On the Prediction of Average Monsoon Rainfall in Bangladesh with Artificial Neural Network

Md. Habibur Rahman Lecturer Department of Statistics Jahangirnagar University Bangladesh

ABSTRACT

The monsoon rainfall has very important affect on the agricultural production, livestock as well as human ecology. In this study, try to fit Artificial Neural Network (ANN) model to predict average monsoon rainfall. For the ANN models monthly average rainfall, sea surface temperature, wind speed, monsoon rainfall, temperature is used as inputs to predict the monsoon rainfall. The feed forward network was trained using a variety of algorithms. For the different networks use sigmoid transfer function, tan sigmoid transfer function and linear transfer function. The total sample was divided into a training set (first 75 percent) and a testing set (last 25 percent). The data pertaining to the years 1961 to 2005 have been explored to develop the predictive models. The model performance is measured by prediction error, mean square error, root mean square error, correlation, similarity, mean percentage error and mean absolute percentage error. Finally, the prediction performance of artificial neural network has compared with polynomial curve fitting, Fourier series, auto regressive moving average model (ARMA), and multiple linear regressions. The average monsoon rainfall prediction based on Artificial Neural Network was found to be superior to that based on polynomial curve fitting, multiple linear regression, ARMA model and Fourier series. Finally, made cluster analysis between actual average monsoon rainfall and predicted average monsoon rainfall by different ANN and other statistical models. From the dendrogram, it is evident that the actual monsoon rainfall and predicted rainfall by ANN fall in one cluster. The ANN model gives more accurate prediction compared to other models.

General Terms

Monsoon rainfall, multiple linear regressions, polynomial curve fitting, Fourier series.

Keywords

Artificial Neural Networks, Similarity, Cluster analysis, Activation function, Prediction.

1. INTRODUCTION

Bangladesh is the country where economy is largely depends on agriculture. The impact of low and high precipitation is very important for agricultural production. Livestock and human ecology depends on the water system especially rainfall. Drought of different intensities occurs across Bangladesh. Rice and other crops productions are severely affected by dryness. To make better understanding of rainfall system which can help the formulation of policies that might include preserve and an efficient use of rainfall water in near future. In order to implement a better policy the policy makers especially farmers need to know the future state of amount of M.A. Matin, PhD Professor Department of Statistics Jahangirnagar University Bangladesh

rainfall. Since in all weather happenings, rainfall plays the most imperative role in human life especially human civilization to a great extent depends upon its frequency and amount to various scales.

Being depended on time and space, it is a chaotic system. The weather forecasting is one of the most imperative and demanding operational responsibilities carried out by meteorological services all over the world. It is a complicated procedure that includes numerous specialized fields of knowhow. The task is complicated because in the field of meteorology all decisions are to be taken in the visage of uncertainty. A number of authors Brown and Murphy (1988), Handerson and Wells (1988), Wilks (1991), Elsner and Tsonis (1992), Jacovides et al. (1994), Kondratyev and Varotsos (2001a, b), Cartalis and Varotsos (1994), Varotsos et al. (2001) have discussed the uncertainty associated with the weather systems. Chaotic features associated with the atmospheric phenomena also have been attracted the attention of the modern scientists Varotsos (2005), Varotsos and Krik-Davidoff (2006), Sivakumar (2000), Sivakumar et al. (1999), Men et al. (2004).

The task also requires sophisticated statistical models, which do not need making assumptions regarding the system. Many studies around the globe have developed stochastic weather models which are basically statistical models that can be used as a random number generators whose output resembles the weather data to which they have been fit (Wilks 1999).

Hu (1964) initiated the implementation of artificial neural network, an important soft computing methodology in weather forecasting (Surajit C. and Manojit C., 2007). There after a voluminous development in the field of applications of ANN has opened up a new avenue in forecasting task involving atmosphere related phenomena (Gardner and Dorling, 1998), Hsieh and Tang (1998). Michaelides et al (1995) compared the performance of ANN with multiple linear regressions in estimating missing rainfall data over Cyprus. Kalogirou et al. (1997) used ANN in rainfall prediction by splitting the available data into homogeneous subpopulations. Wong et al. (1999) constructed fuzzy rules bases with the aid of Self Organization Map (SOM) and back propagation neural networks and then with the help of the rule based developed predictive model for rainfall over Switzerland using spatial interpolation.

The main objective of this study is to develop an artificial neural network model to forecast average rainfall during monsoon in Bangladesh. In doing so a comparison between ANN model and other Statistical predictive models, to training and testing of the model, to evaluate the goodness fit of the model.

2. THEORY BEHIND PREDICTIVE MODEL

Forecasting is most often part of a larger process of planning and managing. A forecast is necessary to provide accurate estimates of the future for the larger process.

2.1 Polynomial Curve Fitting

The Curve fitting technique is used in many applications of data analysis. It is advantageous when the data is follows some pattern with a lesser amount of randomness about the mean. It actually develops an implicit equation for determining one variable with respect to one or more variables (Chernov N. and Lesort C., 2004). Polynomial curve fitting can be considered as Generalized Linear Least-Squares Regression (Arora S. and Khot S., 2003).

A general implicit equation for a variable (θ) in terms of

variables x and y where (x_i, y_i) is experimental data for

i=1, 2, ..., n can be written as $P(x, y, \theta) = 0$. Here *P* refers to polynomial. The coefficients of polynomial need to be determined for fitting the curve while the degree of the curve will determine the number of coefficients. Accordingly the polynomial curve will have only one parameter in estimating θ . In simple statement $\Theta = f(y)$, where Θ is the dependent variable and y is independent variable. Or in other words, $\Theta = a_0 + a_1y + a_2y^2 + a_3y^3 + \dots + a_ny^n$. With available data; polynomial equations could be found in different degrees and corresponding mean square errors and correlation coefficients. If the observed values of θ are θ_i ; $i=1, 2, \dots, n$. The best fitted residual should be least and is called the least squares. The polynomial curve fitting is a generalized form of least squares method where Mean square error is

 $MSE = \frac{\sum (f(y) - \hat{f}(y))^2}{n}$ and Coefficient of correlation is

$$R = \frac{\sum \left[\left(\theta - \hat{\theta} \right) \left(f(y) - \hat{f}(y) \right) \right]}{\sqrt{\left[\left\{ \sum \left(\theta - \hat{\theta} \right)^2 \right\} \left\{ \sum \left(f(y) - \hat{f}(y) \right)^2 \right\} \right]}}$$

2.2 Fourier Series Analysis

Fourier series analysis has advantages in automated in forecasting and is extremely useful in data analysis. The Fourier series is a sum of sine and cosine functions that is used to describe a periodic signal. It is represented either in trigonometric form or in exponential form. The general form of the trigonometric Fourier series is

$$f(t) = \hat{y} = a_0 + \sum_{i=1}^n a_i \sin(n\omega t) + b_i \cos(n\omega t)$$

where, $\hat{y} = \text{fitted or forecasted value at time } t$,

 $a_0 = \text{constant}, n = \text{harmonic of } \omega$

 a_i, b_i = coefficients defining the amplitudes and phases,

 $\omega = 2\pi f / n$, fourier frequency

where, models any DC offset in the signal and is associated with the i = 0 cosine term, ω is the fundamental frequency of the signal, n is the number of terms (harmonics) in the series.

2.3 Multiple Linear Regression

The Regression analysis is used to predict the values of one or more response variables from the collection of predictor variables values. It is also used in assessing the effects of the predictor variable on the responses.

Let $z_1, z_2, z_3, ..., z_r$ be the *r* predictor variables thought to be related to a response variable *Y*. The classical regression model states that *Y* is composed of mean, which depend on the continuous manner on the z_i 's and a random error ε . The linear regression model with single response takes the form as $Y = Z\beta + \varepsilon$ where

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ Y_n \end{bmatrix}_{(n \times 1)} \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \beta_r \end{bmatrix}_{(r \times 1)} \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \varepsilon_n \end{bmatrix}_{(n \times 1)} Z = \begin{bmatrix} 1 & z_{11} & z_{12} & \cdot & z_{1r} \\ 1 & z_{21} & z_{22} & \cdot & z_{2r} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & z_{n1} & z_{n2} & \cdot & z_{nr} \end{bmatrix}_{[n \times (r+1)]}$$

where the error terms are assumed to have the properties $E(\varepsilon) = 0$, and $E(\varepsilon\varepsilon') = \sigma^2 I$. One of the objectives of the regression analysis is to develop an equation to predict the response for given values of predictor variables. Let *b* be the trail values for β . The method of least squares selects *b* so as to minimize the sum of squares of the differences S(b) = (y - Zb)'(y - Zb), the coefficient *b* chosen by the least squares criterion are called least squares estimator of the regression parameters β and denoted by $\hat{\beta} = (Z'Z)^{-1}Z'y$. Hence the fitted regression model becomes $\hat{Y} = Z\hat{\beta}$

2.4 Auto-Regressive Moving Average (ARMA) Models

The Auto-Regressive Moving-Average (ARMA) method is used in modeling which widely used in hydrology, dendrochronology, and many other fields (Chatfield, C., 2004). Auto-regressive Moving-average models are mathematical models of the persistence, or autocorrelation, in a time series (Helmut Lutkepohl and Markus Kratzig). An ARMA model can effectively be used to predict behavior of a time series from the past values alone. The ARMA model is a generalization of Auto-Regressive (AR) and Moving Average (MA) models and is a special case of Auto-regressive Integrated Moving Average (ARIMA) models.

The Auto-Regressive models and moving average models are combined to form one model called Auto-Regressive Moving Average model. The order of AR model is the time steps which the model will go back to predict the future value and the order of MA model is the past difference steps which the model will go to predict the future value. The auto-regressive model includes lagged terms on the time series itself, and that the moving-average model includes lagged terms; it gives what are called auto regressive moving average models. The order of the ARMA model is included in parentheses as ARMA (p, q), where p is the auto-regressive order and q the moving-average order. The simplest and most frequently used ARMA model is ARMA (1, 1) model (Pappas S.SP., Ekonomou, L. Karampelas P., Katsikas S.K and Liatsis. P., 2008).

$$y_t + a_1 y_{t-1} = e_t + b_1 e_{t-1}$$
.

where y_t is the mean-adjusted series in year t, y_{t-1} is the series in the previous year i.e. lag of one, a_1 is the lag-1 autoregressive coefficient, where e_t and e_{t-1} are the residuals or the noise or the random-shock at times t and t-1, and b_1 is

the first-order moving average coefficient. The residuals e_t are assumed to be random in time (not auto correlated), and normally distributed.

The general form of ARMA (p, q) model is given by

$$\left[1-\sum_{s=1}^{p}\Psi_{s}(L)s\right]y_{t}=\left[1-\sum_{s=1}^{q}\Phi_{s}(L)s\right]\varepsilon_{t},$$

where *p* and *q* are called orders of the model, Ψ_s is called AR model coefficients and Φ_s is called the MA model coefficients, *L* is an operator called lag operator when it operates on any data points; the result is the previous data value that is $L^s(y_t) = y_{t-s}$. In short form the ARMA model can be written as $\Psi(L)y_t = \Phi(L)\varepsilon_t$. This is the most general form of ARMA (p, q). Also it can be written as

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + b_1 e_{t-1} + \dots + b_q e_{t-q} + e_t$$

There are different methods to determine the order of the ARMA model. Some of them are auto-correlation function (ACF), partial auto-correlation function (PACF). The Order selection can be made on the basis of the model validity criteria; such as the Akaike's information criteria (AIC) and the minimum description length (MDL) (L. Ljung, 1987).

2.5 Artificial Neural Networks

Forecasting is the part of a larger process of planning and managing most often which forecast is necessary to provide accurate estimates of the future for the larger process. Artificial neural networks are similar to the biological neural networks that are present in human brain. It is the networks constructed by connecting artificial neurons. It is the input out networks (Anil K. Jain, Jianchang Mao and Mohiuddin .K.M, 1999).

Feed forward neural network has become popular tool for solving complex prediction as well as classification problems. In general, a feed forward neural network consists of a lowermost input layer, any number of hidden layers, and an output layer. In a network, the total input received by neuron 'j' in (h+1)th layer is defined as

$$I_j^{h+1} = \sum_i O_i^h W_{ji}^h - \theta_j^{h+1}$$

where, O_i^h =state of the *i* th neuron preceding *h* th layer

 W_{ii}^{h} = weight of the connection between *i* th neuron of

h th layer and *j* th neuron of (h+1) th layer

 θ_j^{h+1} = threshold of *j* th neuron in (h+1) th layer

i = tending node number, and

j = receiving node number

If the threshold $\boldsymbol{\theta}$ is taken to be zero, the above equation modified to

$$I_{j}^{h+1} = \sum_{i} O_{i}^{h} W_{ji}^{h} ,$$

where, $O_{i}^{h} = \left[\frac{1}{\left(1 + e^{-I_{j}^{h}} \right)} \right]$ or any other transfer function. The

output value is compared to the desired output and error is calculated. Based on error, back propagate errors to adjust weight. Then calculate the residual standard error. If RMS is low enough then it should be validated.

3. MATERIALS AND MODEL FITTING PROCEDURE FOR PREDICTION OF AVERAGE MONSOON RAINFALL

The ANN approach has several advantages over conventional phenomenological or semi-empirical models, since they require known input data set without any assumptions. The model building process consists of the following sequential steps: (i) selection of the input and output for the supervised back propagation learning, (ii) selection of the activation function, (iii) training and testing of the model, and (iv) testing the goodness of fit of the model.

In this paper, data pertaining to years 1961 through 2006 are explored to develop the predictive model. The inputs for the different model(s) are summarized below: average monsoon rainfall at time $t(x_1)$, sea surface temperature (x_2) , temperature (x_3) , rainfall (x_4) , wind speed (x_5) , sea level pressure (x_6) , rainfall of June (x_7) , rainfall of July (x_8) , rainfall of August (x_9) , rainfall of September (x_{10}) . The predictand is the average monsoon rainfall at time t + 1(y) in Bangladesh. The weather component measure are recorded at different locations of Bangladesh by the Bangladesh Metrological Department. Table 1 shows the summary statistics of weather component.

Name	Mean	Standard Deviation	Standard Error	Coefficient of variation	Skewness	Kurtosis	Minimum	Maximum
Monsoon Rainfall	407.3	50.2	7.49	12.3	-0.2	-0.5	303	511
Sea Surface Temperature	28.5	0.2	0.03	0.7	0.4	1.1	28	29
Temperature	25.4	0.3	0.05	1.3	1.1	2.1	25	26
Rainfall	162.4	22.4	3.33	13.8	-0.1	-0.5	117	209
Wind Speed	1.32	0.4	0.06	31.3	0.3	-0.4	0.5	2
Sea Level Pressure	1009	0.8	0.12	0.1	1.7	4.3	1008	1012

Table 1. Summary Statistics

It is interesting to see that the time series of Bangladesh monsoon rainfall are not pair wise correlated. The manual

Pearson correlations between predictand and predictors values are presented in Table 2.

Table 2. Following table present the Correlation between predictand and predictors

Name	Sea Surface Temperature	Temperature	Rainfall	Wind Speed	Sea Level Pressure	Monsoon Rainfall (t+1)
Monsoon Rainfall (t)	-0.12	-0.14	0.26	-0.19	-0.18	0.15
Sea Surface Temperatures		0.47	0.05	-0.27	0.11	-0.13
Temperature			-0.14	-0.16	-0.18	-0.33
Rainfall				-0.12	0.01	0.13
Wind Speed					-0.13	0.01
Sea Level Pressure						0.03

All the correlation values are too small, indicating that the relationships are highly nonlinear. The Pearson correlation between the monsoon rainfall and predictors are neither too low nor significantly high. Thus, the necessity of implementing ANN in the prediction problem is felt highly relevant. The first 75 percent data are taken as the training set and the remaining 25 percent data are taken as the test set or validation set.

To avoid the asymptotic effect the raw data are scaled $\begin{pmatrix} x & x \\ y & y \end{pmatrix}$

according to (Surajit, C., 2007)
$$z_i = 0.1 + 0.8 \left(\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \right)$$

where, Z_i denotes the transformed appearance of the raw

data X_i . After the modeling is completed, the scaled data are reversed scaled according to

$$P_i = x_{\min} + \left(\frac{1}{0.8}\right) \left[(y_i - 0.1)(x_{\max} - x_{\min}) \right],$$

where, P_i denotes the prediction in original scale.

After developing the model through training and testing, performance and goodness of fit of the model has to be checked out or validated. The overall prediction error (PE) is measured as (Perez and Reyes, 2001).

$$PE = \frac{\left\langle \left| y_{predicted} - y_{actual} \right| \right\rangle}{\left\langle y_{actual} \right\rangle}$$

where, $\langle \rangle$ implies the average of the whole test set. The

predictive model is identified as a good one if the PE is sufficiently small, that is close to zero. The model with minimum PE is identified as the best prediction model.

4. ARTIFICIAL NEURAL NETWORK MODEL

4.1 ANN Model type 1

The ANN model type 1 fit on the basis of six inputs such as monsoon rainfall, rainfall, sea level pressure, sea surface temperature, wind speed and temperature at time t to predict the monsoon rainfall at time (t+1). The network components are presented in Table 3.

Table 3. Basic network Components of the ANN model Type 1

Network Architecture						
Number of inputs	6	Number of output		1		
Number of Hidden layer	1	Hi	dden layer size	2		
Learning parameter	0.2	Ini rar	tial weight 1ge	±0.5		
Momentum	0.8					
Total rows in data	45	No. of training cycles		300		
Save training weights	With least training error	Training mode On- line		On- line		
Training/Validation set	Partition data into training/validation set					
Activation function in Hidden	Sigmoi	d				
Activation function in Output layer			Sigmoid			

4.2 ANN Model type 2

The ANN model type 2 fit on the basis of four inputs such as monsoon rainfall, rainfall, sea surface temperature, and temperature at time t to predict the monsoon rainfall at time

(t+1). The network components are presented in Table 4.

Here, the hyperbolic tangent function used as activation function in the hidden layer and output layer.

Table 4. Basic network Components of the ANN model Type 2

Network Architecture					
Number of inputs	4	Number of output	1		
Number of Hidden layer	2	Size in Hidden layer 2	2		
Size in Hidden layer 1	3	Initial weight range	±0.5		
Learning parameter	0.2	Momentum	0.7		
Total rows in data	45	No. of training cycles	210		
Save training weights	With least training error	Training mode On lin			
Training/Validation set	Partition data i	nto training/validation set			
Activation function in Hi	Hyperbolic Tangent				
Activation function in Ou	itput layer	Hyperbolic Tangent			

4.3 ANN Model type 3

The ANN model type 3 fit on the basis of four inputs such as rainfall of June, rainfall of July, August, and rainfall September at time t to predict the monsoon rainfall at time

(t+1). The network components are presented in Table 5.

Table 5. Basic network Components of the ANN model Type 3

Network Architecture					
Number of inputs	4	Number of output	1		
Number of Hidden layer	2	Size in Hidden layer 2	2		
Size in Hidden layer 1	3	Initial weight range	±0.5		
Learning parameter	0.2	Momentum	0.75		
Total rows in data	45	No. of training cycles	160		
Save training weights	With least training error	Training mode	On- line		
Training/Validation set	Partition data i	nto training/validation	set		
Activation function in Hide	on in Hidden layer Hyperbolic Tangen		ent		
Activation function in Outp	out layer	Identity			

4.4 ANN Model type 4

In the ANN model type 4 the inputs of the rainfall of June, July, August and September and used Levenberg-Marquardt Algorithm (trainlm) as training function.

 Table 6. Basic network Components of the ANN model

 Type 4

Network Architecture					
Number of inputs	4	Number of output	1		
Number of Hidden layer	1	Hidden layer size	8		
Learning parameter	0.2	Initial weight range	±0.5		
Momentum	0.8	Training option			
Total rows in data	45	No. of training cycles	228		
Save training weights	With least training error	Training function train			
Training/Validation set	Partition data i	nto training/validation	on set		
Activation function in Hidden layer Tan-Sigmoid			id		
Activation function in Out	out layer	Activation function in Output layer Linear			

4.5 ANN model type 5

In the ANN model type 5 the inputs are the monsoon rainfall of previous year, sea surface temperature, temperature, wind speed, rainfall and sea level pressure at time t and the monsoon rainfall at time (t+1) is the target and used Levenberg-Marquardt Algorithm as training function.

Table 7. Basic network Components of the ANN model Type 5

Network Architecture					
Number of inputs	6	Number of output	1		
Number of Hidden layer	1	Hidden layer size	5		
Learning parameter	0.2	Initial weight range	±0.5		
Momentum	0.85	Training option			
Total rows in data	45	No. of training	371		

		cycles	
Save training weights	With least training error	Training function	trainlm
Training/Validation set	Partition data into training/validation set		
Activation function in Hidden layer		Tan-Sigmoid	
Activation function in Output layer		Linear	

5. FITTED PRDICTIVE MODELS

The following statistical models are fitted like the neural net works model, first 75 percent data are used for training, and the last 25 percent data are used for testing. Finally the equations take the forms

Multiple Linear Regression Model

MLR (1):
$$\hat{y} = 30743 - 0.33 x_1 + 11.41 x_2 - 10.78 x_3$$

+ 0.44 $x_4 - 47.43 x_5 - 29.99 x_6$

MLR(2):
$$\hat{y} = 29402.1 - 7.4 x_2 - 7.1 x_3 + 0.12 x_4$$

- 64.57 $x_5 - 27.3 x_6$
MLR(3): $\hat{y} = 514.4 - 0.213 x_7 + 0.05 x_8 - 0.02 x_9 - 0.03 x_{10}$

Polynomial Curve Fitting Model

$$\hat{y} = f(x) = -2.464x^9 - 4.24x^8 + 32.34x^7 + 41.55x^6$$

-87.83x⁵ - 132.9x⁴ + 49.92x³ + 139.6x²
+ 13.19x + 385.1

where, x is normalized by mean 1979 and std 9.958

ARMA (1, 1) Model

 $\hat{y} = y_t = 380.57 + 0.07091 y_{t-1} - 0.26249 e_{t-1}$

Fourier series Model

$$\hat{y} = f(t) = 405.7 \cdot 8.939 \cos(\omega t) + 10.69 \sin(\omega t) -7.264 \cos(2\omega t) -11.73 \sin(2\omega t) -10.59 \cos(3\omega t) + 5.814 \sin(3\omega t) -0.5913 \cos(4\omega t) + 18.28 \sin(4\omega t) - 18.33 \cos(5\omega t) - 10.62 \sin(5\omega t) + 5.715 \cos(6\omega t) - 8.076 \sin(6\omega t) + 6.243 \cos(7\omega t) + 4.39 \sin(7\omega t) -10.33 \cos(8\omega t) - 22.03 \sin(8\omega t) \text{ where, } \omega = 0.2856$$



Figure 1(a): Artificial Neural Network model type 1.



Figure 1(b): Artificial Neural Network model type 2.



Figure 1(d): Artificial Neural Network model type 4.



Figure 1(c): Artificial Neural Network model type 3.



Figure 1(e): Artificial Neural Network model type 5.

measure. There are several types of measure to measure the

accuracy as: mean absolute deviation (MAD), mean square

error (MSE), root mean square error (RMS), mean absolute

prediction error (PE) are widely used statistics in forecasting.

These measures are the powerful diagnostic tools in

percentage error (MAPE), correlation, similarity,

Figure 1: Actual versus predicted average monsoon rainfall in training cases and test cases

6. PERFORMANCE EVALUATION AND ACCURACY MEASUREMENTS FOR DIFFERENT MODEL FOR PREDICTED VALUES IN TRAINING AND TEST

CASES

In this section, calculate the different forecasting accuracy

Model	MAD	MSE	RMS	Correlation	MAPE	Similarity	Prediction Error
ANN_1	33.71	473.15	21.75	0.68	8.209	0.944	0.107
ANN_2	35.98	480.14	21.91	0.62	8.387	0.951	0.101
ANN_3	31.90	529.46	23.02	0.69	7.984	0.944	0.100
ANN_4	9.47	175.07	13.23	0.97	2.387	0.999	0.023
ANN_5	7.15	157.09	12.53	0.97	1.762	0.999	0.018
Polynomial	41.08	2403.39	49.02	0.25	10.54	0.936	0.199
Fourier Series	42.06	2455.28	49.55	0.19	10.73	0.925	0.194
MLR-1	37.84	2127.23	46.12	0.41	9.557	0.937	0.186
MLR-2	39.21	2278.65	47.74	0.32	9.923	0.935	0.186
MLR-3	38.30	2231.78	47.24	0.35	9.696	0.947	0.189
ARMA (1,1)	40.34	2419.36	49.19	0.23	10.27	0.926	0.201

Table 9: Goodness of Fit Test Results for Different Model(s)

forecasting.

and



Figure 2: Comparision of the Prediction Errors of Different Competative Models





Figure 3: Complete Linkage Dendrogram for Manhattan Distances between Actual Rainfall and Predicted Rainfall from Different Models.

on Rainfall



Figure 4: Comparison of Observed Average Monsoon Rainfall and Predicted Average Monsoon Rainfall with Artificial Neural Networks Models (ANN4 and ANN5)

7. RESULT AND DISCUSSION

The correlation between average monsoon rainfall (predictand) and sea surface temperature, wind speed, sea level pressure, temperature, rainfall (predictors) are estimated. The correlation between targeted average monsoon rainfall and sea surface temperature, temperature, rainfall, wind speed, sea level pressure are -0.133, -0.329, 0.131, 0.004, 0.034 respectively. All the correlation values are too small, indicating that the relationships are highly nonlinear. The prediction error for ANN model type 1, ANN model type 2, and ANN model type 3, ANN model type 4, and ANN model type 5 are 0.107, 0.101, and 0.100, 0.023, and 0.018 respectively. For multiple linear regression, polynomial curve fitting, Fourier series, ARMA model the prediction errors are greater than 0.186. Also the similarity measure according to Jaccard is found to be almost 0.99 for ANN model type 4, ANN model type 5. The value of RMS for ANN model(s) is less than the other model. The above results indicate that the artificial neural network models are better than the other statistical models. Also to fit two models ANN model type 4 and ANN model type 5 uses the tan sigmoid function in input layer and the linear function in hidden layer as activation or transfer function. In both the case the "Levenberg Marquardt" algorithm used as Back propagation algorithm which is known as training function. For ANN model type 4 the RMS, correlation, similarity, prediction error are 13.23, 0.97, 0.999,

0.023, and for ANN model type 5 are 12.53, 0.97, 0.999, and 0.018 respectively. The maximum epoch for ANN model type 4 and ANN model type 5 are 228 and 371 respectively. From the above result it is evident that the ANN model type 4 and ANN model type 5 are better than the other model. The performance and accuracy measurement for different models for predicted values in training and test cases evident that the value of mean absolute deviation, root mean square, mean absolute percentage error, prediction error for ANN model type 4 are 9.47, 13.23, 2.387,0.023 and for ANN model type 5 are 7.15, 12.53, 1.762,0.018 respectively. These criteria are lower than other statistical and ANN models. Also the correlation and similarities are relatively higher than the other models. The Figure 3 represents the complete linkage Dendrogram using Manhattan distance which shows that actual monsoon rainfall and predicted average monsoon rainfall by ANN model type 4 and ANN model type 5 falls in one cluster. Thus it can be said that Artificial Neural Network model(s) are the best prediction model.

8. CONCLUSSION

The monsoon rainfall prediction based on Artificial Neural Network was found to be superior to that based on Polynomial curve fitting, multiple linear regression, ARMA model and Fourier series. From cluster analysis between actual monsoon rainfall and predicted monsoon rainfall by different ANN and other Statistical models it is evident that the actual monsoon rainfall and predicted rainfall by ANN fall in one cluster. Thus, the ANN model gives more accurate prediction compared to other models. Performance of the model is evaluated through computation of overall Prediction Error (PE), Sum Square Error (SSE), Mean Square Error (MSE), Root Mean Square Error (RMS), Correlation, Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and different Similarity Measure as Dice, Jaccard, and Cosine are minimum for ANN prediction model, which are less than other models. And it also found that the correlation and different similarities measure are relatively high for ANN(s). In addition to this, the actual average monsoon rainfall and predicted average monsoon rainfall by ANN fall in one cluster. Therefore, Artificial Neural Network is found to be adroit in the prediction of average monsoon rainfall over Bangladesh. Finally, wish to close by mentioning some related areas of further research. In the prediction process the Artificial Neural Network models fitted on the basis of less or equal to six inputs. Further study might focus on fit an ANN model to predict monsoon rainfall on the basis of all variables or weather parameter of all stations and the model accuracy can be improved by Bayesian perspective.

9. REFERENCES

- Anil K. Jain, Jianchang Mao and Mohiuddin .K.M, "Artificial neural networks: a tutorial" Michigon State University March (1999)
- [2] Arora S. and Khot S (2003)., "Fitting algebraic curves to noisy data," *Journal of Computer and System Sciences* 67, 325–340.
- [3] Brown, B.G., and A.H. Murphy, (1988), "The economic value of weather forecasts in wildfire suppression mobilization decisions", Canadian J. Forest Res. 18, 1641-1649.
- [4] Cartalis, C., and C. Varotsos, (1994), Surface ozone in Athens, Greece, at the beginning and at the end of the 20th-century, Atmos. Environ. 28, 3-8.
- [5] Chapman & Hall/CRCChernov N. and Lesort C, (2004), "Statistical efficiency of curve fitting algorithms" *Computational* Statistics & Data Analysis 47, 713 – 728.
- [6] Chatfield, C., (2004), "*The analysis of time series, an introduction*", sixth edition: New York,
- [7] Chernov N. and Lesort C, (2004), "Statistical efficiency of curve fitting algorithms" *Computational* Statistics & Data Analysis 47, 713 – 728.
- [8] Elsner, J.B., and A.A. Tsonis, (1992), Non-linear prediction, chaos, and noise, Bull. Am. Me-teor. Soc. 73, 49-60.
- [9] Gardner, M.W., and S.R. Dorling, (1998), "Artificial Neural Network (Multilayer Perceptron) – A review of applications in atmospheric sciences", *Atmos. Environ.* 32, 2627-2636.
- [10] Helmut Lutkepohl and Markus Kratzig, (2004) "Applied Time Series Econometrics" Cambridge University Press 2004 ISBN 052183919X
- [11] Hsieh, W.W., and T. Tang, (1998), Applying neural network models to prediction and data analysis in

meteorology and oceanography, Bull. Am. Meteor. Soc. 79, 1855-1869.

- [12] Hu, M.J.C., (1964), Application of ADALINE system to weather forecasting, Technical Report, Stanford Electron, Stanford, CA.
- [13] Jacovides, C.P., C. Varotsos, N.A. Kaltsounides, M. Petrakis and D.P. Lalas, (1994), Atmospheric turbidity parameters in the highly polluted site of Athens basin, Renewable Energy 4, 5, 465-470.
- [14] Kalogirou, S.A., C.N. Constantinos, S.C. Michaelides and C.N. Schizas, (1997), A time series construction of precipitation records using Artificial Neural Networks, EUFIT '97, September 8-11, 2409-2413.
- [15] Kondratyev, K.Y., and C.A. Varotsos,(2001a), Global tropospheric ozone dynamics – Part I: Tropospheric ozone precursors, Environ. Sci. Pollution Res. 8, 1, 57-62.
- [16] L. Ljung, (1987) "System Identification: Theory for the User", Prentice Hall, Engle wood Cliis, NJ).
- [17] Men, B., Z. Xiejing and C. Liang,(2004), Chaotic analysis on monthly precipitation on Hills Region in Middle Sichuan of China, Nature and Science 2, 45-51.
- [18] Michaelides, S.C., C.C. Neocleous and C.N. Schizas, (1995), Artificial Neural Networks and multiple linear regression in estimating missing rainfall data, Proc. DSP95 Intern. Confer. "Digital Signal Processing", Limassol, Cyprus. 668-673.
- [19] Pappas S. SP., Ekonomou, L. Karampelas P., Katsikas S.K and Liatsis. P., "modelling of round water resistance variation using ARMA models," Simulation Modelling Practice and Theory 16 (2008) 560–570.
- [20] Sivakumar, B., S.Y. Liong, C.Y. Liaw and K.K. Phoon, (1999), Singapore rainfall behavior: Chaotic, J. Hydrol. Eng., ASCE 4, 38-48.
- [21] Sivakumar, B., (2000), Chaos theory in hydrology: important issues and interpretations, J. Hy-drology 227, 1-20.
- [22] Surajit C. and Manojit C. (2007), "A Soft Computing technique in rainfall forecasting", International Conference on IT, HIT, March 19-21, 2007
- [23] Varotsos, C., (2005), Power-law correlations in column ozone over Antarctica, Intern. J. Remote Sensing 26, 3333-3342.
- [24] Varotsos, C., and D. Krik-Davidoff, (2006), Longmemory processes in ozone and temperature variations at the region 60 degrees S-60 degrees N, Atmos. Chem. Phys. 6, 4093-4100.
- [25] Varotsos, C., K.Y. Kondratyev and M. Efstathiou, (2001), On the seasonal variation of the surface ozone in Athens, Greece, Atmos. Environ. 35, 2, 315-320.
- [26] Wilks, D.S., (1991), Representing serial correlation of meteorological events and forecasts in dynamic decisionanalytic models, Month. Weather Rev. 119, 1640-1662.
- [27] Wong, K.W., P.M. Wong, T.D. Gedeon and C.C. Fung, (1999), Rainfall prediction using neural fuzzy technique, URL:www.it.murdoch.edu.au/~wong/publications/SIC97 .pdf, 213-221.