Modeling Seasonality of Rainfall by Nonlinear curve Fitting to Monthly Rainfall Time Series of Jorhat

Gouri Goutam Borthakur, PhD
Assistant Professor
Department of Physics
Jorhat Institute of Science
and Technology (JIST), Assam, India

Sudeepta Pran Baruah
Undergraduate student
Department of Civil Engineering
Jorhat Engineering College (JEC)
Assam, India

ABSTRACT
In this paper, an attempt to model the temporal variability of rainfall is made by performing a time series analysis on the monthly rainfall data of Jorhat from 1994 to 2013 (excluding 2003). The monthly rainfall time series showed seasonality with a prominent frequency of 0.083 cycles per year. A curve fitting technique by nonlinear regression on the original rainfall time series and on the resulting regular residuals of the subsequent fits is performed to model the seasonality of the rainfall. The selected model is capable of showing the same seasonality and frequency of rainfall variability as that of the original rainfall time series. The selected model has the potentiality to be replicated to model rainfall in places showing similar seasonality as that of the present case.

General Terms
curve fitting, nonlinear regression, time series, seasonality, monthly rainfall

Keywords
monthly rainfall, time series, seasonality, curve fitting, nonlinear regression, Jorhat.

1. INTRODUCTION
Forecasting rainfall and modeling its variability remains crucial for agricultural sector, water resources engineering, in project planning for any construction related activity, in the tourism sector and so forth. However, given the complexity in understanding and elucidating the variability of rainfall over time and space, different approaches of rainfall modeling and forecasting have emerged. The same also varies depending on the aim of the model to be developed, that is, on whether the model is developed to give an estimation of the rainfall amount or the rainfall probability. It is also worth noting that the type of model developed also depends on the amount and type of rainfall data (daily, monthly, yearly etc.) available. Although reviewing all existing approaches to forecast and model rainfall is beyond the scope of this paper, a wide array of approaches mostly used in forecasting and modeling rainfall is reviewed in this section with a greater emphasis on time series analysis.

One of the approaches of rainfall forecasting and modeling is the time series analysis which basically results in a model of precipitation probability. Time series analysis has been performed widely by adopting different models. For eg., P.E. Naill et al. (2009) [1] performed time series analysis using the ARIMA model to forecast rainfall for ten years; A.R. Abdul-Aziz et al. (2013) [2] used the SARIMA model to study the temporal variability of rainfall from 1994 to 2010 in Ashanti region in Ghana. Other stochastic methods such as development of Poisson Cluster models also exist for time series analysis of rainfall. C. Onof et al. (2000) [3] made a

review of the developments in rainfall modeling using Poisson Cluster processes. Yet another approach to time series analysis of rainfall is the curve fitting technique. For eg., Suhaila and Jemain (2009) [4] used the Fourier series as the periodic function to fit a smooth curve to the mean rainfall per day to model daily rainfall amounts. Artificial Neural Network (ANN) approach has also been used as part of time series analysis of rainfall. N.Q. Hung et al. (2009) [5] developed an ANN model to forecast real time rainfall and flood management in Bangkok by four years of hourly rainfall data from 75 gauge stations; Hamada et al. (2013) [6] developed an ANN model to predict the summertime rainfall anomaly of Sahel etc. Additionally, Chatfield (1996) [7] in his book has discussed various approaches to time series analysis in the time and frequency domain. Shumway and Stoffer (2011) [8] have also been discussed similar approaches and models. Kedem and Fokianos (2002) [9] discussed the regression models in time series analysis.

The next broad class of approach in the forecasting and modeling of rainfall is the use of statistical distributions. Different researchers have used different distributions such as Gumbel, Log-Gumbel, Normal, Log-Normal, Pearson, Log Pearson, Exponential, Gamma, Poisson, Gumbel Weibul and the Tweedie family of distributions to model both the occurrence and quantity of precipitation (e.g., Olofintoye et al, 2009 [10]; Olumide et al, 2013 [11]; Oseni et al, 2013 [12]; Berthe et al, 2015 [13]; Dunn, (2014) [14]. Water Resources Systems Planning and Management (2005) [15] also well documents the use of various statistical distributions in modeling natural phenomena.

Regression techniques are also used for early forecast and modeling of rainfall. For eg., Nikhil Sethi et al (2014) [16] and Tahbilder et al (2014) [17] used Multiple Linear Regression. M. Rajeevan et al (2006) [18] used the Multiple Linear Regression and Project Pursuit Regression to model and predict rainfall occurrences. Boochabun [19] used the Dynamic Harmonic Regression.

Other approaches found to be used by researchers in forecasting and modeling rainfall are the Fuzzy Logic (e.g., Jimoh et al, 2013 [20]) and Neuro Fuzzy approach (e.g., Luenam et al, 2010 [21]).

In this paper, the authors use a nonlinear curve fitting technique to model the temporal variability of rainfall of Jorhat, India by performing a time series analysis on the monthly rainfall data of Jorhat from 1994 to 2013 (excluding 2003) sourced from the Indian Meteorological Department (IMD). The method employed is powerful to yield satisfactory results and easily replicable to be used in the study of similar cases.

2. STUDY AREA

Jorhat is a city in the northeastern state of Assam, India with latitude 26°44′N and longitude 94°12′E and total geographical area of 2859.35 km^2. The place experiences South-West monsoons and records an average annual rainfall of 198.76 cm. Rainfall remains low in the month of January and gradually increases till the month of July-August and then starts to decline from the month of September. This pattern then repeats for the subsequent year starting from January. The overall pattern of variation remains the same although some variation occurs from year to year (Baruah, 2014 [22]).

The economy of Jorhat is basically agrarian and a large percentage of cultivated land of Jorhat remains rain fed. Rice is the major crop to be grown and most of its cultivation is to be found in the rain fed areas (Bhowmik et al, 2001 [23]). Besides rice, tea is also cultivated in the district whose production is also chiefly dependent on rainfall availability. As such, besides many other purposes, understanding the temporal variability of rainfall of Jorhat remains crucial for the economy of Jorhat. Although the above seasonal variation of rainfall in Jorhat is acknowledged, there exists no model, mathematical, statistical or otherwise that shows the variation in a tangible manner.

3. MATERIALS AND METHODS

This section of the paper is divided into two sub sections, Data and Methodology. The data section gives details on the source and amount of data. The methodology section elaborates the approach used in the analysis of the rainfall data in the present work and elucidates the curve fitting procedure by nonlinear regression in OriginPro ®.

3.1 Data

Monthly rainfall data of Jorhat for the purpose of the analysis was sourced from the Indian Meteorological Department (IMD) as available from the websites www.indiawaterportal.org/data and www.imd.gov.in/section/hydro/distrainfall/districtrain.html. 228 months of rainfall data were collected from 1994 to 2013. The rainfall data of Jorhat for the year 2003 were unavailable and these missing data were removed from the time series analysis. Modeling the seasonal pattern of Jorhat rainfall being the prime objective and as the seasonal pattern followed by the rainfall is quite similar every year as evident from the time plot of the rainfall data (Figure 1), the missing monthly rainfall data of a particular year (2003) is not considered to hamper the analysis and its results.
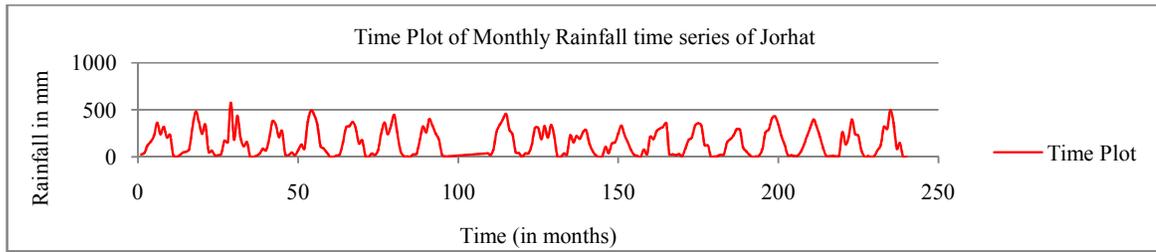
Figure 1: Time plot of monthly rainfall time series of Jorhat of 228 months from 1994 to 2013. The data of 2003 are missed and removed from the plot. The time plot shows seasonality in the time series. Data source: Indian Meteorological Department (IMD)
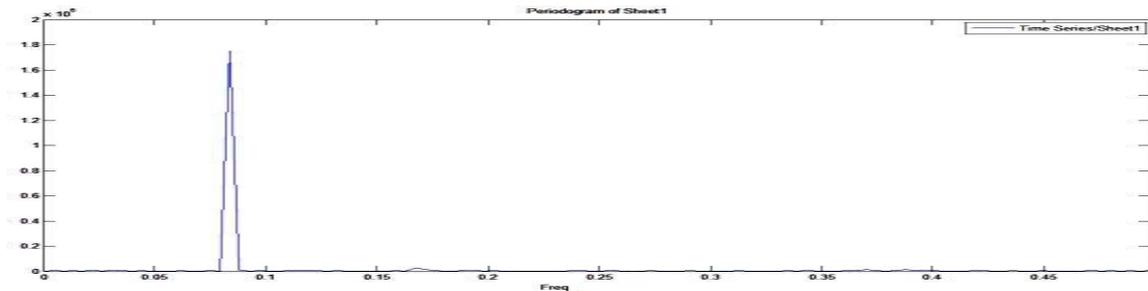
Figure 2: Periodogram of the time series generated from MATLAB ®. The periodogram shows maximum probability at frequency of 0.083 cycles per year, that is, the time series is periodic of period 12 months.

3.2 Methodology

The first step to understanding a time series, which is a collection of observations made sequentially in time, is the generation of the time plot. The time plot can be generated by plotting the observations against time. The time plot of the monthly rainfall data of Jorhat of 228 counts are shown in Figure 1. The time plot shows seasonality in the time series where rainfall increases from the beginning of a year reaches a maximum approximately in the mid year and then decreases either gradually or in a fluctuating manner towards the end of a year. A periodogram (Figure 2) of the time series, constructed in MATLAB ® showed that the time series is mostly periodic of period 12 months.

The waveform variability of the time series with a distinctive period indicates at its probability of being modeled as a periodic function of nonlinear nature by performing curve fitting to the time series. A sine fit was hence performed to the monthly rainfall time series by nonlinear regression in OriginPro ® and a nonlinear model of the time series was formulated.

A general nonlinear model may be understood as one as follows:

$$Y = f(X, \theta) + \varepsilon \qquad (1)$$

where, Y is the dependent variable; $X = x_1, x_2, x_3, \dots, x_n$ are the independent variables; $\theta = \theta_1, \theta_2, \theta_3, \dots, \theta_n$ are the

parameters and ε is the error. Determining a nonlinear model by curve fitting to a set of data points is a problem of nonlinear regression and requires the estimation of the parameters by iterative means for which the fit would converge (Motulsky et al, 1987 [24] and Bates and Watts, 1988 [25]) or which best describe the data. The iterative process of parameter estimation consists of initial approximation followed by successive approximation of the parameter values using different algorithms (Motulsky et al, 1987 [24]). A Newton-Raphson type algorithm may be used in estimating parameter values in nonlinear regression by iterative method (Kedem and Fokianos, 2002 [9]). As this iterative process of estimating the best values of the parameter cannot be performed infinitely, the process terminates when a certain criteria is fulfilled (Motulsky et al, 1987 [24]) as per the algorithm used in the process.

OriginPro ® uses the Levenberg-Marquardt (LM) algorithm in performing curve fitting by nonlinear regression. The LM algorithm is a combination of the steepest descent and the Gauss-Newton method. It behaves as the steepest descent method when the current solution is far from the correct one and as Gauss-Newton when the current solution is close to the correct one (Marquardt, 1963 [26]). The LM algorithm may be given by the following equation

$$[J^T W J + \lambda \, diag(J^T W J)]h_{lm} = J^T W (y - \hat{y}) \qquad (2)$$

$\lambda \to 0$ in the above equation results in Gauss-Newton method and large values of λ results in the gradient descent method. In the above equation, λ is the algorithmic parameter; y is the measured data; $\hat{y}$ is the curve fit function, W is the weighting matrix; J is the Jacobian and h_{lm} is the perturbation.

Starting from some initial value of the parameters, OriginPro ® uses the LM algorithm to minimize the Chi-square value (χ^2) with each iteration and when the difference of the χ^2 value of two subsequent iterations is less than a tolerance value, an absolute minimum value of χ^2 is said to have reached and the fit converges, that is, it is the criteria on which the iterative procedure terminates. The tolerance value may be given as follows:

$$Tolerance = \left| \frac{\chi^{2\prime} - \chi^2}{\chi^{2\prime} + \chi^2} \right| \qquad (3)$$

where χ^2 is the chi-square value of the current iteration and the $\chi^{2\prime}$ is the chi-square value of the last iteration. The parameter values corresponding to the absolute minimum value of χ^2 is reported. Parameter initialization to initiate the fitting process may be done automatically in OriginPro ® or user defined value may also be given as input.

The model so formulated by curve fitting once to the monthly rainfall time series in OriginPro ® showed poor performance with its standardized residuals violating the assumptions of normality and constant variance unlike in the case of the residuals from a good model. The standardized residuals vs. predicted value plot (Figure 6) however inferred that the model could be remedied by performing the same curve fitting procedure to the regular residuals of the fit. This procedure of performing the curve fitting to the resulting regular or ordinary residuals was repeated until the standardized residuals of the consequently resulting fits were found to satisfy the assumptions of normality, constant variance and independence as in the case of residuals from a good model. Consequently, four curve fitting procedures were performed to obtain a satisfactory model for the rainfall time series. Thus, Fit 1 is the nonlinear sine fit done to the original monthly rainfall time series; Fit 2 is the nonlinear sine fit done

to the resulting regular residuals of Fit 1; Fit 3 is the nonlinear sine fit done to the resulting regular residuals of Fit 2 and Fit 4 is the nonlinear sine fit done to the resulting regular residuals of Fit 3 and the rainfall model is formulated as an assemblage of the four fits.

The rainfall model thus formulated is presented and the goodness of fit corresponding to each curve fitting procedures performed on the original monthly rainfall time series and on the resulting regular residuals are presented in the next section.

4. RESULTS AND DISCUSSIONS

This section is divided into two subsections. The first section presents the formulated rainfall model along with its parameter values and standard errors. The second section deals with the Goodness of Fit (GoF) of the formulated nonlinear model. The Goodness of Fit of the formulated model is analyzed by the inspection of the residuals of the four fits from which the rainfall model is developed.

4.1 Formulated model and parameter values

The formulated rainfall model can be written as follows:

$$\begin{aligned}
y = y_1 + y_2 + y_3 + y_4 &+ A_1 sin\left(\pi\left(\frac{x - x_{c1}}{w_1}\right)\right) \\
&+ A_2 sin\left(\pi\left(\frac{x - x_{c2}}{w_2}\right)\right) \\
&+ A_3 sin\left(\pi\left(\frac{x - x_{c3}}{w_3}\right)\right) \\
&+ A_4 sin\left(\pi\left(\frac{x - x_{c4}}{w_4}\right)\right)
\end{aligned} \qquad (4)$$

where, y_1, y_2, y_3, y_4 are the vertical offsets, A_1, A_2, A_3, A_4 are the amplitudes, $x_{c1}, x_{c2}, x_{c3}, x_{c4}$ are the phase shifts and w_1, w_2, w_3, w_4 are the periods of the curve represented by the model.

Parameter values of the formulated model are presented in Table 1 along with the standard error of each of the parameters for the model.

Table 1. Parameter values for the four fits

PARAMETER VALUES	STANDARD ERRORS
y_1 155.92543	9.01055
y_2 0.35119	8.29646
y_3 -0.11756	4.8461
y_4 0.83438	4.4274
A_1 51.05649	12.76313
A_2 25.51195	12.59873
A_3 158.73513	6.85164
A_4 -41.7205	6.27037
x_{c1} 0.90692	1.00783
x_{c2} -46.89632	2.15371
x_{c3} -8.42408	0.17214
x_{c4} 1.75791	0.59862

w_1	6.46277	0.0465
w_2	5.32655	0.06286
w_3	6.00441	0.00698
w_4	6.44451	0.02773

4.2 Goodness of fit of the formulated models

A model formulated by fitting a curve to a set of data is only a close approximation of the data and is never its exact representation. Every model has residuals or errors that differentiate it from the actual data set. However, how well a formulated model represents the set of actual data is often quantified by the Goodness of Fit (GoF) statistics of the model. For models formulated by linear regression, there exist various parameters such as R^2 or χ^2 to quantify the Goodness of Fit. These are also reported alongside the curve fitting results when performed with any standard curve fitting or statistical software. However, the use of the same parameters to quantify the goodness of a fit performed by nonlinear regression is objected in many literatures along with the reasons thereof (Spiess and Neumeyer, 2010 [27] and Andrae, 2010 [28]).

Considering the incompetency of the R^2 and χ^2 to quantify the goodness of a fit performed by nonlinear regression, the authors resorted to check the goodness of the formulated model by the inspection of the residuals resulting from each of the four curve fitting operations. Many literatures suggest the analysis or the inspection of residuals to determine the Goodness of Fit of a model (Motulsky, 1987 [24]; Andrae, 2010 [28] and Tsai et al, 1998 [29] etc.). The assumptions in the residual analysis being that the residuals of a good model are independent, normally distributed and has mean, $\mu = 0$ and constant variance, σ, that is, homoscedastic.

Residual Lag Plots of the standardized residuals (Figure 3) resulting from each curve fitting operation are constructed to check for the independence of the residuals. Kolmogorov-Smirnov normality test was performed on those standardized residuals at 0.05 significance level. Test results (Table 2) shows for the first two curve fitting procedures (Fit 1 and Fit 2), the residuals were not drawn from a Normal distribution at 0.05 significance level. However, for the last two curve fitting procedures (Fit 3 and Fit 4), the residuals were drawn from a Normal distribution at 0.05 significance level. Histogram and Q-Q plots (Figure 4 and Figure 5) of the standardized residuals corresponding to each of the four curve fitting procedures are also constructed to have a visual representation on the normality of the residuals. Scatter plots of the standardized residuals vs. fitted value (Figure 6) are constructed to check if the residuals of the models are homoscedastic.

Table 2. Normality test on the residuals and mean and variance of the residuals

MODELS	KOLMOGOROV-SMIRNOV NORMALITY TEST AT 0.05 SIGNIFICANCE LEVEL
FIT 1	Residuals are not from a normal distribution
FIT 2	Residuals are not from a normal distribution
FIT 3	Residuals are from a normal distribution
FIT 4	Residuals are from a normal distribution

From Figure 3 it follows that the assumption of the independence of the residuals is satisfied for the standardized residuals of the third and forth curve fitting procedures (Fit 3 and Fit 4) where the Residual Lag plots show randomness and no definite pattern. The residuals of the first two fits (Fit 1 and Fit 2) violate the assumption of independence of residuals as a certain pattern is evident from their Residual Lag plots. The histogram of the residuals (Figure 4) for the first two fits (Fit 1 and Fit 2), do not follow a bell shape; however, the histograms of the third and forth fit (Fit 3 and Fit 4) follow a bell shape indicating that the residuals of those fit follow a normal distribution. The Q-Q plots of the residuals in Figure 5 also show that the assumption of normality of the residuals is violated in the case of the first two fits (Fit 1 and Fit 2) but satisfied in the case of the other two fits (Fit 3 and Fit 4). These are also consistent with the K-S normality test on the residuals (Table 2). Scatter plots of the standardized residuals vs. the fitted or predicted values of the models (Figure 6) is random only in the case of Fit 4 which implies that the assumption of constant variance is violated by all the three curve fits (Fit 1, Fit 2 and Fit 3) except the last fit, that is, Fit 4. That is, the residuals of the first three fits are heteroscedastic while only the residuals of the last curve fit is homoscedastic.
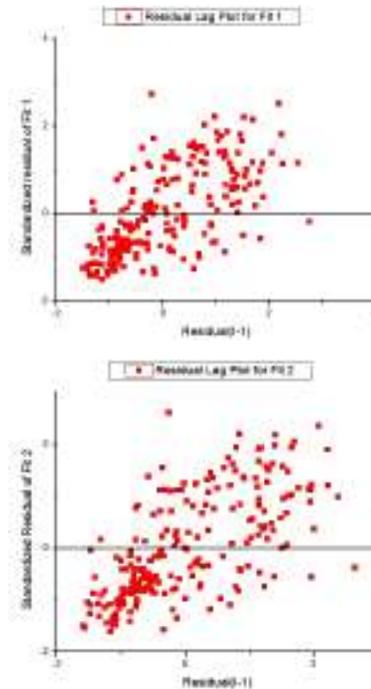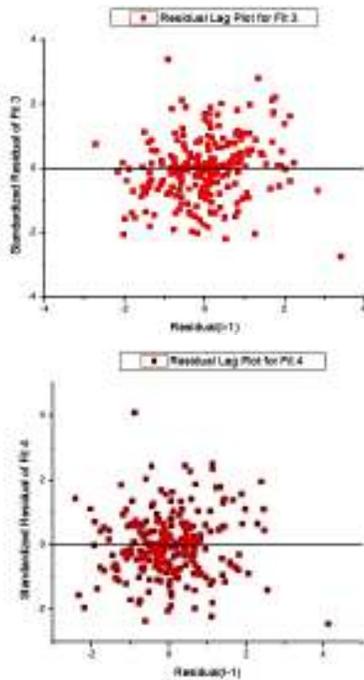
Figure 3: Residual Lag Plots for the four fits. Randomness in the plots is seen in case of Fit 3 and Fit 4 which satisfies the assumption of independence of the residuals.
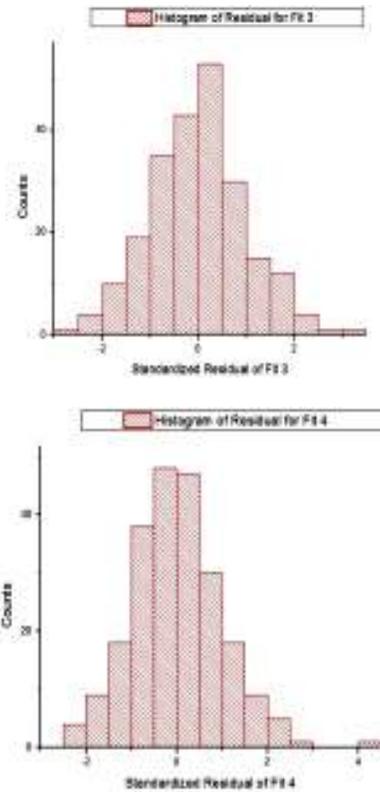
Figure 4: Histogram of the residuals for the four fits. The histogram of the residuals of Fit 3 and Fit 4 follow a bell shape which suggests that those residuals follow a normal distribution. Same is not true for Fit 1 and Fit 2 as seen from the plots.
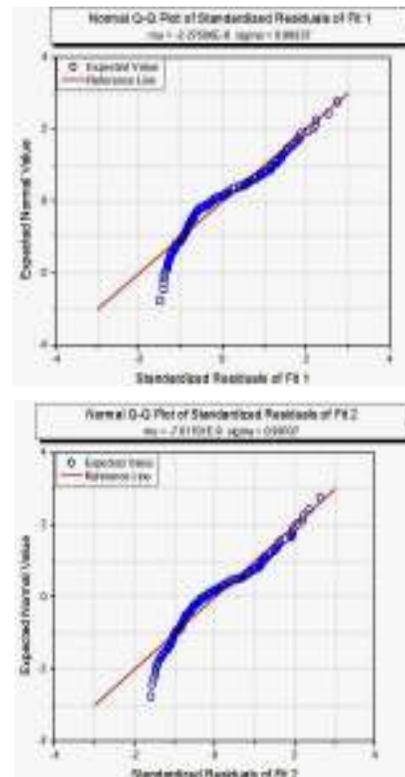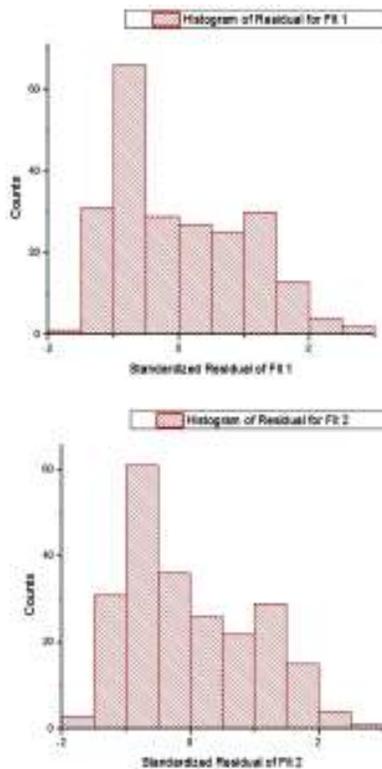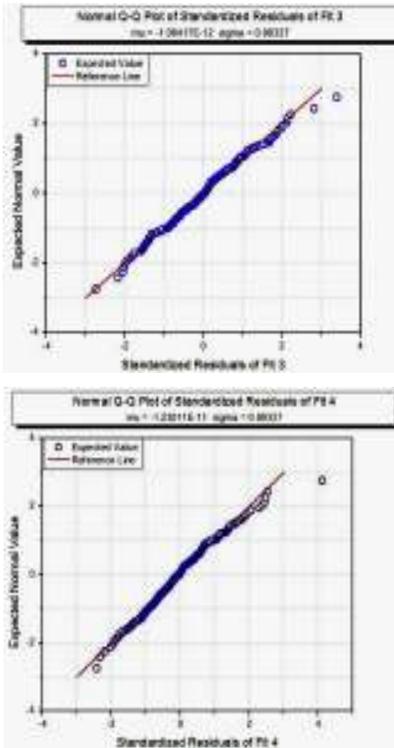
Figure 5: Quantile – Quantile (Q-Q) Plots for the residuals of the Fits. The Q-Q plots of the residuals of Fit 3 and Fit 4 follow a straight line which suggests that those residuals follow a normal distribution. The same is not true for residuals of Fit 1 and Fit 2 as seen from the plots.
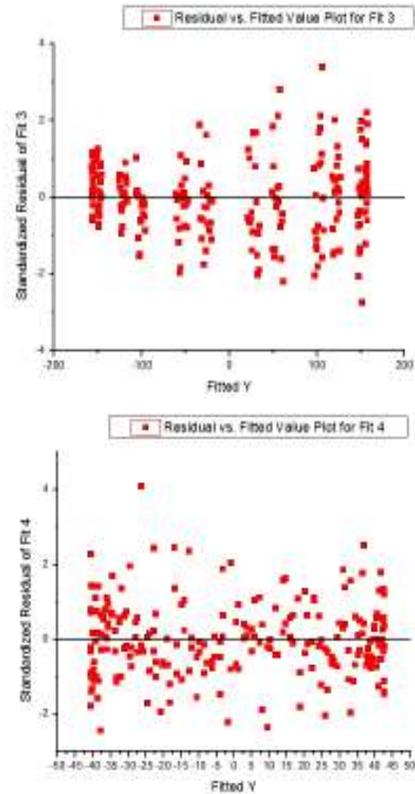
Figure 6: Residual vs. Fitted value plots for the fits. Only the plot of Standard residuals vs. Predicted value for Fit 4 shows randomness and hence satisfies the assumption of constant variance of residuals. All other fits violate this assumption as implied by the plots.
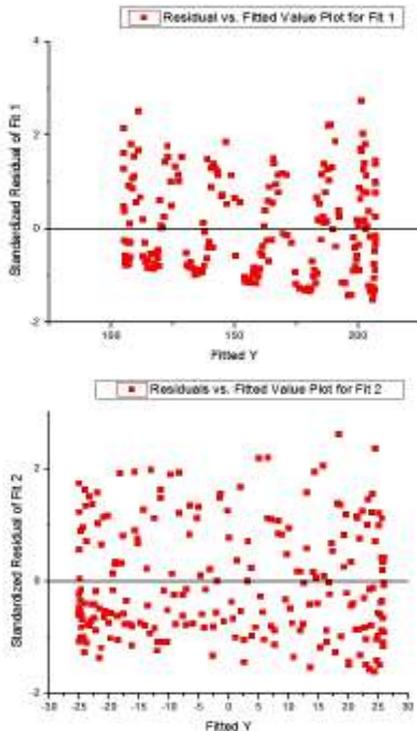
It thus follows from the analysis of the residual plots that the model formulated by nonlinear curve fitting on the monthly rainfall time series once, that is, Fit 1 is not a satisfactory interpretation of the rainfall time series as it violates the assumption of normality and constant variance of the residuals. The residual vs. predicted value plot for Fit 1 (Figure 6) possesses a waveform (nonlinear) pattern and hence a second nonlinear sine fit (Fit 2) was performed to the regular residuals of Fit 1 in order to improvise the model. The same observations were made for the residuals of Fit 2 and thus another nonlinear sine fit (Fit 3) was performed to the regular residuals of Fit 2 in order to further improvise the model. The residuals of Fit 3 however satisfied the assumption of independence and normality but violated the assumption of constant variance. The wedge shaped pattern of the residuals vs. predicted value plot showed that the variance of the residuals of Fit 3 increases with the mean. Repeating the same nonlinear curve fitting procedure to the regular residuals of Fit 3 resulted in Fit 4. The residuals of Fit 4 satisfied all the assumptions of independence, normality and constant variance as can be seen from the residual plots. Further fitting to the resulting regular residuals was thus terminated and the model of the rainfall time series consequently established is given in equation (3). The time plot of the original and modeled rainfall time series is shown in Figure 7 and the periodogram of the modeled rainfall time series is shown in Figure 8. It can be seen from Figure 7 that the plot of the modeled monthly rainfall time series gives a good approximation of the original rainfall time series with a smooth curve and shows the same trend of temporal

variability in the rainfall. From Figure 8 and on comparing it with Figure 2, it is further evident that the model shows the

same seasonality of rainfall as in the case with the original monthly rainfall time series.
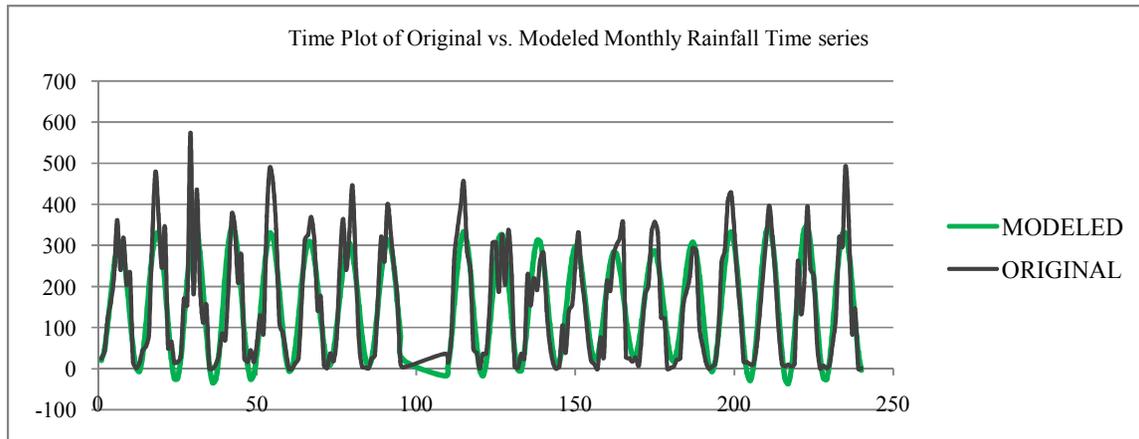
Figure 7: Time Plot of Original vs. Fitted Monthly Rainfall Time series. The model fits well to the original rainfall time series
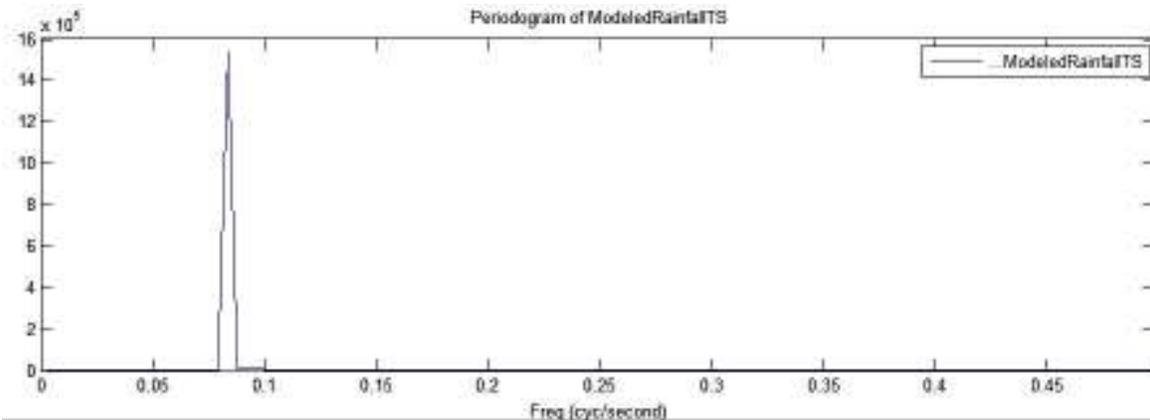
Figure 8: Periodogram of modeled monthly rainfall time series. The periodogram shows maximum probability at the same frequency of 0.083 cycles per year as in the case of the original monthly rainfall time series.

5. CONCLUSION

Nonlinear regression is a robust method for analysis of data which although is difficult to be performed manualy can be performed with comparative ease by the use of computer programs. It requires a clear perception of the equation to be used in the regression and the physical model it represents. The results of the nonlinear regression should be carefully analyzed and visualized graphically (Motulsky, 1987 [24]). In this paper, the seasonality in the monthly rainfall time series of Jorhat is observed and the same is hence modeled by performing sine fits by nonlinear regression in OriginPro ®. The goodness of fits is scrutinized by the analysis of residuals of the fits graphically and also by KS normality tests. The formulated model is an assemblage of four nonlinear sine fits (Fit 1, Fit 2, Fit 3 and Fit 4). The formulated model is found to show the same seasonality and frequency of rainfall variability as that of the original monthly rainfall time series. The model or the method followed in this paper might find utility in modeling monthly rainfall for places having a similar rainfall pattern as in the case of Jorhat.

6. REFERENCES

[1] Momani, P.E. "Time Series Analysis Model for Rainfall Data in Jordan: Case Study for Using Time Series Analysis". *American Journal of Environmental Sciences,* vol. 5 (5), pp. 599-604, 2009.

[2] Aziz, A.A.R.et al. "Modeling and Forecasting Rainfall Pattern in Ghana as a Seasonal Arima Process: *The Case of Ashanti Region*" *International Journal of Humanities and Social Science,* vol. 3 (3), pp. 224-233, Feb. 2013

[3] Onof, C et al. "Rainfall modeling using Poisson-cluster processes: a review of developments". *Stochastic Environmental Research and Risk Assessment,* vol. 14 (2000), pp. 384-411, 2000

[4] Suhaila, J et al. "A comparison of the rainfall patterns between stations on the East and the West coasts of Peninsular Malaysia using the smoothing model of rainfall amounts". *Meteorological Applications,* vol. 16, pp. 391-401, April, 2009

[5] Hung, N.Q. et al. "An artificial neural network model for rainfall forecasting in Bangkok, Thailand". *Hydrology and Earth System Sciences*, vol. 13, pp. 1413-1425, 2009

[6] Badr, H.S. "Application of Statistical Models to the Prediction of Seasonal Rainfall Anomalies over the Sahel". *Journal of Applied Meteorology and Climatology,* vol. 53, pp. 614-636, Oct. 2013

[7] Chatfield, C. *THE ANALYSIS OF TIME SERIES AN INTRODUCTION.* CHAPMAN AND HALL/CRC, 1996

[8] Shumway, R.H and Stoffer, D.S. *Time Series Analysis and Its Applications With R Examples.* Springer, 2011

[9] Kedem, B and Fokianos. *Regression Models for Time Series Analysis.* Wiley Interscience, 2002

[10] Olofintoye, O.O et al. "Best–fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria". *New York Science Journal,* vol. 2 (3), 2009

[11] Olumide, B.A et al. "Evaluation of Best Fit Probability Distribution Method for the Prediction of Rainfall and Runoff Volume (Case Study of Tagwai Dam, Minna-Nigeria)". *International Journal of Engineering and Technology,* vol. 3 (2), pp. 94-98, Feb.2013

[12] Oseni et al. "Fitting the Statistical Distribution for Daily Rainfall in Ibadan, Based On Chi-Square and Kolmogorov-Smirnov Goodness-Of-Fit Tests". *West African Journal of Industrial and Academic Research,* vol. 7 (1), pp. 93-100, Jun. 2013

[13] Berthe, K.A et al. "Gumbel Weibul distribution function for Sahel precipitation modeling and predicting: Case of Mali". *African Journal of Environmental Science and Technology,* vol. 9 (5), pp. 405-412, May.2015

[14] Dunn, P.K. "Occurrence and quantity of precipitation can be modeled simultaneously". *International Journal of Climatology,* vol. 24, pp. 1234-1239, May. 2004

[15] *Water Resources Systems Planning and Management.* UNESCO, 2005

[16] Sethi, N and Garg, K. "Exploiting Data Mining Technique for Rainfall Prediction". *International Journal of Computer Science and Information Technologies,* vol. 5 (3), pp. 3982-3984, 2014

[17] Dutta, P.S and Tahbilder, H. "Prediction of Rainfall using Datamining Technique over Assam". *Indian Journal of Computer Science and Engineering,* vol. 5 (2), Apr-May. 2014

[18] Rajeevan, M et al. "New statistical models for long-range forecasting of southwest monsoon rainfall over India". *Clim Dyn,* Sep. 2006

[19] Boochabun, K et al. "Statistical Modeling of Rainfall and River Flow in Thailand" *Journal Geological Society of India.* Available: http://www.es.lancs.ac.uk/people/nickc/boochabun.pdf

[20] Jimoh, R.G. et al. "Modeling Rainfall Prediction using Fuzzy Logic". *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 1 (4), pp. 929-936, Jun.2013

[21] Luenam, P et al. "A Neuro-Fuzzy Approach for Daily Rainfall Prediction over the Central Region of Thailand". *Proceedings of the International MultiConference of Engineers and Computer Scientists, 2010,* vol. I, Mar.2010

[22] Baruah, M.B. "Investigation of Rainfall Trend in Jorhat Town, Assam, India". *European Academic Research,* vol. II (7), pp. 8945-8954, Oct.2014

[23] Bhowmik, B.C. et al. "Socio-economic dynamics of rice production system in eastern India" presented at the 4th workshop of the NATP project RRPS at the Central Rice Research Institute, Cuttack, 2001

[24] Motulsky, H.J et al. "Fitting curves to data using nonlinear regression: a practical and non mathematical review". *The FASEB Journal*, vol. 1 (5), pp. 365-374, Nov. 1987

[25] Bates, D.M and Watts, D.G. *Nonlinear Regression Analysis and Its Applications.* Wiley Interscience, 1988

[26] Marquardt, D.W. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters". *Journal of the Society for Industrial and Applied Mathematics*, vol. 11 (2), pp. 431-441, Jun.1963

[27] Spiess, A.N and Neumeyer, N (2010). "An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach". *Spiess and Neumeyer BMC Pharmacology.* Available: http://www.biomedcentral.com/1471-2210/10/6

[28] Andrae, R et al (2010, Dec.). "Dos and don'ts of reduced chi squared". *arXiv:1012.3754v1 [astro-ph.IM].* Available: http://arxiv.org/pdf/1012.3754

[29] Tsai, C-L et al. "The Examination of Residual Plots". *Statistica Sinica,* vol. 8, pp. 445-465, 1998