International Conference "Water, Environment, Energy and Society" (WEES–2009)
New Delhi, 12–16 January 2009

Prediction of Rainfall in West Bengal for the Next Two Years

J. Pal Choudhury
Information Technology Department
Kalyani Govt. Engg. College, Kalyani, Nadia, West Bengal - 741 235, INDIA
E-mail: jnpc193@yahoo.com

Kalyan Chakrabrty
AICRP
Bidhan Chandra Krishi Viswabidhyalaya
Kalyani, Nadia, West Bengal - 741 235, INDIA

ABSTRACT: The rainfall plays a major role in agricultural field that generates various agricultural products that generates revenue through marketing not only in Indian market, but also through export in foreign markets. It also affects the economic and financial condition of the state and the country. If the information related to rainfall is available before time, the planners of the state in various fields find it easy to perform their work in various fields related to them. In this paper an effort has been made to predict the rainfall in the state West Bengal in the next two years. Prediction is made based on previous years available data. The method of least square using linear, exponential, curvilinear(parabolic) equation, using orthogonal polynomial and using neural network have been used.

Keywords: Fuzzy Time Series, Neural Network, Forecasting Error, Average Error.

INTRODUCTION

The rainfall plays a major role in agricultural field that generates various agricultural products that generates revenue through marketing not only in Indian market, but also through export in foreign markets. It also affects the economic and financial condition of the state and the country. If the information related to rainfall is available before time, the planners of the state in various fields find it easy to perform their work in various fields related to them. In this paper an effort has been made to predict the rainfall in the state West Bengal in the next two years. Prediction is made based on previous years available data. The method of least square using linear, exponential, curvilinear (parabolic) equation, using orthogonal polynomial and using neural network have been used.

Before using neural network, the raw data is fuzzified and the fuzzy data is applied to neural network as input. The output of neural network contains fuzzy data which has to be defuzzified to get the original output data. The neural network can not generate futuristic terms. The statistical models of least square using linear, exponential, curvilinear (parabolic) equation, using orthogonal polynomial are applied to output data of neural network. Based on minimum error, the particular statistical model has been chosen for the futuristic terms. The previous years of mango generation data is very uncertain, that is the reason for using the technique of fuzzy logic.

Q. Song and Chissom [3] explained the definition of fuzzy time series and discussed the models using Fuzzy Relational equations. Song and Chisson [1] used a time invariant fuzzy time series model. J. Sullivien and W.H. Woodall [2] made a comparative study of Fuzzy Forecasting and Markov Model and suggested that Markov Model would give better prospects. They illustrated the methodology by forecasting the enrollment at the University of Alambama from 20 years of data. H. Bintley [4] applied fuzzy logic and approximate reasoning to a practical case of forecasting. Q. Song and B.S. Chissom[5] used first order time variant models and utilized 3 layer back propagation neural network for defuzzification. G.A. Tagliarini, J.F. Christ and E.W. Page [6] demonstrated that artificial neural networks could achieve high computation rates by employing massive number of simple processing elements of high degree of connectivity between the elements. This paper presented a systematic approach to design neural networks for optimizing applications. T.K. Bhattacharya and T.K. Basu [7] described that a time series model of multiplicative SARIMA (Seasonal Autoregressive Integrated Moving Average) type suffered from a divergent error level in the multistep ahead forecast of all the specific features of the different days of the week. The authors proposed that the data had to be grouped into various subgroups known as Walsch transform, the components of which were then processed by fading memory Kalman filter algorithm and forecasts were made by taking the inverse transform of the predicted value of the component. F.G. Donaldson and M. Kamstra [8] investigated the use of Artificial Neural Network (ANN) to combine time series forecasts of stock market volatility from USA, Canada, Japan and UK. The authors presented combining procedures to a particular class of nonlinear combining procedure based on artificial neural network. J.V. Hansen and R.D. Nelson [9] presented the neural network techniques which provided valuable insights for forecasting tax revenues. The pattern finding ability of neural networks gave insightful and alternate views of the seasonal and cyclical components found in economic time series data. It was found that neural networks were stronger than exponential smoothing and ARIMA (Autoregressive Integrated Moving Average). S.F. Brown, A. Branford and W. Moran [10] proposed that artificial neural networks were powerful tool for analyzing data sets where there were complicated nonlinear interactions between the measured inputs and the quantity to be predicted. M. Sugeno and K. Tanaka [11] proposed successive identification method of a fuzzy model. The structure and initial parameters were determined to identify a model called 'initial model', which was identified by the off-line fuzzy modeling method using some pairs of input-output data. L. Zuoyong, C. Zhenpei and L. Jitao [12] proposed a method of classification of weather forecasts by applying fuzzy grade statistics. The rainfall in a certain region could be forecasted as one of three grades. The range of rainfall was chosen depending on the historical data. The membership functions of the fuzzy sets were also designed. L. Feng and Xu Xia Gaung [13] described the model of fuzzy self-regression. The main steps were the making of the form of self-related sequence number according to the observed number, the calculation of self-related coefficient and the ascertaining of the forecasting model of fuzzy self-regression. M. Ishikawa and T. Moriyama [14] presented various methods of learning and the process of predicting time series analysis, which were ranged from traditional time series analysis to recent approaches using neural networks. It described that back propagation learning had a difficulty in interpreting hidden inputs. In order to solve these problems, a structural learning method was proposed which was based on an infórmation criterion.

STATISTICAL METHODS

Least Squares Method

Let us consider a system of m equations with unknowns $x_1, x_2, x_3, \ldots, x_m$ as follows,

$$\sum_{j=1}^{n} ai_j \; x_j = bi \; (i = 1, 2, 3, ..., m)$$

where ai_j and bi are constants. When $m > n$, the system of equations, has no unique solution. In that case, we shall find the best or most plausible values of the unknowns x_j (j = 1, 2, 3, …., n) which satisfy the system of equations as closely as possible to its accurate values and that can be done by the method of least squares which states that the best or more accurate values of the unknown quantities are that values for which the sum of squares of the residuals is least.

Linear Equation

Let the equation be,

$$y = a + bx$$

So,

$$\sum y = \sum a + \sum bx$$

$$\sum y = na + b\sum x \text{ [where } n = \text{number of records]} \quad \ldots (1)$$

Again,

$$y = a + bx$$

$$xy = ax + bx^2 \text{ [Multiplied both sides by } x\text{]}$$

so,

$$\sum xy = a\sum x + b\sum x^2 \quad \ldots (2)$$

if x in chosen in such a way that $\sum x = 0$

Then,

$$a = \frac{\sum y}{n}$$

and

$$b = \frac{\sum xy}{x^2}$$

Putting the values of a and b in equation,

$$y = a + bx \quad \ldots (3)$$

we get,

$$y = \frac{\sum y}{n} + x\frac{\sum xy}{x^2} \quad \ldots (4)$$

For different values of x, the different values of y have been calculated.

Exponential Equation

Let the equation be,

$y = A\,B^x$

Taking log in both sides,

$\log y = \log A + x \log B$

Let us assume,

$a = \log A$, $b = \log B$ and $y = \log y$

Thus the above equation be,

$y = a + bx$

which is a linear equation.

Using the values of a and b the values of A and B have been calculated as:

A = antilog a, B = antilog b

Curvilinear (parabolic) equation

Let the equation be,

$y = a + bx + cx^2$

Using Least Square Technique,

$$Y = na + b\sum x + c\sum x^2$$

$$\sum xy = a\sum x^2 + b\sum x^2 + c\sum x^3$$

$$\sum x^2 y = a\sum x^2 + b\sum x^3 + c\sum x^4$$

x is chosen in such a way that $\sum x = 0$ and accordingly $\sum x^3 = 0$

The value of b can be calculated as $b = \sum xy/\sum x^2$

The values of a and c can be calculated accordingly.

Orthogonal Polynomial

If the values of x are equally spaced, the manual fitting of polynomial,

$y = b_0 + b_1 x + b_2 x^2 + \ldots$

Is simplified by use of tables of orthogonal polynomial [21].

To replace x^i by a polynomial of degree x_i, the coefficients of polynomials are chosen so that,

$\sum X_i = 0$ and $\sum x_i\, y_i = 0$ (I not $= j$)

The difference polynomials are orthogonal to one another.

The polynomial can be calculated as,

$Y_0 = B_0 + B_1 X_1 + B_2 X_2 + \ldots$

Because of orthogonalty of X_i the following values are obtained,

$B_0 = -Y$, $B_i = \sum X_i\, Y/\sum X_i^2$

Fuzzy Logic

In fuzzy logic ([1]–[5]), unlike standard conditional logic, the truth of any statement is a matter of degree. The notion central to fuzzy systems is that the truth values (in fuzzy logic) or membership values (in fuzzy sets) are indicated by a value on the range [0.0, 1.0], with 0.0 representing absolute False and 1.0 representing absolute Truth. For example, let us take the statement:

"John is old."

If John's age was 75, we might assign the statement the truth value of 0.65. The statement could be translated into set terminology as follows:

"John is a member of the set of old people."

This statement would be rendered symbolically with fuzzy sets as:

*m*OLD (John) = 0.65

where *m* is the membership function, operating in this case on the fuzzy set of old people, which returns a value between 0.0 and 1.0.

At this juncture it is important to point out the distinction between fuzzy systems and probability. Both operate over the same numeric range, and at first glance both have similar values: 0.0 representing False (or non-membership), and 1.0 representing True (or membership). However, there is a distinction to be made between the two statements: The probabilistic approach yields the natural-language statement, "There is an 65% chance that John is old," while the fuzzy terminology corresponds to "John's degree of membership within the set of old people is 0.65." The semantic difference is significant: the first view supposes that John is or is not old (still caught in the Law of the Excluded Middle); it is just that we only have a 65% chance of knowing which set he is in. By contrast, fuzzy terminology supposes that John is "more or less" old, or some other term corresponding to the value of 0.65.

Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a model that emulates a biological neural network. The nodes in an ANN are based on simplistic mathematical representation, which real neurons look like. Today's neural computing uses a limited set of concepts from biological neural systems to implement software simulations of massively parallel processes involving processing elements(also called artificial neurons or neurodes) interconnected in a network architecture.

The neurode (artificial neuron) is analogous to the biological neuron, receiving inputs that represent electrical impulses that dendrites of biological neurons receive from other neurons. The output of the neurode corresponds to a signal sent out from a biological neuron over its axon. The axon of the biological neuron branches to the dendrites of other neurons, and the impulses are transmitted over synapses. A synapse is able to increase or decrease its strength, thus affecting the level of signal propagation and is said to cause excitation or inhabitation of a subsequent neuron. An ANN (Artificial Neural Network) is composed of basic units called artificial neurons on neurodes that are Processing Elements (PEs) in a network. Each neurode receives input data, processes it and delivers a single output. The input data can be raw data or output of other Processing Elements (PEs). The output can be the final product or it can be an input to another neurode.

An ANN (Artificial Neural Network) is composed of collection of interconnected neurons that are often grouped in layers; however in general no specific architecture should be assumed. In terms of layered architecture, two basic structures are considered. In one type, two layers are seen: input and output. In other type there are three layers: input, intermediate (called hidden) and output. An input layer receives data from the outside world and sends signals to subsequent layers. The outside layer interprets signals from the previous layer to produce a result that is transmitted to the outside world. In three layer ANN (Artificial Neural Network) architecture, the concept of hidden layer is assumed in order to control the weights between input to hidden layer and hidden layer to output layer.

Feed Forward Back Propagation Neural Network

The feed forward back propagation neural network (FFBP NN) does not have feedback connections, but errors are back propagated during training. Errors in the output determine measures of hidden layer output errors, which are used as a basis for adjustment of connection weights between the input and hidden layers. Adjusting the two sets of weights between the pairs of layers and recalculating the outputs is an iterative process that is carried on until the errors fall below a tolerance level. Learning rate parameters scale the adjustments to weights. A momentum parameter can be used in scaling the adjustments from a previous iteration and adding to the adjustments in the current iteration. The layout of neural network has been furnished in Figure 1.

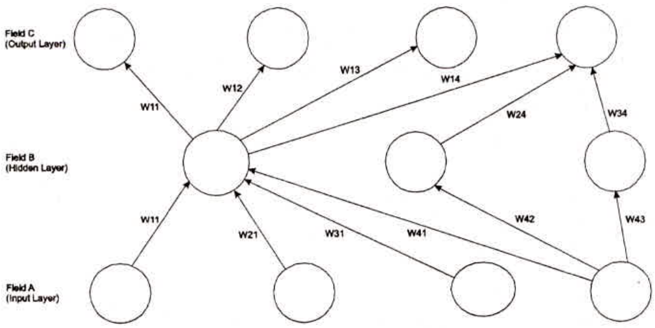

Fig. 1: Layout of a Feed Forward Back Propagation Neural Network

Equations

$M_1[i][j]$ be the matrix of weights from input layer to hidden layer and $M_2[i][j]$ be the matrix of weights from hidden layer to output layer.

Output of j^{th} hidden layer neuron,

$$Y_j = f((\Sigma X_i M_1 [i][j]) + A_j),$$

where A_j is the threshold value or bias for j^{th} hidden layer neuron.

Output of j^{th} output layer neuron,

$$Z_j = f((\Sigma Y_i M_2 [i][j]) + K_j)$$

Where, K_j is the threshold value or bias j^{th} output layer neuron.

I^{th} component of vector of output difference is,

Desired value – Computed value $= P_i - Z_i$

I^{th} component of output error at the output layer is,

$$E_i = (P_i - Z_i)$$

I^{th} component of output error at the hidden layer is,

$$T_i = Y_i (1 - Y_i) (\Sigma M_2 [i][j] E_j)$$

Adjustment for weight between i^{th} neuron in hidden layer and j^{th} output neuron is,

$$\Delta M_2[i][j] = B_1 Y_i E_j + \alpha \Delta M_{2prev} [i][j]$$

where, B_l is the learning parameter for hidden layer and α is the momentum parameter.

Adjustment for weight between i^{th} input neuron and j^{th} neuron in hidden layer is,

$$\Delta M_1[i][j] = B_h X_i T_j + \alpha \Delta M_{1prev} [i][j]$$

where, B_h is the learning parameter for output layer and α is the momentum parameter.

It is no note that $\Delta M_{2prev}[i][j]$ and $\Delta M_{1prev}[i][j]$ are the previous values of $\Delta M_1[i][j]$ and $\Delta M_2[i][j]$ respectively.

Adjustment to the threshold value or bias for j^{th} output neuron is,

$$\Delta K_j = B_l E_j$$

Adjustment to the threshold value or bias for j^{th} hidden layer neuron is,

$$\Delta A_j = B_h T_j$$

Adjustment of weights from input to hidden layer and hidden to output later is,

$$M_2[i][j]\text{new} = M_2[i][j]\text{ old} + \Delta M_2[i][j] \text{ and}$$
$$M_1[i][j]\text{ new} = M_1[i][j]\text{old} + \Delta M_1[i][j]$$

Error Analysis

Forecasting error = | Forecasted Value – Actual Value | /(Actual Value) × 100%

Average Forecasting error = (Sum of Forecasting errors)/(Total no of Errors).

IMPLEMENTATION

The available rainfall figures for previous years have been collected and using these, the prediction of rainfall for futuristic years has to be ascertained using least square techniques, fuzzy logic based neural network. The selection of model is made based on minimum average error.

Statistical Model Using Least Square Technique

The estimated value and average value based on least square technique using linear, exponential, curvilinear (parabolic) equations and based on orthogonal polynomial are calculated. The average error are 7.02%, 7.106%, 6.94%, 6.96%, 7.07% for linear, exponential, parabolic, asymptotic, logistic equations respectively and that of orthogonal polynomial is 5.23%. Now an effort is being made whether the performance of the model can be improved (in terms of average error) using fuzzy logic and fuzzy logic based neural network can be used for better performance using these data.

Neural Network Using Fuzzy Logic

Step 1: The available data are fuzzified based on Gaussian function ([1], [3]). The membership function is chosen as Gaussian based on the trend of data. The actual data and the fuzzy set are furnished in Table 1.

Table 1: Actual Rainfall and Fuzzy Set

Actual Rainfall	*A1*	*A2*	*A3*	*A4*	*A5*	*A6*	*Fuzzy Set*
303.6	0	0	0.4	1	0.6	0	A4
253.5	1	0.2	0	0	0	0	A1
256.2	1	0.4	0	0	0	0	A1
305.2	0	0	0.4	1	0.6	0	A4
311.2	0	0	0.4	1	0.6	0	A4
262.4	1	0.8	0	0	0	0.	A1
311.3	0	0	0	0.8	1	0.2	A5
312.5	0	0	0	0.8	1	0.2	A5
338.2	0	0	0	0	0.2	1	A6
315.3	0	0	0	0.6	1	0.4	A5
295.6	0	0	0.8	1	0.2	0	A4

Step 2: The fuzzy data are applied to neural network. Under artificial neural network system, a feed forward back propagation neural network is used which comprises of a 5 noded input layer, 5 noded output layer and 2 noded hidden layer.

The value of the artificial neural network parameters are furnished in Table 2. It is to mention the M_1 array be the matrix of weights from input to the hidden layer, M_2 array be the matrix of weights from hidden to output layer, A array be the threshold value for j^{th} hidden layer, K array be the threshold or bias to j^{th} output layer as furnished in Table 2. The value of learning rate B_l is taken as 1.5 and B_h as 2.0. The momentum parameter α is taken as 0.7.

Table 2: The Value of Parameters (Feed Forward Back Propagation NN)

M_1 array	0.06	0.04	M_2 array	0.09	0.43	0.25	0.5	0.3	A array	0.2	0.3
	0.2	0.08		0.11	0.07	0.07	0.05	0.2			
	0.05	0.3									
	0.3	0.07	K array	0.15	0.25	0.32	0.35	0.28			
	0.3	0.2									

It has been observed that if momentum parameter is very low, the number of iterations required to achieve the output in comparison to targeted output will be more. As for an example for a low value of momentum of 0.3, number of iterations required for the network to be trained is around 30 whereas in case of momentum of 0.7, the number of iterations required is 10. Learning rate is the degree by which the network is trained. If it is very high, it is very difficult to achieve the point where the network is perfectly is trained, because the network is approaching towards the target and after optimum point again it diverges.

Step 3: The forecasted output is interpreted. The results are all fuzzy sets. Now it is necessary to translate the fuzzy output into a regular number (equivalent scalar). This step is called defuzzification. The following principles are used to interpret the forecasted results:

1. If the membership of the output has only one maximum, the mid point of the interval corresponding to the maximum is selected as the forecasted value.
2. If the membership of an output has two or more consecutive maximums, the midpoint of the corresponding conjunct intervals is the forecasted value.
3. Otherwise, the fuzzy output is standardized and the midpoint of each interval is used to calculate the centroid of the fuzzy set as the forecasted value.

The predicted values for the rainfall are calculated which are furnished in Table 3.

Table 3: Forecasted Value and Forecasted Error Using Artificial Neural Network with Fuzzy Input

Actual Value	*Forecasted Value*	*Forecasting Error (%)*
303.6	–	–
253.5	259.56	2.39
256.2	261.84	2.2
305.2	304.01	–0.389
311.2	312.93	0.556
262.4	264.03	0.619
311.3	312.93	0.524
312.5	312.94	0.139
338.2	330.36	–2.32
315.3	316.01	0.225
295.6	297.74	0.724

Average error 1.01%

RESULTS

It has been observed that the average error for neural network (1.01%) is less that that of statistical models (7.022%, 7.106%, 6.94%, 6.96%, 7.07%, 5.23%) Thus neural network can be used for the futuristic prediction of rainfall data. Since neural network can not generate futuristic terms statistical models (least square technique using linear, exponential, parabolic equations and orthogonal polynomial) can be used.

CONCLUSIONS

Based on average error it has been found that orthogonal polynomial gives minimum value. Therefore the orthogonal polynomial can be used for futuristic prediction of data. The yearly data up to the year 2005 are available, the futuristic data for two years (i.e. 2006 and 2007) are 321.88 and 323.21 respectively.

REFERENCES

[1] Song, Q. and Chissom, B.S. (1993). "Forecasting enrollments with fuzzy time series part I", *Fuzzy Sets and Systems* 54, 1–9.

[2] Sullivan, J. and Woodall, William H. (1994). "A Comparison of Fuzzy Forecasting and Markov Modelling", *Fuzzy Sets and Systems* 64, 279–293.

[3] Song, Q. and Chissom, B.S. (1993). "Fuzzy Time Series and its Models", *Fuzzy Sets and Systems* 54, 269–277.

[4] Bintley, H. (1987). "Time Series analysis with REVEAL", *Fuzzy Sets and Systems* 23, 97–118.

[5] Song, Q. and Chissom, B.S. (1994). "Forecasting enrollments with fuzzy time series—part II", *Fuzzy Sets and Systems* 62, 1–8.

[6] Tagliarini, G.A., Christ, J.F. and Page, E.W. (December 1991). "Optimization using Neural Networks, *IEEE Transactions on Computers,* Vol. 40, No. 12, 1347–1358.

[7] Bhattacharya, T.K. and Basu, T.K. (1993). "Medium range forecasting of power system load using modified Kalman filter and Walsch transform", *Electric Power and Energy Systems,* Vol. 15, No. 2, 109–115.

[8] Donaldson, F.G. and Kamstram M. (1996). "Forecast combining with Neural Networks", *Journal of Forecasting*, 15, 49–61.

[9] Hansen, J.V. and Nelson, Ray D. (July 1997). "Neural Networks and Traditional Time Series Methods: A Synergistic combination in state Economic Forecasts", *IEEE Transactions on Neural Networks*, Vol. 8, No. 4.

[10] Brown, S.F., Branford, A.J. and Moran, W. (Sept. 1997), "On the use of Artificial Neural Networks for the analysis of Survival Data", *IEEE Transactions on Neural Networks,* Vol. 8, No. 5.

[11] Sugeno, M. and Tanaka, K. (1991). "Successive Identification of a Fuzzy Model and its applications to prediction of a complex system", *Fuzzy Sets and Systems* 42, 315–334.

[12] Zuoyong, L., Zhenpei, Chen and Jitao, Li (1998). "A model of Weather Forecast by Fuzzy Grade Statistics", *Fuzzy Sets and Systems* 26, 275–281.

[13] Feng, L. and Guang, Xu Xia (1993). "A Forecasting Model of Fuzzy Self Regression", *Fuzzy Sets and Systems* 58, 239–242.

[14] Ishikawa, M. and Moriyama, T. (1996). "Prediction of Time Series by a Structural Learning of Neural Networks", *Fuzzy Sets and Systems* 82, 167–176.

[15] Zimmermann, H.J. (1996). *Fuzzy Set Theory and it applications, second revised edition*, Allied Publishers Limited.

[16] Klir, G.J. and Yuan, B. (1997), *Fuzzy Sets and Fuzzy Logic Theory and Application,* Prentice Hall of India Pvt. Limited, New Delhi.

[17] Chen, S.M. (1996). "Forecasting enrollments based on fuzzy time series", *Fuzzy Sets and Systems* 81, 311–319.

[18] Choudhury, J. Paul, Sarkar, Bijan and Mukherjee, S.K. (October–November 1999). "Some Issues in building a Fuzzy Neural Network based Framework for forecasting Engineering Manpower", *Proceedings of 34th Annual Convention of Computer Society of India*, pp. 213–227.

[19] Choudhury, J. Paul, Sarkar, Bijan and Mukherjee, S.K. (October–November 1999). "Rule Base of a Fuzzy Expert Selection System", *Proceedings of 34th Annual Convention of Computer Society of India*, Mumbai, pp. 98–104.

[20] Choudhury, J. Paul, Sarkar, Bijan and Mukherjee, S.K. (March 2000). "A Fuzzy Time Series based Framework in the Forecasting Engineering Manpower in comparison to Markov Modelling", *Proceedings of Seminar on Information Technology*, The Institution of Engineers (India), Computer Engineering Division, West Bengal State Center, pp. 39–45.

[21] Snedecor, G.W. and Cochran, W.G. (1994). *Statistical Methods*, eight edition, East Press, New Delhi.