# EXPERIMENTS ON DOUBLE DELAYED EQUATIONS OF PRECIPITATION FORECAST*

*LIN Zhenshan* (林振山)

Department of Physical Geography, Nanjing Normal University, Nanjing 210097

and *BAO Ming* (鲍 名)

Department of Atmospheric Sciences, Nanjing University, Nanjing 210093

Received June 22, 1998; revised September 4, 2000

ABSTRACT

In this article, the delayed equation of precipitation forecast has been raised, and experiments have been made on the monthly mean precipitation in the stations of Nenjiang and Shanghai. The results show that after being debugged, the anomalous symbol of precipitation can reach a forecasting accuracy of 67%−73% with the average relative error less than 15%.

**Key words**: delayed equation, precipitation, forecast

## I. INTRODUCTION

Climatic forecast is an important branch of atmospheric sciences, concerning the national economy and people's livelihood. In the past 10 years, floods and droughts in China and many other countries have brought great losses of billions of US dollars, which has called much attention from governments and meteorological departments around the world. At present common measures taken to forecast climates include dynamic forecast, statistic forecast and statistic-dynamic forecast.

Dynamic forecast costs a lot in calculation, and its present accuracy in forecast can not fulfill the requirement of professional forecast, especially precipitation forecast. Statistic-dynamic forecast still stays in experimental stage. Traditional statistic forecast, usually lays more emphases on data handling completely based on statistics, such as anomaly and variance, neglects all the dynamic factors in the system, which is the biggest problem so that dynamic characters of the system can not be described or revealed.

The development of atmospheric sciences will certainly be influenced by the development of other natural sciences, especially mathematics and physics. While nonlinear science has made great achievements, dynamics theories, statistical principles and nonlinear scientific theories should be combined organically to open a new way to climatic forecast. In this article it will be mainly discussed how to make precipitation forecast with the delayed equation in nonlinear theory.

* This research is supported by the Ministry of Science and Technology of China under grant NKPDBS G199804900-1.

## II. DELAYED EQUATION

It is known that there are three kinds of observed nonlinear dynamic systems with chaotic rule (see Lin 1993): (1) autonomous system of equations with more than three variables; (2) non-autonomous system of equations with more than two variables; and (3) delayed equation with one or more variables. Non-autonomous system of equations will become autonomous equations after a new variable is led into the time item. Delayed operation $x(t-\tau)$ in the delayed equation will be changed into differential operator with infinite steps:

$$x(t-\tau)=\sum_{n=0}^{\infty}\frac{(-\tau)^n}{n!}\frac{\mathrm{d}^n}{\mathrm{d}t^n}x(t). \tag{1}$$

Then the delayed equation will become an autonomous equation with infinite steps.

Equation (1) shows that a simple delayed equation can describe a complicated nonlinear dynamic system. It also reveals that a delayed equation may equal an autonomous system of differential equations with three variables. A lot of research (see Lin 1992) has made it clear that the fractal dimension of attractors in weather-climatic system usually remains from 4 to 7 dimensions, no more than 9 dimensions. Therefore, at most three delayed equations can describe the evolutionary character of this system. If the parameter or the function is chosen well, usually only one or two delayed equations are needed to forecast some dynamic rules, which is the advantage of using delayed equations to conduct forecast. Besides, because: (1) climatic system is a nonlinear system; (2) the observational data of such climatic factors as precipitation and temperature are scattered; (3) draughts or floods usually have effect of time delay, theoretically it is better to use delayed equations to forecast precipitation than continuous system of dynamic equations and statistical method. As a theoretical research with common use, the following delayed models are presented:

(1) single delayed equation

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t}=f[x_i(t),x_i(t-\tau)]. \qquad i=1,2,\cdots,n \tag{2}$$

(2) double delayed equations

If $X(t)=x_i(t)$, $Y(t)=x_i(t-\tau)$, then the following double delayed equations are formed:

$$\begin{cases}\dfrac{\mathrm{d}X(t)}{\mathrm{d}t}=f_1[Y(t),X(t)],\\ \dfrac{\mathrm{d}Y(t)}{\mathrm{d}t}=f_2[X(t),Y(t)].\end{cases} \tag{3}$$

(3) triple delay equations

Make $X(t)=x_i(t)$, $Y(t)=x_i(t-\tau)$, $Z(t)=x_i(t-2\tau)$, then the following triple delayed equations will be written as

$$\begin{cases}\dfrac{\mathrm{d}X(t)}{\mathrm{d}t}=f_1[X(t),Y(t),Z(t)],\\ \dfrac{\mathrm{d}Y(t)}{dt}=f_2[X(t),Y(t),Z(t)],\\ \dfrac{\mathrm{d}Z(t)}{dt}=f_3[X(t),Y(t),Z(t)].\end{cases} \tag{4}$$

The above $x_i$s $(i=1,2,\cdots,n)$ represent the time sequence of some climatic factors while the functional form of $f$ and $f_i$ $(i = 1, 2, 3)$ is only decided by data, people and experience, which is the most serious shortcoming in delayed equation shared by traditional statistical methods (such as regression and auto-regression methods). However, the above form can not be found in any traditional statistical methods (e. g. auto-regression), so delay equations contain the advantages of dynamics and statistical principles.

## III. DETERMINATION OF FUNCTIONAL FORM AND PARAMETER

It is known that when a dynamic model is built upon a physical process to describe the evolutionary character of the system, whether it is successful is determined by the fact that whether the forecast of the model is consistent in theories, experiments and observed evolutionary rules, or not. When forecasting the state of the system, because of the objective existence of the long-term forecast error, the accurate forecast to the system can not be expected. Therefore, the accuracy can only be judged from geometric or statistical criterion.

While delayed forecast model is being set up, it should be taken into consideration whether to set up a single delayed equation, or double delayed equations, or triple delayed equations. There are two ways: (1) if the length of data is above 780, then compute the fractal dimension of the climate system corresponding to the data; if the fractal dimension $D \leqslant 6$, then choose double delayed equations, or triple delayed equations; (2) according to adjusted model and experiments (choose the one with good forecast effect), determine the equation numbers of model.

The second step to set up the model is to choose the functional form and to determine the parameters. At this moment the minimum variance criterion can be used:

$$\sigma_i^2 = \frac{1}{N}\sum_{i=1}^{N}[x_i - X_i]^2, \tag{5}$$

where $x_i$ is the date number, and $X_i$ is forecast output.

Or use the criterion of root mean square

$$r_i = \left[\frac{\sum(x_i - X_i)^2}{\sum x_i^2}\right]^{1/2} \tag{6}$$

to determine the functional form. Because priority is given to drum dynamic forecast (only one step is forecasted, and take the latest data as the latest reference state), only several $(3-5)$ forecast numbers are needed to test. After the determination of the functional form, in order to improve the accuracy of forecast, usually one or two adjustable parameters are set in the function $f$, and the parameter numbers are determinated by the least square method.

In principle the functional form is selected mainly according to the "form" of outputs and test numbers, in stead of accuracy that will be handled through the set of parameters.

Since during practical forecast, the length of many precipitation data is hardly as long as $900-780$ (except dekad precipitation data), there are some difficulties in computing fractal dimension. However, the basis of the available fractal dimension of precipitation system with a time scale of dekad, month and year should be less than 6. Therefore, double delayed equations should be taken.

Make experiments on Nenjiang monthly mean precipitation data, and take tanh $(x)$ as the functional form. The forecasting equations are

$$\begin{cases} \dfrac{dX}{dt} = -X(t) + c_1 f[a_1 Y(t-\tau)] - c_2 f[a_2 X(t-\tau)], \\ \dfrac{dY}{dt} = -Y(t) + c_3 f[a_3 X(t-\tau)] - c_4 f[a_4 X(t-\tau)]. \end{cases} \tag{7}$$

Now use two-step Runge-Kutta method to work out the above equations. In order to do this, make all the right items in Eq. (7) as follows:

$$X_1 = -X(t) + c_1 f[a_1 Y(t-\tau)] - c_2 f[a_2 X(t-\tau)], \tag{8}$$

$$Y_1 = -Y(t) + c_3 f[a_3 X(t-\tau)] - c_4 f[a_4 X(t-\tau)]. \tag{9}$$

Then add a step length of $h$ to the two Eqs. (8) — (9), and the number is:

$$\begin{cases} X_2 = -[X(t) + hX_1] + c_1 f[a_1 Y(t-\tau+h)] - c_2 f[a_2 X(t-\tau+h)], \\ Y_2 = -[Y(t) + hY_1] + c_3 f[a_3 X(t-\tau+h)] - c_4 f[a_4 Y(t-\tau+h)]. \end{cases} \tag{10}$$

According to Runge-Kutta method, get the following difference equation:

$$\begin{cases} X(t+h) = X(t) + h\dfrac{X_1 + X_2}{2}, \\ Y(t+h) = Y(t) + h\dfrac{Y_1 + Y_2}{2}. \end{cases} \tag{11}$$

The key to carry out forecast through Eq. (11) is the chosen number of all the parameters. In Eq. (7) the chosen number of $a_i$ $(i = 1, 2, 3, 4)$ should try to make the chosen number area of the strain number in $f$ between 0.3 and 3.0 to avoid the linear part $(x \in 0, 0.3)$ and the part of constant 1 $(x \in 3.0, \infty)$ in tanh $(x)$. Set $R_{max}$ as the maximum monthly mean precipitation and $R_{min}$ as the minimum monthly mean precipitation in the taken data, and get

$$a_1 = a_4 \frac{3 \div 1.5}{R_{max}} = \frac{2}{R_{max}},$$

$$a_2 = a_3 = \frac{0.3 \times 1.5}{R_{min}} = \frac{0.45}{R_{min}}. \tag{12}$$

When the numbers of parameter $a_i$ $(i=1, 2, 3, 4)$ are available from Eq. (12), basically most $x$'s numbers in $f(x)$ are ranging from 0.3 to 3.0, it is acceptable that there are a few $x$'s numbers beyond the range of 0.3 — 3.0 because th $(x < 0.3)$ and th $(x > 3.0)$ are meaningful (but if most numbers of th $(x)$ are constant, then it is not worth testing). For example, in the light of Nenjiang's monthly mean precipitation of April, May, June, July (18, 36, 71, 130) (dimension is mm, and data come from Domros and Peng 1988, the same below), assume

$$a_1 = a_4 = \frac{2}{130} = 0.0154,$$

$$a_2 = a_3 = \frac{0.48}{18} = 0.026, \tag{13}$$

and $c_i$ $(i=1, 2, 3, 4)$ can be worked out by the least square method:

$$c_1 = -19.2, \quad c_2 = 41.3,$$

$$c_3 = 127.5, \quad c_4 = 9.8, \tag{14}$$

where

$$\tau = 2, \qquad m = 2, \qquad h = 0.2. \tag{15}$$

It is necessary to point out that the selection of functional form and parameter $a_i$ and $c_i$ varies with different places and data, the key lays in debugging and testing process. Fortunately, nowadays with the rapid development of the computer, it is not difficult to undertake debugging activity after the selection of several functional form and several groups of parameters.

## IV. TEST

According to the evaluating criteria of long-term professional forecast, the following is the standard of symbol correlation to test whether the forecast is correct:

(1) The forecast is correct when the symbols of the forecast numbers' and the practical numbers' anomalous percentage are same;

(2) The forecast is correct when the symbols of the forecast numbers' and the practical numbers' anomalous percentage are different or one of the two numbers is zero, but the absolute value of difference between the two numbers is less than 20 percent;

(3) The forecast is incorrect when the above two conditions are not satisfied.

Table 1 shows the comparison between the forecast and practical data of Nenjiang's May, June, July in 1976—1980, among which $\tau$ is 2 $\Delta t$, $\Delta$t's dimension is month. The data come from Domros and Peng (1988), the unit is mm.

**Table** 1. The Comparison between Delayed Precipitation Forecast and Practical Data

| | Practical data | | | Forecast | | |
|---|---|---|---|---|---|---|
| | May | June | July | May | June | July |
| 1976 | 68 | 74 | 172 | 61 | 80 | 189 |
| 1977 | 19 | 83 | 146 | 22 | 91 | 152 |
| 1978 | 39 | 156 | 132 | 43 | 145 | 126 |
| 1979 | 2 | 100 | 82 | 15 | 91 | 90 |
| 1980 | 7 | 108 | 162 | 9 | 117 | 147 |
| mean of 30 years | 36 | 71 | 130 | | | |

In the above experiment, except the forecast of May 1979 and May 1980 (for these two months, practical numbers are too small), all the other forecasts have relative errors less than 15% (relative error = | forecasting number − practical number | ÷ practical number). In the two experiments of May 1979 and May 1980, although the relative errors are large, the absolute errors are respectively 13 and 2 (absolute error = | forecasting number − practical number | ) that are small values. Therefore, the result of above experiments is cheering.

Tables 2—4 give the comparison between the forecast and practical data of Shanghai's February, June and October in 1976—1980. In these tables the time sequence of February (1901—1975) is taken in the forecasting experiment of February and the time sequence of May and June (1901 — 1975) is used in June's forecasting experiment while the time sequence of September, October and November (1901 — 1975) is taken in October's experiment. The respective $\tau$ is $1, 2, 3\Delta t$.

**Table 2.** Comparison between Experiment and Test of February Data

| Time | Feb. 1976 | Feb. 1977 | Feb. 1978 | Feb. 1979 | Feb. 1980 |
|---|---|---|---|---|---|
| Practical data | 4 | −29 | −30 | −52 | −24 |
| Forecast | −7 | −21 | −15 | −45 | −19 |

**Table 3.** Comparison between Experiment and Test of June Data

| Time | June 1976 | June 1977 | June 1978 | June 1979 | June 1980 |
|---|---|---|---|---|---|
| Practical data | −5 | 9 | −95 | −37 | −37 |
| Forecast | 8 | 16 | −111 | −50 | −43 |

**Table 4.** Comparison between Experiment and Test of Oct ober Data

| Time | Oct. 1976 | Oct. 1977 | Oct. 1978 | Oct. 1979 | Oct. 1980 |
|---|---|---|---|---|---|
| Practical data | −6 | −47 | −28 | −67 | −6 |
| Forecast | 20 | −51 | −13 | −55 | 3 |

From above experiments, it is shown that the effect of forecast is satisfying between 65%—73%.

It is necessary to point out that $\tau$'s number chosen will greatly influence the forecasting effect. It is suggested to select from 1—5 experimental numbers. And function is also assumed as $\sin x$ ($5° < x < 85°$) to make experiment.

We would like to thank PENG Yongqing and YAN Shaojin for helpful comments and discussions on the earlier version of this paper, and ZHANG Hongzheng and YE Ruozhou provided editorial assistance.

REFERENCES

Lin Zhenshan (1992), *Theory and Model of Phase Space for Long-Term Forecasting*, China Meteor. Press, Beijing, pp. 1—41 (in Chinese).

Lin Zhenshan (1993), *Nonlinear Mechanics and Atmospheric Sciences*, China Meteor. Press, Beijing, pp. 1—62 (in Chinese).

Domrös' M. and Peng Gongbing (1988), *The Climate of China*, Springer-Verlag, pp. 77—207.