration is needed to reveal the difference between ANN and EDA (exploration date analysis) within a nonlinear framework.

## IV. CONCLUDING REMARKS

Because of the length limitation of the continuous SH measurements, no calculation is performed of the chaos indices (e. g., incidence dimension and Lyapunov exponent) in designing the prognostic model so that no study was made of the dimension of its "attractive system" and no fuller arguments are given in defining the dimension for model establishment. Evidently, a high-quality model depends substantially on all of the considered factors related to the predictand's future state. For this reason, rainfall and insolation duration are included in model (5). Will the forecast accuracy be improved if other contributing meteorological factors, dynamic and thermal, are involved? This is a problem that remains to be investigated.

## REFERENCES


Jin Long and Luo Ying (1992). Study of a computation method for drought index and its application. *J. Nanjing Inst. Meteor.*, **15**: 584—590 (in Chinese).

Jin Long, Luo Ying and Yuan Chengsong (1996). An important method to flight and relieve agrometeorological drought damages. *J. Nanjing Univ.*, **32**: 149—155 (in Chinese).

Lee, K. Y., Cha, Y. T. and Park, J. H. (1992). Short-term load forecasting using the artificial neural networks. *IEEE Transactions on Power System*, **7**: 124—132.

Michaels, P. J. and Gerzoff, R. B. (1984). Statistical relations between summer thunderstorm patterns and continental midtropospheric heights. *Mon. Wea. Rev.*, **112**: 778—789.

Yin Y. Y., Xu X. M., (1991). Applying neural network technology for multiobjective land use planning. *J. Environment Management*, **32**: 319—356.

Zhang, S. P., Watanable, H. and Yamada, R. (1993). Prediction of daily water demands by neural networks. *Proc. International Conference on Statistical Methods in Hydrology and Environmental Engineering held at Ontario*, Canada, pp. 217—227.

Zhou Jiabin and Huang Jiayou (1990). Status of forecasting methods of drought and flood. *The Climatological Research of Drought and Flood*. China Meteor. Press, Beijing, pp. 134 — 142 (in Chinese).