Jñānābha, Vol. 50(1) (2020), 243-252
(Dedicated to Honor Professor H.M. Srivastava on His 80th Birth Anniversary Celebrations)

TIME SERIES ANALYSIS OF RAINFALL USING HETEROSKEDASTICITY MODELS

By
Rashmi Bhardwaj and Varsha Duhoon
University School of Basic and Applied Sciences, Non-Linear Dynamics Research Lab, Guru Gobind Singh Indraprastha University, Dwarka, Delhi-110078, India.
Email:rashmib22@gmail.com, varshaduhoon5@gmail.com

(*Received : February* 04, 2020 *; Revised: June* 21, 2020)
DOI: https://doi.org/10.58250/jnanabha.2020.50123

Abstract

Weather forecasting is predicting present state of weather by the help of analyzing collected data such as temperature, humidity, wind etc. to analyze atmospheric processes and determine how weather condition is going to change in future. Weather forecasting is important not only for prediction but also to prepare for future coming events if any such as cyclone, heavy rainfall, hails which can cause harm to the agricultural production of the country and hence affect the livelihood of the farmers. Continuous change in variance of time series over time; such a process is termed as Volatility. Heteroskecdascity refers to increasing variance in a way, such as increase in trend, this property of series is termed as Heteroskecdascity. Objective of the paper is to analyze, model and predict Rainfall time series of Delhi region from January 01, 2017 to December 31, 2018 using Heteroskecdascity model such as *ARCH, GARCH, TARCH/ GJR-GARCH, EGARCH* models to select most suited model on basis of probability value of the model hence calculated. Further, to analyze and choose model has full filled required conditions the model is checked for Serial Correlation, *ARCH* Effect, Normal distribution of Residuals, *ARCH-LM* test is applied, *AIC*, *SIC* values are calculated. *GJR-GARCH* model is most suited model among all models tested for modeling and analyzing rainfall. Model selection is done based on *AIC* value and *SIC* value calculated.

2010 Mathematics Subject Classifications: 33C60, 33C45.
Keywords and phrases: Volatility, Heteroscedascity, *ARCH*, *GARCH*, *TARCH*, *EGARCH*, Time Series, Rainfall.

1 Introduction

Climate change has adverse effect on rainfall as, monsoons are shifted, rainfall as in average has decreased per year and agricultural sector of the economy is adversely affect in many parts of country due to change in pattern of rainfall causing drought and forcing the farmer to find an alternate for their living. In general, climate is highly non-linear phenomena in nature. Climate change is a serious environmental threat to human kind nowadays. Change in weather pattern has and is affecting livelihood of people, it can also be seen that there are longer summers and shifted monsoon every year. Over a long period of time it has been observed that every year temperature is raising by some degrees and overall rainfall expectancy has decreased. The objective in the paper is to test whether volatility which is applied in the field of economics to study the nature of time series can also be applied to weather data or not and to further see which model will give better forecasting of rainfall.

Ardia et al. [1] studied single regime & markov switching *GARCH* models by comparing forecasting efficiency in terms of risk management. Bhardwaj et al. ([2],[3],[4]) studied various methods for weather forecasting by using data mining techniques which included *MLP*, Gaussian Process, *SMO* Regression, Linear Regression among which best suited tool was selected using the statistical calculations *CC, MAE, RMSE, RRSE, RAE*; also studied *ANFIS-SUGENO* model using subtractive clustering technique for the formation of membership function to forecast minimum and maximum temperature; further studied behavior of temperature from 1981-2015 by calculating hurst exponent, fractal dimension, predictability index using R/S method ([2],[3],[4]). Bentes [5] studied GARCH, *IGARCH, FIGARCH* models for predicting volatility of gold returns. Bollerseley [6] studied generalized *ARCH(GARCH)* model for parsimonious representation of *ARCH*. Dickey [7] studied likelihood ratio test and also studied least square estimator of alpha. Engles ([8],[9]) studied *ARCH* and *GARCH* models for analysing time seriesdata for financial application and also studied symmetry effect in exchange rate data by applying *ARCH* and *GARCH* and *EGARCH* model for studying volatility. Gabriel [11] studied forecasting efficiency of GARCH model. Herwartz et al. [12] studied *GARCH* model to predict stock return. Huang et al. [13] studied *HAR-GARCH* model for modelling long memory volatility by the help of measures of volatility. Jordan et al. [14] studied *ARCH/GARCH* model to study volatility in price of field crops. Kobayashi et al. [15] studied *EGARCH* model and compared it to stochastic volatility models. Liu et al. [16] studied *GARCH* model for predicting stocks. D.N. Fente et al. [10] studied different weather parameter and used *LSTM* Technique for forecasting the weather parameters. N. Singh et al. [18] studied weather parameters and forecasted the parameters using machine learning algorithm such as random forest. K. Sushmitha et al. [19] studied weather parameters such as temperature, wind, humidity, rainfall. The parameters were studied and forecasted using *ARIMA*. B. Munmun et al. [17] studied and forecasted weather parameters using data mining technique such as Naïve Bayes and Chi Square algorithm.

GJR-GARCH, ARCH, GARCH, EGARCH models are studied step wise analyzing that whether residuals show that *ARCH* family models can be applied or not and so further Residuals are diagnosed by applying heteroscedascity test and best model is initially selected using the *AIC*, *SIC* values and further the diagnostic test to re check if selected model is actually full filling the conditions required.

2 Methodology

2.1 Time Series Analysis

Time series analysis is done in order to analysis sequence of observation of the time pattern. Analysis is done to study various pattern of the data such as data has a seasonal pattern or random pattern. This study of pattern of rainfall helps us know whether data has repetitive pattern or not. The time series process is analysed based on the past values to study the pattern if any and if the pattern exist then the time series is modelled and forecasting is done.

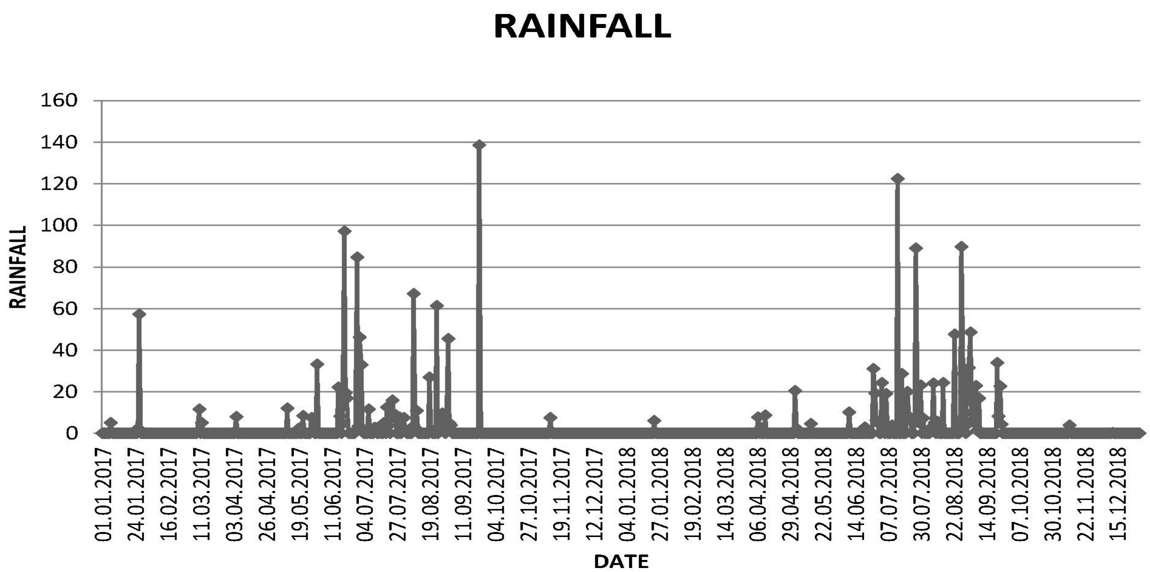

***Figure* 2.1:** Plot of rainfall values from January 01, 2017 till December 31, 2018

2.2 *ARCH*

Auto Regressive Conditional Heteroscedasticity *(ARCH)* model includes single or many data values in series further variance of existing error term is a function of actual size of previous time periods error terms. *ARCH* model models time series. *ARCH(q)* model is used to model time series by *ARCH* process, ϵ_t: error term,ϵ_t are split into stochastic piece, λ_t: standard deviation.

Step 1: consider a time series t, for all values from 1 to n.

Step 2: estimate best fit *AR(q)* by calculating: $g_t = x_0 + x_1 g_{t-1} + ... + x_q g_{t-q} + \epsilon_t$

Step 3: calculate the error term ϵ_t , with the help of formula:
$\epsilon_t = Z_t \lambda_t$, Z_t: white noise process.

Step 4: model time series by calculating $\lambda_t^2 = x_0 + x_1 \epsilon_{t-1}^2 + ... + x_q \epsilon_q^2]$, Such that is positive or equal to 0.

Further, *ARCH (q)* model is analyzedby ordinary least squares.

2.3 *GARCH*

ARMA model is used as error variance, Generalized Autoregressive Conditional Heteroscedasticity model. *GARCH(p, q)* model, p:order of *GARCH* terms λ^2& q:order of *ARCH* ϵ^2.

Step 1: consider time series t, for all values from 1 to n.

Step 2: estimate best fit $AR(q)$ by calculating:$g_t = x_0 + x_1 g_{t-1} + ... + x_z g t - q + \epsilon_t$

Step 3: calculate $Y_t = X_t' b + \epsilon_t$, such that $\epsilon_t / \phi_t \; N(0, \lambda_t^2)$, where ϵ_t is the error term.

Step 4: now calculate standard deviation for time series t for order q with ϵ_t as error term

$$\lambda_t^2 = W + \alpha_1 \epsilon_{t-1}^2 + ... + \beta_1 \lambda_{t-1}^2 + ... = W + \sum_{i=1}^{q} \lambda_i \epsilon_{t-1}^2 + \sum_{i=1}^{q} \beta_i \epsilon_{t-1}^2$$

GARCH model is applied for reducing error in forecasting time series and further improve accuracy of forecast.

2.4 *EGARCH*

Exponential Generalized Autoregressive Conditional Heteroskedastic model is extension of *GARCH* model which studies time series values by calculating variance of present error term, such that "p'':degree of *GARCH* polynomial and "q'' is degree of *ARCH* & Leverage polynomial.

Formally, $EGARCH(p,q)$:
Step 1: consider time series t, for all values from 1 to n.
Step 2: estimate best fit $AR(q)$ by calculating: $g_t = x_0 + x_1 g_{t-1} + ... + x_q g_{t-q} + \epsilon_t$
Step 3: now, calculate $log\lambda_t^2 = W + \sum_{i=1}^{q} \beta_i(z_{t-i}^i) + \sum_{i=1}^{p} \alpha_i log\lambda_{t-i}^2$, $log\lambda_t^2$ can be negative, hence no restrictions on parameters.
Step 4: further calculate $g(z_t) = \phi z_t + \tau(|z_t| - E(|z_t|))$, λ_t^2, $g(z_t, z_t)$, $z_t$sign and magnitude have separate effect on volatility.

Above is conditional variance $w, \beta, \alpha, \phi, \tau$: coefficients,$z_t$:standard normal variable,also known as error distribution. Positive or negative sign is not a limitation in *GARCH* model. Test of *ARCH* and *GARCH* errors is important for studying time series.

2.5 *T-GARCH*

Threshold-*GARCH* model which has another name that is *GJR-GARCH* model. Model includes calculation of conditional Standard Deviation instead of calculating conditional variance. $T - GARCH(p,q)$ model is calculated as follows:
Step 1: consider time series t, for all values from 1 to n.
Step 2: estimate best fit $AR(q)$ by calculating:$g_t = x_0 + x_1 g_{t-1} + ... + x_q g_{t-q} + \epsilon_q$
Step 3: calculate error term using $\epsilon_{t-1}^+ = \epsilon_{t-1}$ if $\epsilon_{t-1}^+ = 00, \epsilon_{t-1} > 0$ and vice versa.
Step 4: calculate standard deviation using: $\lambda_t = K + \delta\lambda_{t-1} + ... + \alpha_i + \epsilon_{t-1}^+$

3 Resluts and Discussion

Daily data of Temperature, rainfall for Delhi with coordinates Longitude 770 09' 27" Latitude280 38 '23" N Altitude:228.61m from January 01, 2017 till December 31, 2018 is taken. The time series of daily values of rainfall have been taken and the parameters are set accordingly with respect to the resulting probability values and values of *AIC, SIC* the best fit model is selected. The best suited model is selected among *ARCH, GARCH, TARCH/ GJR-GARCH, EGARCH*.

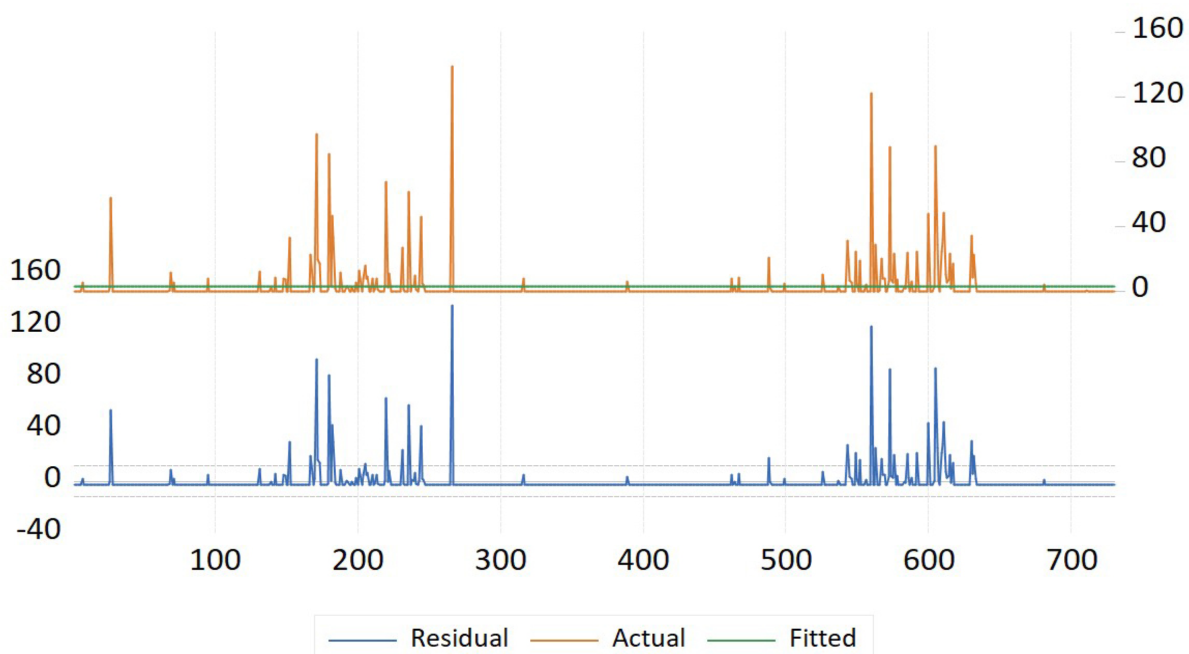

***Figure* 3.1:** Plot of rainfall values from January 01, 2017 till December 31, 2018

In ***Figure* 3.1**, residuals have been plotted of rainfall time series to understand pattern of time series, it shows that periods of low values is followed by high values and similarly further low values for longer duration is followed by the period of high values thus it implies that we can apply the *ARCH* family models.

Further, the residuals are diagnosed whether *ARCH* family models can be applied or not and so further Residuals are diagnosed by applying Heteroskecdascity test: *ARCH*, in which the following hypothesis is considered:

Null Hypothesis: Absence of *ARCH* effect. (If probability value < 5%)

Alternative Hypothesis: Presence of *ARCH* effect. (If probability value > 5%)

***Table* 3.1:** Residual Diagnostic using Heteroskecdascity test for Rainfall

Heteroskecdascity test	Observed R^2	Chi-probability
ARCH	**9.681242**	**0.0019 (< 5%)**

Now, it is clear from ***Table* 3.1**, that probability value comes 0.0019, that is lesser than 5%; hence rejecting the Null Hypothesis, absence ofARCHeffect. Now, there exists *ARCH* effect in time series of Rainfall therefore, *ARCH* family models such as *GARCH, TGARCH, EGARCH* models can be applied.

***Table* 3.2:** AIC, SIC values as per the model for Rainfall time series

Heteroscedascity test	Observed R^2	Chi-probability
ARCH	**0.04**	**0.82 (> 5%)**

The above ***Table* 3.2**, shows that *GJR-GARCH* or *TGARCH* model is best suitedon basis of *AIC*, SIC values calculated for each model. Now, *GJR-GARCH* model will be tested whether it full fills all the conditions required.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.000	-0.000	5.E-05	0.994
		2	-0.001	-0.001	0.0016	0.999
		3	0.004	0.004	0.0113	1.000
		4	-0.002	-0.002	0.0138	1.000
		5	0.002	0.002	0.0161	1.000
		6	-0.003	-0.003	0.0226	1.000
		7	-0.008	-0.008	0.0688	1.000
		8	-0.006	-0.006	0.0990	1.000
		9	0.024	0.024	0.5356	1.000
		10	-0.007	-0.007	0.5696	1.000
		11	0.003	0.003	0.5773	1.000
		12	-0.007	-0.008	0.6179	1.000
		13	0.036	0.036	1.5851	1.000
		14	-0.013	-0.013	1.7082	1.000
		15	-0.010	-0.010	1.7865	1.000
		16	-0.002	-0.002	1.7893	1.000
		17	0.011	0.011	1.8789	1.000
		18	-0.013	-0.014	2.0084	1.000
		19	-0.004	-0.004	2.0209	1.000
		20	-0.010	-0.010	2.0901	1.000
		21	-0.009	-0.009	2.1560	1.000
		22	0.007	0.005	2.1939	1.000
		23	-0.013	-0.012	2.3167	1.000
		24	-0.009	-0.009	2.3824	1.000
		25	-0.007	-0.006	2.4167	1.000
		26	-0.005	-0.007	2.4346	1.000
		27	-0.005	-0.004	2.4540	1.000
		28	-0.008	-0.008	2.5048	1.000
		29	-0.009	-0.009	2.5703	1.000
		30	0.058	0.057	5.1531	1.000
		31	-0.012	-0.011	5.2603	1.000
		32	0.013	0.014	5.3867	1.000
		33	-0.009	-0.010	5.4546	1.000
		34	-0.005	-0.004	5.4766	1.000
		35	-0.002	-0.003	5.4784	1.000
		36	-0.008	-0.007	5.5310	1.000

***Figure* 3.2:** Correlogram plot of GJR-GARCH Model

Now, ***Figure* 3.2**, shows probability values greater than 5% hence, we accept Null Hypothesis which is: Absence of Serial Correlation.

Hence, there is no Serial Correlation.

***Table* 3.3:** Diagnostic using ARCH-LM test for Rainfall for ARCH effect

Diagnostic using *ARCH-LM* test for Rainfall for *ARCH* effect

Model	***Equation***	***AIC***	***SIC***
***ARCH*(5)**	**X(2)+X(3)*RESID(-1)^2+ X(4) * RESID(-1)^@+X(5)*RESID(-3)^2+X(6)*RESID(-4)^2+X(7)*RESID(-5)^2**	**7.670149**	**7.714192**
***GARCH*(1,1)**	**X(2)+X(3)*RESID(-1)^2+X(4)*GARCH(-1)**	**7.233153**	**7.258321**
***GJR-GARCH/ GARCH*(1,1)**	**X(2)+X(3)*RESID(-1)^2+X(4)*RESID(-1)^2*(RESID(-1)<0)+X(5)*GARCH(-1)**	**6.489396**	**6.520855**
***EGARCH* (1,1)**	**X(2)+X(3)*ABS(RESID(-1)SQRT(GARCH(-1)))+X(4)*RESID(-1)SQRT(GARCH(-1)+X(5)*LOG(GARCH(-1))**	**7.493835**	**7.525294**

Null Hypothesis: *ARCH* Effect-NOT PRESENT. (prob >5%)

Alternative Hypothesis: *ARCH* effect-PRESENT. (prob < 5%)

Hence, Null Hypothesis is accepted since probability value is greater than 5%; no *ARCH* effect.

Now, checking if residuals are distributed normally or not using Histogram plot and checking the probability value of the thus calculated.

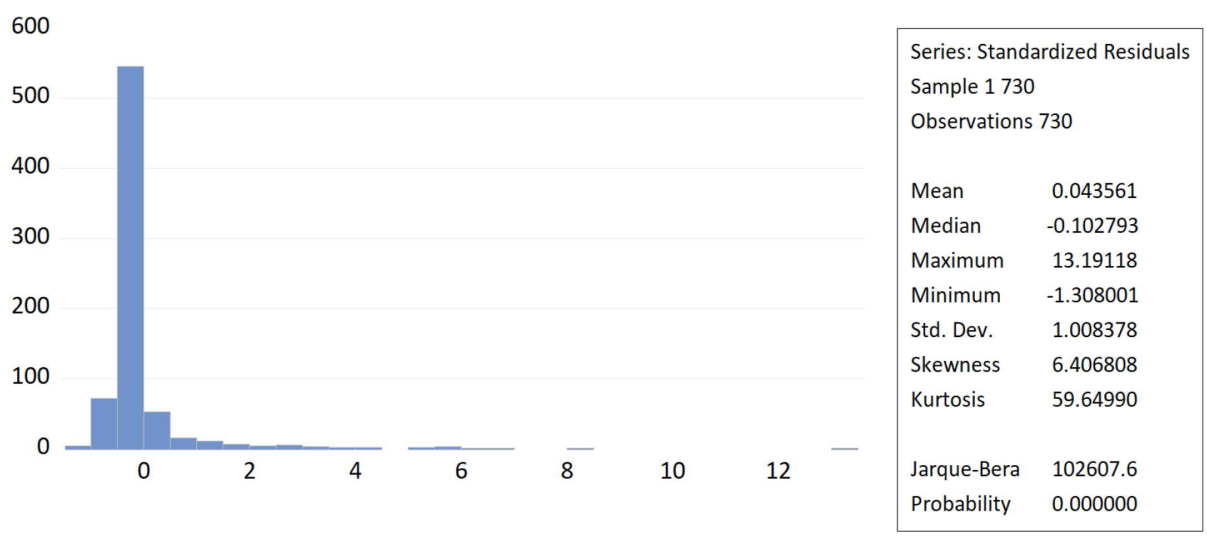

***Figure* 3.3:** Histogram plot to check the normal distribution and probability value

Now, since value of probability is less than5% thus, we reject Null Hypothesis (X_0)&accept

Alternative Hypothesis (X_1).

X_0: Normally Distributed.

X_1: Not Normally Distributed. Hence, distribution of residuals is not normal. As there is absence of *ARCH* effect therefore *GJR-GARCH* is to be considered.

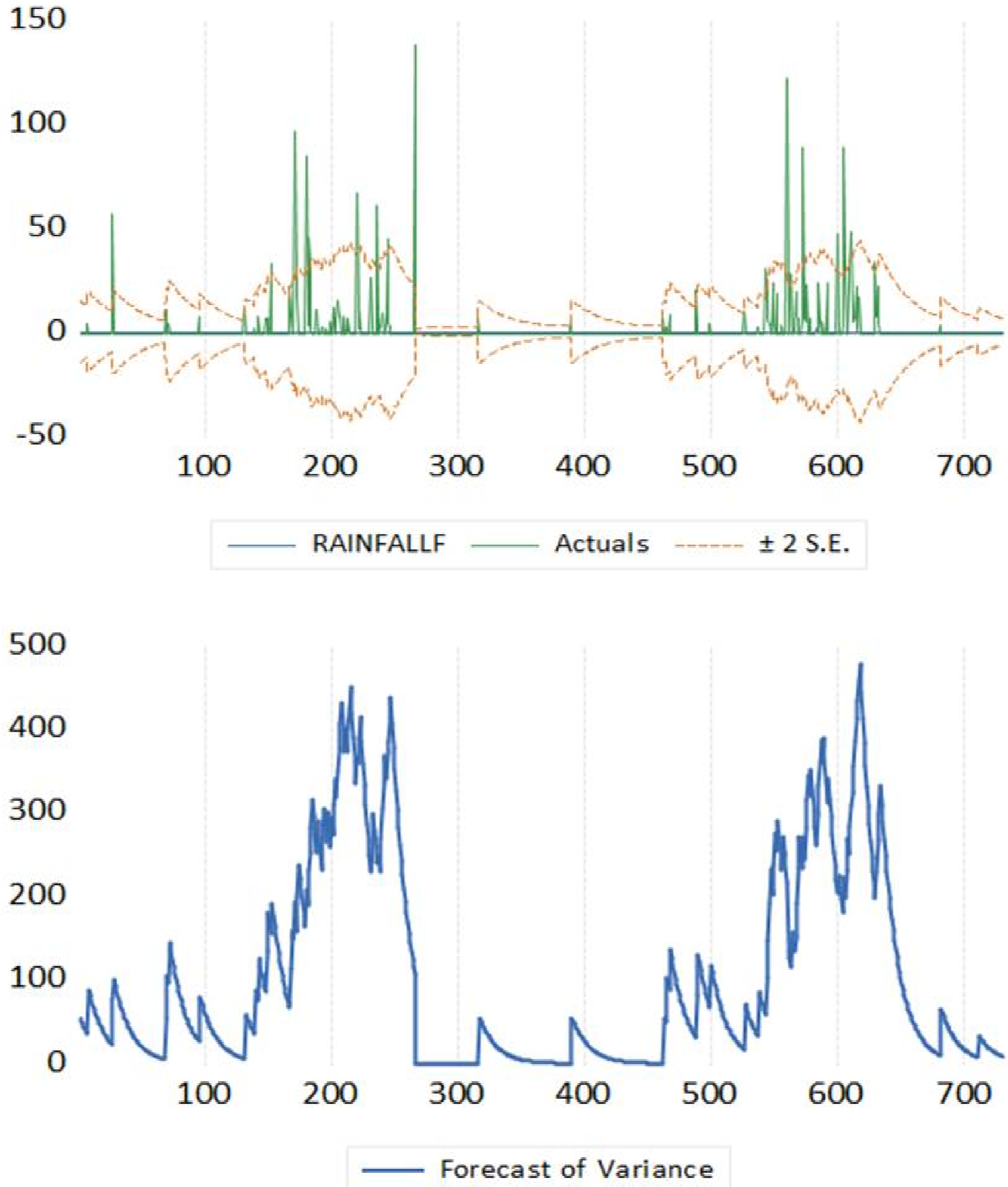

***Figure* 3.4:** Forecast of variance and Rainfall

***Figure* 3.4**, shows the forecast of variance of rainfall and standard error calculated. *GJR-GARCH* is considered on the basis of tests and *AIC, SIC* values and hence the above plot is obtained using the model.

4 Conclusion

The above study has been done to analyze, model and predict Rainfall time series of Delhi region using Heteroskecdascity model such as *ARCH, GARCH, TARCH/ GJR-GARCH, EGARCH* models by selecting most suited model on basis of calculations. Further, the objective was to analyze whether model has full filled required conditions by checking for Serial Correlation, *ARCH* Effect, Normal distribution of Residuals, *ARCH-LM* test, *AIC, SIC* values calculated. *GJR-GARCH* model is the most suited model among all models tested for modelling rainfall. Model selection is done based on *AIC* value and *SIC* value calculated and all other tests done. It can also be concluded that Heteroskecdascity model which is used in economics to study the volatility of time series can be applied in area of weather forecasting by studying the time series behavior of weather parameters and checking if the time series shows required behavior for application of such models as applied in the study above.

Hence, *GJR-GARCH/ TARCH* is the suited for forecasting daily rainfall. Since, the residuals have no *ARCH* effect also no Serial correlation thus its observed that *GJR-GARCH* model is the suited model.

5 Acknowledgements

The authors are thankful to Guru Gobind Singh Indraprastha University, Delhi (India) for providing research facilities and financial support. They are also thankful to the referee and Editor for their valuable suggestions to bring the paper in its present form.

References

[1] D. Ardia, K. Bluteau, K. Boudt and Leopoldo Catania, Forecasting risk with Markov-switching *GARCH* models: A large-scale performance study, *International Journal of Forecasting*, **34 (4)** (2018), 733-747

[2] R. Bhardwaj and V. Duhoon, Weather forcasting using Soft Computing Technique, International Conference on Computing, Power, and Communication Technologies (GUCON), IEEE Explore Digital Library, (2019), 1111-1115.

[3] R. Bhardwaj and V. Duhoon, Real Time Weather Parameter Forecasting Using Anfis-Sugeno, International Journal of Engineering and Advanced Technology *(IJEAT)*, Blue Eyes Intelligence Engineering & Sciences Publication *(BEIESP)*, **9(1)** (2019), 461-469.

[4] R. Bhardwaj and V. Duhoon, Time Series Analysis of Heat Stroke, Jnbha, Vijnana Parishad of India, **49(1)** (2019), 01-10.

[5] R. S. Bentes, Forecasting volatility in gold returns under the *GARCH, IGARCH* and *FI-GARCH* frameworks: New evidence, Physica A: Statistical Mechanics and its Applications, **438** (2015), 355-364

[6] T. Bollerslev, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, **31(3)** (1986), 307-327.

[7] D. A. Dickey and W. A. Fuller, Likelihood ratio statistics for autoregressive time series with unit root, *Econometrica*, **49** (1981), 1057-1072.

[8] R. Engle, GARCH 101: The use of ARCH/GARCH models in applied econometrics, Journal of Economic Perspectives, **15(4)** (2001), 157-168.

[9] R. Engle, Risk and volatility: Econometric models and financial practice, *American Economic Review*, **94(3)** (2004), 405-416.

[10] D. N. Fente and D. Kumar Singh, Weather Forecasting Using Artificial Neural Network, *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, (2018), 1757-1761.

[11] S. A. Gabriel, Evaluating the Forecasting Performance of *GARCH* Models, *Procedia - Social and Behavioral Sciences*, **62** (2012), 1006-1010

[12] H. Herwartz, Stock return prediction under GARCH - An empirical assessment, *International Journal of Forecasting*, **33 (3)** (2017), 569-580

[13] Z. Huang ,H. Liu and T. Wang , Modeling long memory volatility using realized measures of volatility: A realized *HAR GARCH* model, *Economic Modelling*, **52(B)** (2016), 812-821.

[14] H. Jordaan, B. Grove, A. Jooste, Z.G. Alemu, Measuring the price volatility of certain field crops in South Africa using *ARCH/GARCH* approach, **46(3)** (2007), 306-322.

[15] M. Kobayashi, X. Shi, Testing for *EGARCH* against stochastic volatility models, *Journal of Time Series Analysis*, **26** (2005), 135-150.

[16] C. H. Liu and C. J. Hung, Forecasting S&P-100 stock index volatility: The role of volatility asymmetry and distributional assumption in *GARCH* models, *Expert Systems with Applications*, **37 (7)** (2010), 4928-4934

[17] B. Munmun, D. Tanni and B. Sayantanu, Weather Forecast Prediction: An Integrated Approach for Analyzing and Measuring Weather Data. *International Journal of Computer Applications*, **182** (2018), 20-24. 10.5120/ijca2018918265.

[18] N. Singh, S. Chaturvedi and S. Akhter, Weather Forecasting Using Machine Learning Algorithm, 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, (2019), 171-174, doi: 10.1109/ICSC45622.2019.8938211

[19] K. Sushmitha and T. Shashikumar, A real-time weather forecasting and analysis, (2017), 1567-1570. 10.1109/ICPCSI.2017.8391974.