IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 9, Issue.9, September 2021 E-ISSN: 2347-2693

# Quantile Regression Models for Rainfall Data

**S. Damodharan[1], S. Venkatramana Reddy[2], B. Sarojamma[3*]**

[1,3]Department of Statistics, S.V. University, Tirupati-517502, Andhra Pradesh, India
2: Department of Physics, S.V. University, Tirupati-517502, Andhra Pradesh, India

**Corresponding Author: saroja14397@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v9i9.8385 | Available online at: www.ijcseonline.org

Received: 17/Sept/2021, Accepted: 20/Sept/2021, Published: 30/Sept/2021

***Abstract***: Rainfall is important for human beings, animals and plants for their survival. Rainfall depends on many variables such as wind speed, temperature, humidity etc. Mathematical modelling of rainfall data is a stochastic process. Several mathematical models based on the probability concept are available. These models help in knowing the probable weekly, monthly or annually rainfall. Over the past decade or so, a number of models have been developed to generate rainfall and runoff. Monthly rainfall and temperature were analyzed using time series analysis. In this paper we are fitted linear regression model and quartile regression model at various values of tau 0.25, 0.5 and 0.75 for North west India (NWI), West Central India (WCI), North East India(NEI), Central North East India (CNEI) and Peninsular India (PI). Best model among fitted four models is choosing by using root mean square error (RMSE) criteria.

***Keywords:*** Rainfall, Quantile Regression, Linear regression, RMSE

## I. INTRODUCTION

Rainfall plays vital role in agricultural practices and for human beings. Rainfall is mainly depend on the atmospheric variables such as temperature, wind speed, humidity etc. The monsoon season is the principal rain bearing season and in fact a substantial part of the annual rainfall over a large part of the country occurs in this season. In general monsoon starts around mid June and goes on until end of September. Small variations in the timing and the quantity of monsoon rainfall have the potential to impact on agricultural output. Rainfall is the most important climatic element that influences agriculture. Monthly rainfall forecasting plays an important role in the planning and management of agricultural scheme and water resources systems.

Mathematical modelling of rainfall data is a stochastic process. Several mathematical models based on the probability concept are available. These models help in knowing the probable weekly, monthly or annually rainfall. Over the past decade or so, a number of models have been developed to generate rainfall and runoff. Monthly rainfall and temperature are analyzed using time series analysis. ARIMA model is a common linear model that have been used in time series forecasting during past three decades and also ARCH and GARCH models are extensively used. Various methods for statistical forecasting are regression analysis, classical decomposition method, Box and Jenkins and exponential smoothing techniques. Being a number of exponential smoothing methods, Holts-Winter exponential smoothing is one of the popular approaches for forecasting time series, due to its robust and accurate characteristic.

Nail Homani[1] published a paper "Time series analysis models for Rainfall data in Jordan". He used Box and Jenkins model of Autoregressive Integrated Moving Average for forecasting the monthly rainfall data. Mostata Dastorani et al[2] discussed about "Comparative Study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition". In this paper, they used Auto Regressive Integrated Moving Average and Seasonal Auto regressive Integrated Moving Average with different structures of trial and error and it was examined for North khorasan province from 1989 to 2012 using R software. Nasimul Hasan et al [3] used support vector regression model and forecasted 7 days ahead results in their paper "A support vector regression model for forecasting Rainfall". In this paper they also. Luk et al [4] in there article "An Application of Artificial Neural networks for rainfall forecasting", used artificial neural networks like multilayer feed forward neural networks, partial recurrent neural networks and time delay neural networks. Hung et al [5] used ANN for estimation of rainfall using meteorological parameters like relative humidity, air pressure, wet bulb temperature and cloudiness and published a paper on "explains an artificial neural network model for rainfall forecasting in Bangkok, Thailand". Darji et al [6] in their paper "rainfall forecasting using Artificial Neural Networks". Analyse crop productivity and use of water resources and different accuracy measures are used to test performance of ANN. Varsha M and Poornima, have tried to solve the problem of predicting paddy blast disease by using machine learning approaches for the given weather datasets. In proposed solution we considered both disease data and weather datasets that influence in occurrence of blast disease. To improve efficiency model, we have used filter-based feature selection methods to select subset of features and according to the combined results of feature selection methods Minimum Temperature, Humidity, Rainfall has high influence in occurrence of blast disease. To propose solution, with selected features we used Ensemble classifiers to train classifier and it is observed that all 3 classifiers. Kartik Jawanjal et.al [7], A basic implementation of two Panchang types (In python language) has been done. One being Traditional Panchang and the other- Tamil Panchangam. Both the methods mostly use similar rules and logic, but Traditional Panchang allows prediction for a shorter time period accurately and Tamil Panchangam provides accurate predictions for the yearly time period. Varsha and Poornima concluded that the bagging classification model has performed well in predicting paddy blast disease[8].

## II. METHODOLOGY

There are several regression models in literature. Some of the popular regression models are simple linear regression, multiple regression model, polynomial regression model, and ridge regression model etc,. In this paper we are using linear regression and quantile regression for rainfall data in India. Linear regression: by taking rainfall data as dependent variable and time as independent variable. The linear model fitted is as follows.

$R_t = A + B\,T$

Where rainfall at time ' t ' is $R_t$

Time is denoted with T

A and B are coefficients

Coefficients are estimated using mean values and using least square method or any other method.

**Quantile Regression:** Quantile Regression is extension of linear regression. Means are used for prediction of dependent variable for given independent variables. In general mean has its own limitations and it influences prediction also. For estimation of quantile regression, median plays main role in estimation of relation between dependent and independent variables. $r^{th}$Quantile regression model is as follows

$$Z_i = W_i^I \beta + \varepsilon_i$$

Objective function we consider to minimize is

$r\Sigma|\varepsilon_i| + (1 - r)\,\Sigma|\varepsilon_i|$

where r is $r^{th}$quantile

**Root Mean Square Error:** Arithmetic Mean of the square of a set of numbers to errors.

## III. EMPIRICAL INVESTIGATIONS

By taking rainfall as dependent variable and time as independent variable, we are fitted linear regression and Quantile regression for 30 sub divisions of India as North West India contains 6 sub divisions, North east india contains 4 sub divisions, West central India contains 9 sub divisions, central north east India contains 5 sub divisions and peninsular India contains 6 sub divisions.

Fitted linear regression for North West India is as follows

$R_t = 444.53767 + 0.\,01349\,T$

Quantile Regression for North West India at Tau 0.25, 0.5 and 0.75 Coefficients of slope and intercepts are as follows.

Table – 1

| Model | Intercept | Slope |
|---|---|---|
| Linear | 444.5378 | 0.01349 |
| Quantile Regression 0.25 | 0.0021 | 15.9493 |
| Quantile Regression 0.5 | - 0.0017 | 86.1909 |
| Quantile Regression 0.75 | 0.0777 | 517.1386 |

The linear regression and Quantile Regression slope and intercept for West Central India by taking month wise time as independent variable and rainfall as dependent variable, as tabulated in Table-2.

Table – 2 North West India

| Model | Intercept | Slope |
|---|---|---|
| Linear | 901.2592 | - 0.0065 |
| Quantile Regression 0.25 | 15.9426 | 0.0211 |
| Quantile Regression 0.5 | 86.1909 | - 0.0017 |
| Quantile Regression 0.75 | 517.1386 | 0.7773s |

Fitted linear regression and Quantile Regression for rainfall data for 4 sub divisions of North East India, slope and intercept are listed in below.

Table – 3 North East India

| Model | Slope | Intercept |
|---|---|---|
| Linear | - 0.0488 | 1752.1681 |
| Quantile Regression 0.25 | - 0.0174 | 268.7766 |
| Quantile Regression 0.5 | - 5.4773 | 1.3153 |
| Quantile Regression 0.75 | - 0.1790 | 3232.5802 |

Central North India contains 5 sub divisions by taking rainfall as dependent variable and month wise time as independent variable, we performed linear regression model and Quantile regression model slope and coefficients are listed in table – 4

Table – 4 Central North India

| Model | Slope | Intercept |
|---|---|---|
| Linear | - 0.03315 | 1021.3171 |
| Quantile Regression 0.25 | 0. 0039 | 89.4752 |
| Quantile Regression 0.5 | 0.0001 | 3.2378 |
| Quantile Regression 0.75 | 3.5151 | 1.8244 |

6 sub divisions in peninsular India, the linear model and Quantile Regression model fitted to data of rain fall as dependent and monthly time as independent, intercepts and slope are as follows

 

Table – 5 Peninsular India

| Model | Intercept | Slope |
|---|---|---|
| Linear | 942.8509 | 0.0298 |
| Quantile Regression 0.25 | 188.3502 | 0.0220 |
| Quantile Regression 0.5 | 769.7350 | 0.0919 |
| Quantile Regression 0.75 | 1.6199 | 0.3941 |

**Root Mean Square Error:**

Root mean Square Error values for North West India (NWI), West Central India (WCI), North East India(NEI), Central North East India (CNEI) and Peninsular India (PI) by fitting linear and Quantile Regression at 0.25, 0.5 and 0.75 is as follows.

Table – 6

| Model | NEI | CNEI | NWI | WCI | PI |
|---|---|---|---|---|---|
| Linear | 1709.525 | 992.4019 | 456.4129 | 895.5626 | 969.171 |
| 0.25Quantile | 253.6749 | 92.9681 | 17.8344 | 65.018 | 207.997 |
| 0.5 Quantile | 1310.524 | 326.6627 | 84.6738 | 229.7516 | 851.6106 |
| 0.75 Quantile | 3077.008 | 1855.332 | 586.5888 | 1721.072 | 1626.906 |

## IV. SUMMARY AND CONCLUSIONS:

For month wise rainfall data of India, from 30 subdivisions of India is divided into North East India (NEI), West Central India (WCI), North West India (NWI), Central North East India (CNEI)and Peninsular India (PI). We are fitted linear regression model and Quantile regression at Tau 0.25, 0.5 and 0.75. The best model for above subdivisions of India is choose, by using Root mean Square Error criteria.

Table – 7

| area | model | RMSE |
|---|---|---|
| NEI | Quantile Regression Model at tau=0.25 | 253.6749 |
| CNEI | Quantile Regression Model at tau=0.25 | 92.9681 |
| NWI | Quantile Regression Model at tau=0.25 | 17.8344 |
| WCI | Quantile Regression Model at tau=0.25 | 65.018 |
| PI | Quantile Regression Model at tau=0.25 | 207.997 |

## REFERENCES

[1] P.N.M. Momani, "Time Series Analysis Model for Rainfall Data in Jordan: Case Study for Using Time Series Analysis" *American Journal of Environmental Sciences*, Vol **5**, No**5**, pp.**599-604**, **2009**.

[2] Mostata Dastorani, Mohammad Mirzawad, Mohammad Taghi Dastorani, and Syed Javad Sadatinejad, "Comparative study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition**"**, *Natural Hazards*, **Vol. 81, No.3, pp. 1811–1827, 2016.**

[3] N. Hasan, N. C. Nath and R. I. Rasel, "A support vector regression model for forecasting rainfall," *2015 2nd International Conference on Electrical Information and Communication Technologies (EICT)*, pp. 554-559, 2015, doi: 10.1109/EICT.2015.7392014.

[4] Kin C. Luk, J.E. Ball and A. Sharma, "An Application of Artificial Neural Networks for Rainfall Forecasting", *Mathematical and Computer Modelling*, **Vol.33, No.6-7, pp. 683-693, 2001.**

[5] N.Q.Hung, M.S. Babel, S. Weejakul and N.K. Tripathi, "An Artificial Neural Network model for rainfall forecasting in Bangkok, Thailand", *Hydrology and earth sciences*, Vol**.13,** No.**8**, pp**1413-1425**, **2009**.

[6] M. P. Darji, V. K. Dabhi and H. B. Prajapati, "Rainfall forecasting using neural network: A survey," *2015 International Conference on Advances in Computer Engineering and Applications*, pp. 706-713, 2015. doi: 10.1109/ICACEA.2015.7164782.

[7] Kartik Jawanjal, Mayank Modi, Aditya Panchwagh, Nagesh Jadhav, "Weather Prediction Using Indian Almanac Rules," *International Journal of Computer Sciences and Engineering*, **Vol.9, No.6, pp.19-24, 2021.**

[8] M.Varsha, B. Poornima, "A Comparative Study of Ensemble Classifiers for Paddy Blast Disease Prediction Model", *International Journal of Computer Sciences and Engineering*, **Vol.8, No.9, pp.91-100, 2020.**

## AUTHOR'S PROFILE

Mr. **S. Damodharan** pursed M.Sc Statistics in Sri Venkateswara University, Tirupati, A.P, India. Later he joined as Research Scholar under the guidance of Dr. B. Sarojamma, Associate Professor, Department of Statistics in the same University. He has published 7 papers in various journals. He participated in 13 conferences and 5 workshops in various Universities. He has 6 years teaching experience.

Dr. S. Venkatramana Reddy is working as an Associated Professor in the Department of Physics, S.V.University, Tirupati. He has 21 years' teaching and 27 years research in experience. He has published 110 research papers in various internationally reputed Journals and presented more than 110 papers in various National and International Conferences. Received the Best Paper Award for presentation of the Research Paper entitled "Reducing the effect of ground clutter from wind profiler radar signal using wavelet transforms", in the National Conference. 8 Ph.D. and 4 M.Phil. degrees are awarded under his Supervision. He was Co-ordinator for 5yr Integrated M.Sc. Course in Physics and presently Coordinator for M.Sc. Electronics, S.V.U. University, Tirupati.

Dr. B. Sarojamma is working as an Associate Professor in Dept. of Statistics, Sri Venkateswara University, Tirupati, A.P, India. She has 14 years of Teaching and 19 years Research Experience. She Published 88 Research papers and 2 Books. She Presented 130 Research Papers in Various National and International Conferences, Seminars and Symposiums. She attended 57 Wokshops. She organized 4 National Seminars and workshops. 5 Ph.Ds and 2 M.Phil Degrees are awarded under her guidance.

 