International Journal of Innovative Research in Science and Engineering
Vol. No.3, Issue 02, February 2017
www.ijirse.com
ISSN (O) 2454-9665
ISSN (P) 2454-0663

DEVELOPMENT OF EMPIRICAL MODELS FOR RAINFALL PREDICTION

Ashwanth Athi[1], N Deepa[2], Balakrishnan Athiyaman[3]

[1,2]School of Information Technology and Engineering, Vellore Institute of Technology (VIT), Vellore
[3] National Centre for Medium Range Weather Forecasting, Noida, India

ABSTRACT

Weather plays an important role in every aspect of the economy from commerce to travel to agriculture to manufacturing etc. Being a natural non-linear phenomenon, predicting rainfall or precipitation is a challenging task for the climate and weather researchers. Accurate prediction implies predicting the exact amount both spatially and temporally, which is complicated because of the multiple weather process and patterns or systems involved. India, primarily being an agriculture economy, accurate rainfall prediction is all the more important for overall wellbeing of the nation.

The present work is an attempt to develop and evaluate three common statistical techniques like: Multiple Linear Regression (MLR), Data Mining (DM) and Neural Networks (NN) using observations from Automatic Weather Stations (AWS) collected from National Climatic Data Centre (NCDC) for the period 2000-2014. The three statistical models were developed using SPSS, WEKA &Java and MATLAB, respectively.

Past works have commonly used five critical meteorological parameters like: Temperature, Dew Point, Mean Sea-level Pressure (MSLP), Wind speed, Humidity to correlate it with precipitation. In the present study, the data and the techniques were evaluated for four selected cities in India located in four different geographical regions of the country. The results were quite interesting and the rainfall prediction made through Neural Network for all four regions were reasonably accurate compared to the other models.

I. INTRODUCTION

India is primarily an agriculture based economy and an agrarian society for which weather plays an important role in determining the success or the failure of agriculture and also agro-related enterprises. According to Government of India's Economic Survey 2012, agriculture and its related sectors contribute to 17% to the national gross domestic product (GDP), while employing 51% of the total national workforce. The country is heavily dependent on the agricultural usage and particularly on rains for harvesting. Nearly 65% of the total population in rural area are dependent on agricultural related activities, and a significant percentage of rural population is indirectly dependent on the throughput from the cultivable land. An important point to note is that the amount of cultivable land too is a variable of rainfall, as only a small portion of cultivatable area is today under canal irrigation and most areas are rain-fed.Hence a reliable and accurate rainfall prediction would help farmers to plan their agriculture activities well in advance and takes precaution of heavy or scanty rainfall. An

accurate rainfall prediction (from weekly to seasonal time scales) would also help many utility departments like: power, civil aviation, shipping, fisheries, space program etc. to plan their activities well in advance.
To help the numerical weather forecasting agencies and as part of wider National and World Meteorological Organization (WMO) goals, India has been continuously upgrading its automatic weather stations (AWS) both qualitatively and quantitatively. As of today, there are a total of 550 AWSs operating in the country of which 127 are exclusively Agro-met stations (Ranalkar, et al. 2010), along with 1350 automatic rain gauges (ARGs) stations.

Nearly 80% of the total annual rainfall in the country is due to the Indian Summer Monsoon (ISM) or the Southwest monsoon, which is a major part of the global circulation feature called the Asian Summer Monsoon (ASM).There is a large variability in the rainfall during the ISM due to the forcing from sea surface temperature (SST) of primarily Indian Ocean and internal dynamics of the ISM dictated and modulated by the Madden-Julian Oscillation (MJO) and the monsoon intra-seasonal oscillations (ISO)(Ajaya Mohan and Goswami 2003). Due to this complex nature of the ISM, India Meteorological Department (IMD) has been using statistical/empirical methods to provide long range forecasts(LRF) of the rainfall due to ISM since their first introduction by Blanford (1884).Since then many researchers have proposed various statistical techniques, of which major works are: Autoregressive Integrated Moving Average (ARIMA) [(Thapliyal 1981); and(Tektas 2010)]; multiple linear regression [Shukla and Mooley, 1987]; power regression model [Gowrikar et al., 1991; and Rajeevan et al., 2004]; and artificial neural networks [Goswami and Srividya, 1996; Goswami et al., 1999; and Sahai et al., 2000]; singular value decomposition (SVD) and principal component analysis (PCA) [Mohanty et al., 2013].

The present paper presents results from the application of three standard techniques for forecasting rainfall: multiple linear regression (MLR); data mining (DM) and neural networks (NN) to the observation data obtained from AWS. The following section briefly summarizes the relevant literature survey; while the third section describes the study area. The fourth section presents the adopted methodology while the fourth section presents the results.The final section discusses these results and presents a conclusion on the selected techniques.

II. LITERATURE SURVEY

Murphy (1998) summarized the history of probabilistic forecasts by classifying the developments as pre-and post-synoptic era. The pre-synoptic era lacked the infrastructure of weather stations and predictions were made mostly based on probability as odds of raining or not. Again the post-synoptic era presented by Murphy (1998) can be classified into two phases as: pre-and post-application of numerical era, roughly before and after 1925.The present section briefly presents the past studies on weather prediction using statistical regression or MLR, NN and data mining. Any prediction or forecast can be accomplished from and by the data (statistical or empirical) or from the dynamics that produced the data (modern numerical weather prediction (NWP) models based on physical laws). This literature survey does not encompass the literature on NWP models or the dynamics, but focuses on numerically simpler statistical techniques and is primarily confined to post-synoptic era only.

ISSN (O) 2454-9665
ISSN (P) 2454-0663

A report on long range weather forecasting by the committee on atmospheric sciences under the aegis of National Research Council of United States (Report of the Panel on Long Range Weather Forecasting 1975) found that the statistical approach to determine the weather has definite advantages over numerical models, as the former deals with real weather, rather than a simulation and linear regression methods ignore the nonlinear nature of weather dynamics. Thus linear regression methods might not be useful in predicting medium range of 1 to 5 days, but are relatively more useful for long range prediction.Shukla and Mooley (1987) highlighted the fact that empirical methods have helped in discovering important processes in the atmosphere citing the example ofWalker (1923), who explained the nature of Southern Oscillation due to fluctuations in Indian Monsoon.

In terms of statistical or empirical methods, MLR is the simplest and most commonly used forecasting method and in fact, IMD's current operational forecasting model for southwest monsoon is based on multiple regression (Rajeevan, et al. 2007). Since Walker's first objective model (Walker 1923), many studies have made used of regression techniques [Shukla and Mooley 1987; Gowariker, et al. 1989; Gowariker, Thapliyal and Kulshrestha, et al. 1991; Thapliyal and Kulshrestha 1992; Thapliyal 1997; Rajeevan 2001; Thapliyal and Rajeevan 2003 and others] for predicting the Indian Summer Monsoon. All these studies have used specific phenomena around the world as predictors instead of numerical data like: sea surface temperature (SST) over Arabian Sea, South Indian Ocean and North Atlantic and their indices and tendencies, Eurasian Snow Cover, temperatures over NW Europe, atmospheric pressure and its tendencies at selected levels over the Northern Hemisphere, and wind patterns etc.

Three prominent studies applying DM for weather prediction using observation data are:

Sivaramakrishan *et al., (*2011)used association rule and classifier approach for rainfall analysis using the data for the period of 1961-2010 for the coastal station of Cuddalore, India. They used the following meteorological parameters: *temperature, dew point, wind speed, visibility* and *rainfall* for predicting rainfall. The authors also used predictive *apriori* algorithm for generating the best rules. The outcome from the association rule mining and instance based classifier approach were found to give accurate results for one-day forecasts of rainfall.**Taksande *et al.,* (2014)**analysed commonly adopted data mining algorithms for rain fall prediction. The five commonly adopted data mining algorithms are: Neural Network (NN), Random Forest, Classification and Regression Tree (CRT), Support Vector Machine (SVM) and k-nearest neighbour. They adopted an Artificial Neural Network (ANN) and Genetic Algorithm (GA)based hybrid model for rainfall forecasting. Using data of temperature, air pressure, rainfall, humidity and wind speed, they observed that the neural nets perform the best in the group. Their work developed a Hidden Markov Model (HMM) based Genetic Algorithm and compared the predictability with the best performer of the commonly adopted data mining ANN models. They observed that the result from HMM based GA outperformed the ANN models.**Hemachandra *et al.,*(2013)**developed an algorithm to predict electric load forecasting using Neuro-Fuzzy Systems. They have discussed various techniques like Regression method, Fuzzy logic approach, Neural Network approach, Neuro-fuzzy approach etc. They suggesteda hybrid model of Neural Network and Fuzzy (Neuro-Fuzzy) for short-term load forecasting. The mean error from the Neuro-fuzzy model was found to be less than that of multiple linear regression models,

ISSN (O) 2454-9665
ISSN (P) 2454-0663

thereby indicating the superiority of the Neuro-fuzzy model vis-a-vis the multiple linear regression in forecasting at short-terms.

III. STUDY AREA

The present study focuses on observation data from India. The selected study sites are from four distinct geographic locations, with each city experiencing rainfall in different monsoon phases and exhibiting different climatic zones. The selected cities are: Patiala, Kolkata, Mumbai and Chennai, located in Northern, Eastern, Western and Southern regions of the country, respectively.

The city of Patiala is located in Punjab at about 30 °N latitude and according Köppen-Geiger climate classification (KGCCS; [Peel et al., 2007]), it is classified as *BSh*, which corresponds to *Steppe* or *semi-arid climate*. It experiences very harsh summers with temperatures touching 40 °C and pleasant winters with an average temperature of 8 °C. The average annual rainfall is about 754 mm, with most of the rains occurring in July, August and September months, i.e., SW monsoon period.

Kolkata is located at about 22 °N latitude and according to KGCCS; it is classified as *Aw*, which corresponds to *Tropical Wet and Dry Climate* or *Tropical Savanna*. The temperature ranges from a record high of 43.7 °C in June to a record low of 6.7 °C in January. The average annual rainfall is about 1735 mm, with rainfall over 100 mm occurring between May and October [IMD].

Mumbai is located at latitude of approximately 19 °N and according to KGCCS, Mumbai's climate is classified as *Am*, which corresponds to *tropical wet climate* or *tropical monsoon and trade wind littoral climate*. The coastal and tropical nature of the city modulates the temperatures, hence the mean maximum summer temperature is about 32 °C with a mean minimum winter temperature is about 30 °C, while the record maximum and minimum temperatures are 42.2 °C in April and 7.4 °C in January, respectively [IMD]. Mumbai city experiences lot of rainfall with an annual average of 2258 mm, with more than 100 mm between May and October.

The city of Chennai is located at approximately 13 °N latitude and according to KGCCS, it is classified as *Aw*. Chennai is located on the "thermal equator", with a record maximum and minimum of 45 °C in May and 13.9 °C in January, respectively. The mean annual rainfall is about 1400 mm, with peak rain season during the retreating monsoon phase (NE monsoon) during November. This is in marked contrast with the other three selected stations.

IV. METHODOLOGY

4.1 Criteria for selection of the critical parameters:

From the recent works 13 parameters were found to influence rainfall amount, of which top five independent variables (parameters) were considered based on reported correlation. They are: Temperature, Dew Point, Mean Sea Level Pressure (MSLP), Wind Speed (WS) and Humidity.

Data collection and preparation: The data collection and preparation is very important for rainfall analysis. Fifteen years data have been collected and prepared for ready to use. All five stations different data sets were used. The numbers of records are same for all five stations. The total daily data was collected for the period from (2000-2014).

4.2 Multiple Linear Regression (Step wise Regression)

Based on the literature review and the relevant research papers from 2006 to 2015 were reviewed and step-wise regression used to find correlation between the parameters. The correlation between dependent parameter and independent parameters was calculated and each station the dependent and independent parameters correlations were tabulated. Most of the researcher used the regression model for predicting the rainfall. This method is useful for evaluation and comparison purpose.

The data reqired forMultiple regression technique and step-wise algorithm was prepared in a spreadsheet (MS Excel form). Out of 15 year data 13 year was used as training data, which results in 4749 data points, while two year data was used as validation test data, which results 730 data points. Step-wise regression was performed in SPSS and from which the best fit regression models are:

CHENNAI MODEL: Y=932.191+12.961*T+4.413*H+0.518*WS-1.238*P-15.217*TD

KOLKATTA MODEL: Y=327.95+12.332+3.485+0.08*WS-0.616*P-13.458*TD

MUMBAI MODEL Y= 1263.585+9.948*T+3.732*H+0.722*WS-1.505*P-12.956*TD

PATIALA MODEL: Y=377.518+1.057*T+0.495*WS+0.488*H-0.41*P-1.321*TD

4.3 Data Mining

Data mining is one of the important tool to mining the very large volume of data. It was proved in the literature review section. The weather prediction researchers are mostly using this tool to predict the rainfall. The accuracy of the data mining prediction is comparable with the other model prediction.

The quality controlled observation data required to be discretized to facilitate the application of data mining predictive and cluster algorithm. In this process the data shall be classified and converted from numeric to nominal. An open source data mining tool Waikato Environment for Knowledge Analysis (WEKA) will be used to classify the data. The fifteen-year data was segregated into two parts training (13 years) and testing (2 year). The classification algorithm will classify and make cluster with respect to rainfall as a prognostic variable. WEKA has several apriori algorithms to generate associative rules based on the observed data. Aprori rules (Predictive rules) have been commonly applied in Business Analytics (BA) to determine the consumer behaviour based on their recent purchases. With the application of these techniques, few of the topmost rules are as follows:

- Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' WindSpeed[knots]='(4-8]' Humidity[%]='(55-70]' 229 ==> Rainfall[mm]='(0-50]' 102 acc:(0.4752)
- Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' Humidity[%]='(55-70]' 286 ==> Rainfall[mm]='(0-50]' 127 acc:(0.46681)

ISSN (O) 2454-9665
ISSN (P) 2454-0663

- Temperature[°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' WindSpeed[knots]='(8-12]' Humidity[%]='(70-85]' 26 ==> Rainfall[mm]='(50-100]' 3 acc:(0.14494)
- Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' MSLP[hPa]='(1008-inf)' Humidity[%]='(55-70]' 269 ==> Rainfall[mm]='(0]' 256 acc:(0.9510)

The interpretation of the first rule is as follows: when the temperature is between 20.06°C to 30.04°C and whenthe Dew Point is between 20.06°C to 30.04°C; and when the MSLP is between 996 hpa and 1008, and when Wind Speed between 4knots and 8knots and Humidity is between 55% and 70% then the predicted rainfall would be maximum of 50 mm with an accuracy of 47%.

An expert system search engine was developed to extract the rule that suggest rainfall with high accuracy. The search engine algorithm basically takes the observation from the validation/test data to predict the rainfall. The search algorithm sorts the obtained list of rules based on the number of parameters and the level of accuracy. For example, if two rules match, then the one with more parameter was considered as top, followed by the accuracy. The predicted rainfall was compared with the actual observation to derive RMSE values presented in Table 1.2

19-5-2010 -Three Parameter Match Rule No. :-191. Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' 916 ==> Rainfall[mm]='(50-100]' 23 acc:(0.01427)

Three Para Are :-Temperature = 24.83 MSLP = 999.8 DewPoint = 23.17

23-4-2011 - Three Parameter Match Rule No :- 80**.** Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' WindSpeed[knots]='(4-8]' Humidity[%]='(55-70]' 229 ==> Rainfall[mm]='(0-50]' 102 acc:(0.4752)

Three Para Are :-Temperature = 27.39 DewPoint = 24.67 WindSpeed = 5.0

17-10-2012 – Three Parameter Match Rule No. :- 80. Temperature [°C]='(20.06-30.04]' DewPoint [°C]='(20.06-30.04]' MSLP[hPa]='(996-1008]' WindSpeed[m/s]='(4-8]' Humidity[Degree]='(55-70]' 229 ==> Rainfall [mm]='(0-50]' 102 acc:(0.4752)

Three Para Are :-Temperature = 25.89 DewPoint = 23.33 WindSpeed = 4.6

15-02-2013 - Three Parameter Match Rule No 25. Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' WindSpeed[knots]='(0-4]' Humidity[%]='(55-70]' 123 ==> Rainfall[mm]='(0]' 107 acc:(0.85891)

Three Para Are :-Temperature = 26.33 DewPoint = 23.61 WindSpeed = 3.4

18-10-2014 - Three Parameter Match Rule No 208. Temperature[°C]='(20.06-30.04]' DewPoint[°C]='(20.06-30.04]' WindSpeed[knots]='(0-4]' 1255 ==> Rainfall[mm]='(100-150]' 12 acc:(0.00642)

Three Para Are:-Temperature = 25.56 DewPoint = 24.67 WindSpeed = 2.4

The Predictive algorithm predicted results are shown in the Table 1.XX. Using the generated rules, the predicted rainfall was estimated and compared with actual rainfall to calculate the RMSE for the model for each

ISSN (O) 2454-9665
ISSN (P) 2454-0663

station. Comparing these results with previously tested regression model, the Data mining results gives the better results compare to regression results

Data mining predictive algorithm predicts the rules with various combinations of the selcted (05) parameters and there by using the rule search engine, it predictsthe rainfall with the determined accuracy of the rule. It is observed that with a combination of three parameter of temperature,dew point and wind speed , the accuracy have been found to be 0.8589 and minimum accuracy level with a combination of three parameter Temperature, Dewpoint and Wind Speeed increases by 0.00642.

4.4 Neural Network

Neural Network is nonlinear model that is easy to use and understand compared to statistical methods. ANN is non-parametric model while most of statistical methods are parametric model that need higher background of statistic. ANN with Back propagation (BP) learning algorithm is widely used in solving various classifications and forecasting problems. Even though BP convergence is slow but it is guaranteed. However, ANN is black box learning approach, cannot interpret relationship between input and output and cannot deal with uncertainties. To overcome this several approaches have been combined with ANN such as feature selection and etc.

Based on the literature review the Neural Network model handles the bigger volume of data. Most of the authors used Neural Network methods for analysing the rainfall data. Considering the outcome of the literature review the researcher has evaluated the rainfall prediction data using the Neural Network method. The Neural Network learns the input-output relationship through the training process. The learning process in the Neural Network is an interactive procedure in which its connection weights are adapted through the presentation of a set of input-output training example pairs. The "Feed-forward back-prob" and "Traingdx" techniques will be used during the Neural Network designing and implementation. The present work made use of MSE performance function as a metric. The number of Network Neurons used in this evaluation is ten. The test results are De-Normalized and compared with the observed data and calculated the difference and MSE. The same procedure adopted all four geographical locations.

V. RESULTS AND ANALYSIS

5.1 Result of Step-wise Regression output

The data required for Multiple regression technique and step-wise algorithm was prepared in the excel form. There are 4749 data points were considered as input data. The remaining 730 data points used for validation purpose. The data base records were organised with proper field types according to the data types. The nominal and numeric fields are identified and segregated accordingly. The dependent and independent parameters were identified and supplied the parameter. Finally the step-wise algorithm was applied station wise and calculated the regression outputs.

ISSN (O) 2454-9665
ISSN (P) 2454-0663

Table 1.1: Performance of multiple regression algorithms using the test data

SL.NO	STATION	RMSE
1	CHENNAI	3.21
2	KOLKATA	2.65
3	MUMBAI	4.67
4	PATIALA	0.65

5.2 Result of Data Mining Approach

Data mining predictive algorithm predicts the rules with various combinations of the selected (05) parameters and there by using the rule predict rainfall with the determined accuracy of the rule. It is observed that with a combination of four parameter of temperature,dew point, wind speed and humidity, the accuracy have been found to be 0.462 with a combination of three parameters: Dew point, MSLP and wind speed the accuracy of prediction increase to 0.994.

The Predictive algorithm predicted results are shown in the Table 1.3 Using the generated rules, the predicted rainfall was estimated and compared with actual rainfall to calculate the RMSE for the model for each station. Comparing these results with previously tested regression model, the Data mining results gives the better results compare to regression results.

Table 1.2: Performance of DM algorithms using the test data

SL.NO	STATION	RMSE
1	CHENNAI	2.49
2	KOLKATA	1.45
3	MUMBAI	1.90
4	PATIALA	1.95

5.3 Result of Neural Network Approach

The Neural Network (NN) Model predicted the rainfall for all five stations. The NN predicted rainfall was compared with the observed rainfall of all five stations. The difference between predicted and observed rainfall has been calculated and prepared the error table. The results were tabulated in Table 1.4 and from which it can be noted that the results predicted by the Neural Network is better than the previously tested models of data mining and fuzzy exert system results.

Table 1.3: Performance of Neural Network algorithms using the test data

SL.NO	STATION	RMSE
1	CHENNAI	0.42
2	KOLKATA	0.85
3	MUMBAI	0.26
4	PATIALA	0.68

Table 1.4: Comparison of three selected algorithms using the test data

Station	Regression	Data Mining	Neural Network
Chennai	3.21	2.49	0.42
Kolkatta	2.65	1.45	0.85
Mumbai	4.67	1.90	0.26
Patiala	0.65	1.95	0.68

Table 1.5 Predicted results for the Algorithms

Station Name	Date	Rainfall[mm] Observed	DM Predicted Rainfall[mm]
Chennai	19-05-2010	78.99	9.1
Chennai	23-04-2011	19.05	10.5
Chennai	17-10-2012	20.07	11.1
Chennai	15-02-2013	11.9	4.3
Chennai	18-10-2014	115.0	9.8
Kolkata	28-06-2010	17.0	23.0
Kolkata	17-06-2011	44.9	46.0
Kolkata	29-07-2012	30.9	23.0
Kolkata	17-02-2013	9.90	0.0
Kolkata	03-09-2014	24.8	23.0
Mumbai	08-07-2010	53.0	92.2
Mumbai	07-09-2011	14.9	0.0
Mumbai	07-09-2012	22.0	0.0
Mumbai	03-08-2013	9.90	46.1
Mumbai	29-07-2014	48.0	46.1
Patiala	08-08-2010	3.04	34.79

Patiala	09-09-2011	33.0	34.79
Patiala	14-12-2012	14.9	17.39
Patiala	05-02-2013	13.97	17.39
Patiala	18-03-2014	8.89	0.0

Table: 1.6: Chennai (India) and Mannar (Srilanka) validation

Date (Random)	INDIA	Actual Rainfall observed	Rainfall predicted by Neural Network	SRILANKA	Validated with other countries
07-09 -2013	Chennai	0.0	2.26	Mannar	2.91
02-11-2013	Chennai	3.05	0.88	Mannar	2.61
11-11-2013	Chennai	4.0	0.099	Mannar	0.70
19-11-2013	Chennai	3.05	0.096	Mannar	1.79
13-12-2013	Chennai	0.0	0.100	Mannar	2.59

Table: 1.7: Mumbai (India) and Karachi (Pakistan) validation

Date (Random)	INDIA	Actual Rainfall observed(mm)	Rainfall predicted by Neural Network	PAKISTAN	Validated Rainfall Prediction (mm)
03-08-2013	Mumbai	9.90	16.16	Karachi	21.7
27-09-2013	Mumbai	13.97	5.71	Karachi	15.7
10-10-2013	Mumbai	3.0	0.690	Karachi	4.7
03-03-2014	Mumbai	5.0	7.105	Karachi	43.3
02-08-2014	Mumbai	0.0	0.158	Karachi	3.5

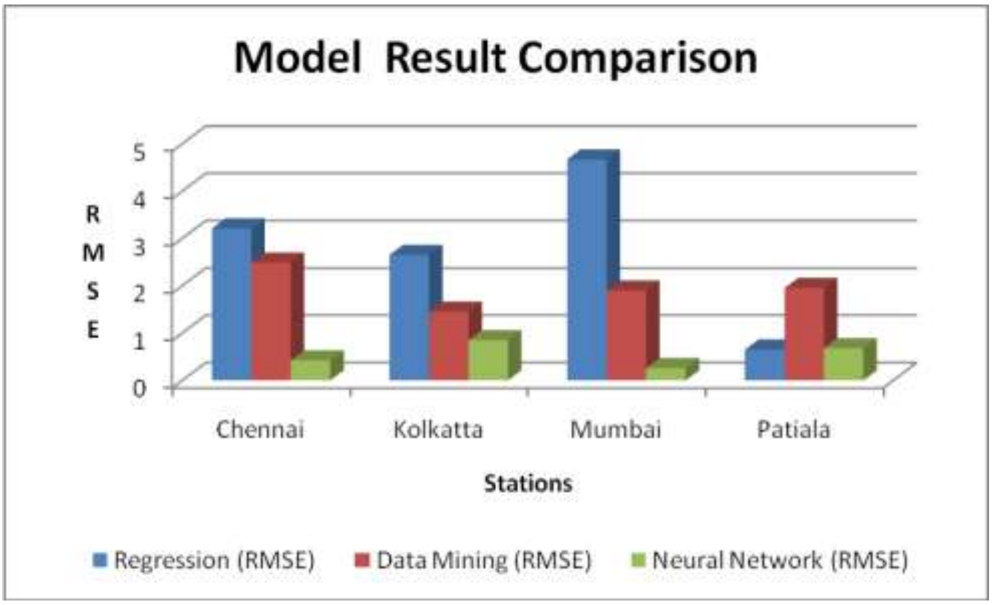
Table: 1.8: Kolkata (India) and Jessore (Bangladesh) validation

Date (Random)	INDIA	Actual Rainfall observed(mm)	Rainfall predicted by the model Neural Network (mm)	BANGLADESH	Validated Rainfall Prediction (mm)
31-05-2013	Kolkata	7.112	2.508	Jessore	4.6
24-06-2013	Kolkata	19.05	8.364	Jessore	1.6
29-09-2013	Kolkata	14.98	23.51	Jessore	21.3

27-05-2014	Kolkata	14.98	27.10	Jessore	12.3
05-09-2014	Kolkata	10.50	17.04	Jessore	13.1

Table 1.9: Patiala (India) and Multan (Pakistan) validation

Date (Random)	INDIA	Actual Rainfall observed(mm)	Rainfall predicted by the model Neural Network (mm)	PAKISTAN	Validated Rainfall Prediction (mm)
05-02-2013	Patiala	13.97	20.21	Multan	16.5
24-03-2013	Patiala	9.08	13.21	Multan	13.3
14-08-2013	Patiala	8.40	14.64	Multan	15.7
07-02-2014	Patiala	9.906	17.40	Multan	14.3
18-03-2014	Patiala	8.89	0.19	Multan	4.6



VI. CONCLUSION

The present study has examined three different models namely Regression, Data mining, NeuralNetwork using fifteen years daily data for four different geographical locations. Six parameters were used in this study namely

Temperature, Dew Point, MSLP, Wind Speed, Humidity and Rainfall. Out of three model prediction the Neural Network model predict the rainfall with high accuracy. The same has been validated with four International locations which is in the similar climatic zone. The rainfall predicted by the model and the actual rainfall observed in the International zones has been verified. Random dates have been chosen for all four geographical locations and tested and found that the neural network predicts the rainfall more accurately.

REFERENCES

[1] Ajaya Mohan, R. S., and B. N. Goswami. 2003. "Potential Predictability of the Asian Summer Monsoon on Monthly and Seasonal Time Scales." *Meteorology and Atmospheric Physics* (Springer-Verlag) 84 (1-2): 83-100. doi:10.1007/s00703-002- 0576-4.

[2] Blanford, H. F. 1884. "On the connection of Himalayan snowfall and seasons of drought in India." *RoyalSociety London.* London. 3-22. doi:doi:10.1098/rspl. 1884.0003.

[3] Gowariker, V., V. Thapliyal, R. P. Sarker, G. S. Mandal, and D. R. Sikka. 1989. "Parametric and powerregression models: New approach to long range forecasting of monsoon rainfall in India." *Mausam* 40: 115-122

[4] Gowariker, V., V. Thapliyal, S. M. Kulshrestha G. S. Mandal, N. Sen Roy, and D. R. Sikka. 1991. "A power regression model for long range forecast of southwest monsoon rainfall over India. ." *Mausam* 42: 125-130.

[5] Murphy, Allan H. 1998. "The Early History of Probability Forecasts: Some Extensions and Clarifications."*Weather and Forecasting* 13 (1): 5-15. doi:10.1175/1520-0434(1998)013<0005:TEHOPF>2.0.CO;2.

[6] Rajeevan, M. N., D. S. Pai, R. Anil Kumar, and B. Lal. 2007. "New statistical models for long-range forecasting of southwest monsoon rainfall over India." *Climate Dynamics* 28 (7): 813-828. doi:10.1007/s00382006-0197-6.

[7] Rajeevan, M. N. 2001. "Prediction of Indian summer monsoon: Status, problems and prospects. ." *Current Science* 81 (11): 1451-1457.

[8] Ranalkar, M. R. , R. P. Mishra, U. K. Shende, and R. D. Vashistha. 2010. "Establishing a network of 550Automatic Weather Stations and 1350 Automatic Rain Gauge Stations across India: Scheme, Scope andStrengths." *WMO Technical Conference on Instruments and Methods of Observations, 30 August 2010 – 1September 2010.* Helsinki, Finland: WMO. 15.

[9] 1975. "Report of the Panel on Long Range Weather Forecasting." Committee Report, Committee onAtmospheric Sciences, National Research Council, National Academy of Sciences, 114. Accessed Jan 21,2016.

[10] Shukla, J., and D. A. Mooley. 1987. "Empirical Prediction of the Summer Monsoon Rainfall over India."*Monthly Weather Review* 115 (3): 695-704. doi: http://dx.doi.org/10.1175/1520-0493(1987)115<0695:EPOTSM>2.0.CO;2.

[11] Tektas, Mehmet. 2010. "Weather Forecasting Using ANFIS and ARIMA MODELS. A Case Study for Istanbul." *Environmental Research, Engineering and Management* 51 (1): 5–10.

[12] Thapliyal, V. 1981. "ARIMA model for long-range prediction of monsoon rainfall in Peninsular India." India Meteorological Department, Monograph Climatology No. 12/81.

[13] Thapliyal, V. 1997. "Preliminary and final long range forecasts for seasonal monsoon rainfall over India."*Journal of Arid Environments* (Elsevier) 36 (3): 385-403. doi:10.1006/jare.1996.0233.

[14] Thapliyal, V., and M. N. Rajeevan. 2003. "Updated operational models for long-range forecasts of Indiansummer monsoon rainfall." *Mausam* 54 (2): 495-504.

[15] Thapliyal, V., and S. M. Kulshrestha. 1992. "Recent models for long-range forecasting of southwest monsoon rainfall over India." *Mausam* 43: 239-248.

[16] Walker, Gilbert T. 1923. "Correlations in Seasonal Variations of Weather, III. A preliminary Study ofWorld Weather." *Memoirs of India Meteorological Department* XXIV: 75-131.

[17] Gowariker, V., V. Thapliyal, S. M. Kulshrestha, G. S. Mandal, N. Sen Roy, and D. R. Sikka, (1991),“A power regression model for longrange forecast of southwest monsoon rainfall over India”, Mausam, Vol. 42, pp. 125–130.

[18] Murphy, A. H., (1998), “The Early History of Probability Forecasts: Some Extensions and Clarifications”, Weather and Forecasting, Vol. 13, pp. 5-15.

[19] Peel, M. C., B. L. Finlayson, and T. A. McMahon, (2007), “Updated world map of the Köppen-Geiger climate classification”, Hydrol. Earth Syst. Sci., Vol. 11, pp. 1633-1644, doi:10.5194/hess-11-1633-2007.

[20] Ranalkar, M. R., R. P. Mishra, U. K. Shende, and R. D. Vashistha, (2010), “Establishing a network of 550 Automatic Weather Stations and 1350 Automatic Rain Gauge Stations across India: Scheme, Scope and Strengths”, WMO Technical Conference on Instruments and Methods of Observations, Helsinki, Finland, 30 August 2010 – 1 September 2010, pp. 15.

[21] Shukla, J., and D. A. Mooley, (1987), “Empirical Prediction of Summer Monsoon Rainfall Over India, Mon. Weather Rev., Vol. 115, pp. 695-703.

[22] Walker, G. T., (1923), “Correlations in Seasonal Variations of Weather”, Mem. India. Meteor. Dept., XXIV, pp. 75-131.