Rainfall Prediction in India

Aashish Kumar

UT EID : ak38299

Time Series Final Project

Contents

1. Introduction:

India receives most of its rainfall during the monsoon season. Monsoon starts from around May and lasts till August. Prediction of accurate amount of rainfall is critical for proper planning of usage of resources. Farmers in India are heavily dependent on the rainfall for irrigation. Predicting the accurate amount of rainfall in a particular month could help farmers plan their crop growth and provide necessary irrigation. Rainfall also fills the dams, rivers, recharges ground water and is a critical source of water in India. If the water resource is not managed properly, then it could lead to cases of water scarcity and drought in some regions. Water scarcity in big cities like Bangalore is a common occurrence. Rainfall prediction could help government devise proper strategy and plan to use the resources and tackle these situations.

I have the data for monthly rainfall in India(in mm) from Jan 1901 to Dec 2015. My goal is to give accurate prediction of the rainfall in the coming years using the time series models and techniques learned during the course.

My project is centered on four sections. The first section is analysis of data. The second section is stationary time series modeling of the rainfall data. The third section is the non-stationary time series modeling of the rainfall data. The fourth section is ARMAV modeling of the data. Finally, all the models are analyzed and their predictions are compared to give better insight into the capabilities and limitations of the models derived.

2. Data Analysis and Collection

Series Title: India area weighted monthly rainfall(in mm) from 1901-2015

Data Source: https://data.gov.in/catalog/rainfall-india

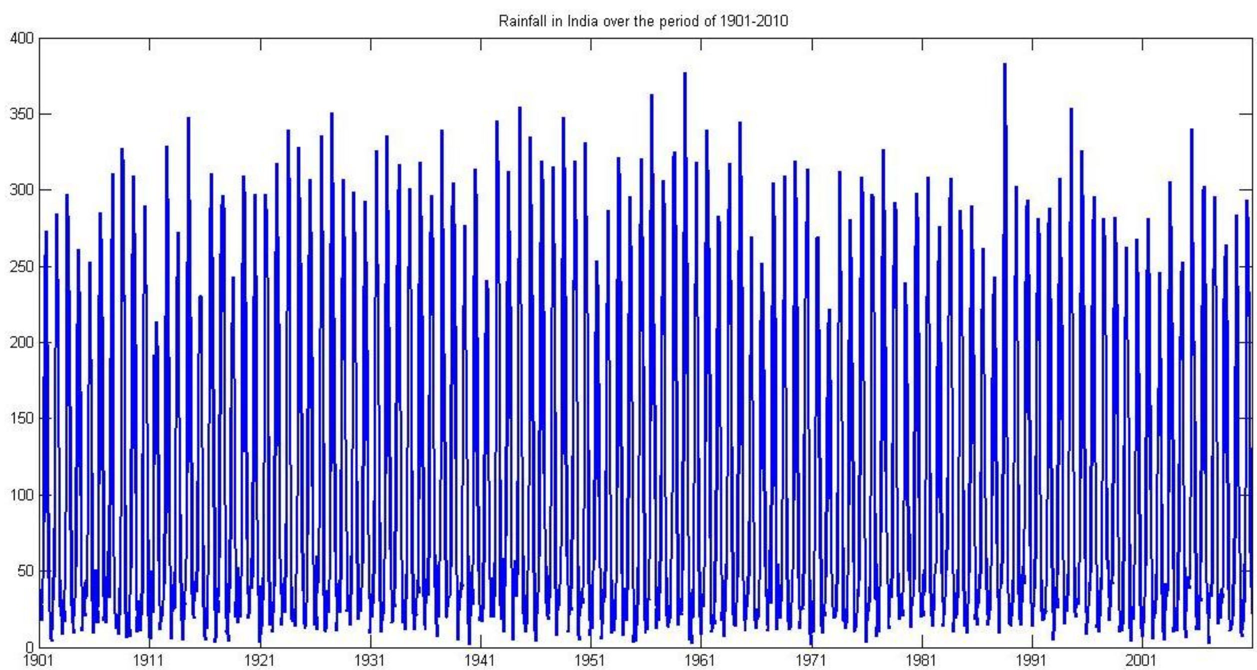
Time Span: Jan 1901-Dec 2015. Total 1380 Data points.

Modeling : Jan 1901-Dec 2010. Total 1320 Data points

Forecasting : Jan 2011-Dec 2015. Total 60 Data points

Units: Rainfall in mm

Rainfall in mm



Time

Figure 1 : Rainfall in India over the period of 1901-2010

2.1 Analysis of Data

Data has average rainfall of India on monthly basis in mm. India receives bulk of its rainfall during the monsoon period, which occurs once every year from May-August, so a seasonality of 12 is expected to be seen in the data, which needs to be confirmed by the models. For modeling, data from Jan 1901 to Dec 2010 is used which has 1320 data points. Data from Jan 2011 to Dec 2015, is kept to compare with the prediction obtained through the model and calculate the accuracy of the model. The prediction is made up to 5 years ahead, so that the Government can effectively plan its short term and long term policies and strategy.

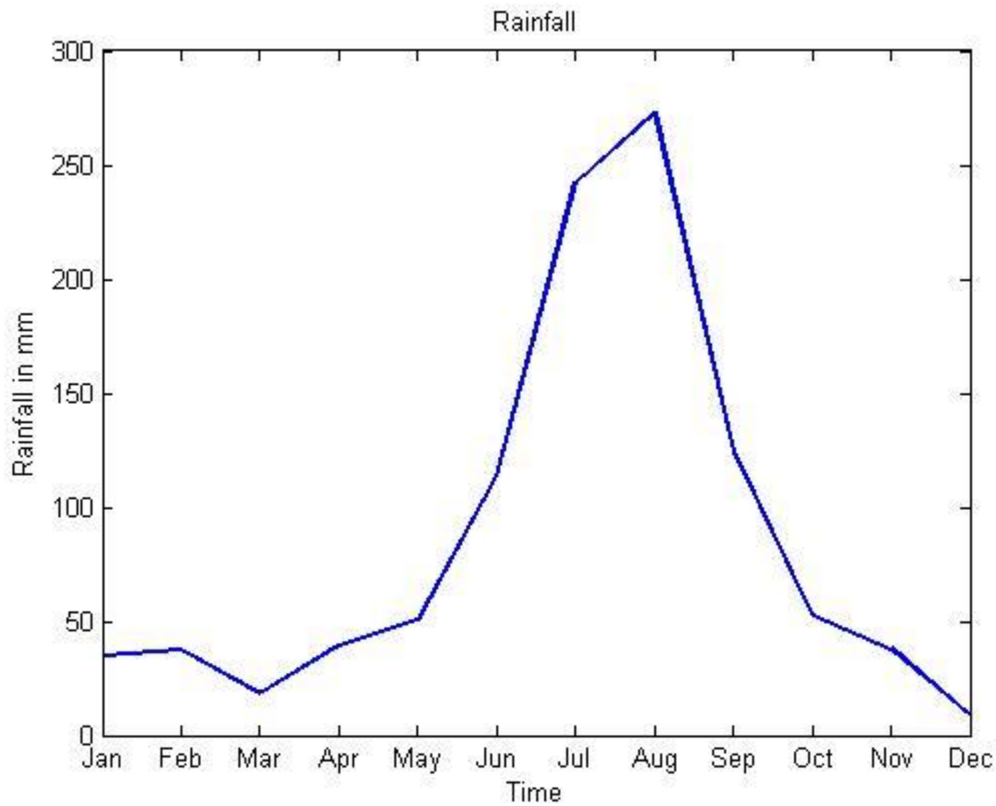


Figure 2 : A typical rainfall pattern in a year in India

3. Stationary Time Series Modeling

3.1 Modeling Process

ARMA(2n,2n-1) modeling strategy is used to obtain the adequate model. The modeling strategy starts from n = 1, fits an ARMA(2,1) model and calculates the RSS. It then increases the value of n by 1 and fits another ARMA model and calculates the RSS. It then calculates whether the change is significant using the F-test and compares it with the and $F^{95}_{s,inf}$ (95% value of F-test). If the change is insignificant then the model stops and the restricted model is the correct model, otherwise the step repeats and value of n is incremented by 1. When an adequate model is found then, the data is fit to odd order model, i.e. ARMA(2n-1,2n-2). Again F-test is used to calculate the adequate model. After that, the modeling strategy tries to drop the leading values of Moving Average part, and compares the adequacy of the model using the F-test.

Using the above strategy, ARMA(10,9) model was found to be the adequate model.

ARMA(10,9) model :

Autoregressive Part: $X_t - 0.89 X_{t-1} + 1.409 X_{t-2} - 0.707 X_{t-3} + 0.6833 X_{t-4} + 0.3656 X_{t-5} - 0.4489 X_{t-6} + 1.254 X_{t-7} - 0.8564 X_{t-8} + 1.071 X_{t-9} - 0.1327 X_{t-10}$

Moving Average Part: $a_t - 0.76 a_{t-1} + 1.334 a_{t-2} - 0.567 a_{t-3} + 0.6703 a_{t-4} + 0.3826 a_{t-5} - 0.3196 a_{t-6} + 1.144 a_{t-7} - 0.6532 a_{t-8} + 0.9473 a_{t-9}$

Residual Sum of Squares for ARMA(10,9) model = 8.1381e+05. The confidence interval for θ_9 (0.9473 ± 0.0490) does not include zero, and thus the Moving Average part is not dropped which is also confirmed by using the above strategy mentioned. Please refer to Table 1 (below) for the parameters, confidence intervals and residual sum of squares for the ARMA(10,9) model, together with other ARMA models used as a comparison to derive the adequate ARMA model.

Table 1 : Estimates of the parameters of ARMA model

Parameter	Order of the Model		
	(8,7)	(10,9)	(12,11)
ϕ_1	1.8516 ± 0.0984	0.89 ± 0.0718	0.5836 ± 0.3691
ϕ_2	-2.2357 ± 0.2198	-1.4085 ± 0.0854	-0.4843 ± 0.4542
ϕ_3	1.0490 ± 0.3406	0.707 ± 0.1480	-0.1545 ± 0.5741
ϕ_4	0.510 ± 0.3486	-0.6833 ± 0.1481	0.3256 ± 0.5409
ϕ_5	-1.9325 ± 0.2885	-0.3656 ± 0.1710	-0.8149 ± 0.3687
ϕ_6	1.8367 ± 0.1983	0.4489 ± 0.1439	0.6668 ± 0.3152
ϕ_7	-1.1293 ± 0.1228	-1.2542 ± 0.1246	-0.7460 ± 0.2085
ϕ_8	0.0909 ± 0.0672	0.8564 ± 0.0953	0.1989 ± 0.4441
ϕ_9		-1.0714 ± 0.0550	-0.0086 ± 0.4770
ϕ_{10}		0.1327 ± 0.0549	-0.6109 ± 0.4406
ϕ_{11}			0.6566 ± 0.3561
ϕ_{12}			0.0471 ± 0.0995
θ_1	1.6735 ± 0.0819	0.7633 ± 0.0490	0.4013 ± 0.3668
θ_2	-1.9431 ± 0.1950	-1.3336 ± 0.0836	-0.4722 ± 0.4057
θ_3	0.7795 ± 0.3232	0.5676 ± 0.1453	-0.1960 ± 0.5066
θ_4	0.5583 ± 0.3611	-0.6703 ± 0.1670	0.2750 ± 0.4747
θ_5	-1.6893 ± 0.3067	-0.3826 ± 0.1936	-0.7427 ± 0.3058
θ_6	1.4708 ± 0.1758	0.3196 ± 0.1681	0.5470 ± 0.2798
θ_7	0.8118 ± 0.0669	-1.1439 ± 0.1460	-0.6372 ± 0.1611
θ_8		0.6532 ± 0.0847	0.1287 ± 0.3619
θ_9		-0.9473 ± 0.0490	0.0147 ± 0.3791
θ_{10}			-0.6143 ± 0.3376
θ_{11}			0.4682 ± 0.2711
RSS	9.2811e+05	8.1381e+05	8.6831e+05

Adequate Model is ARMA(10,9).

Once the adequate model is obtained, correlations between the residuals are checked to make sure they are indeed white. The plot below shows that the residuals are uncorrelated and thus the model is adequate and correct.

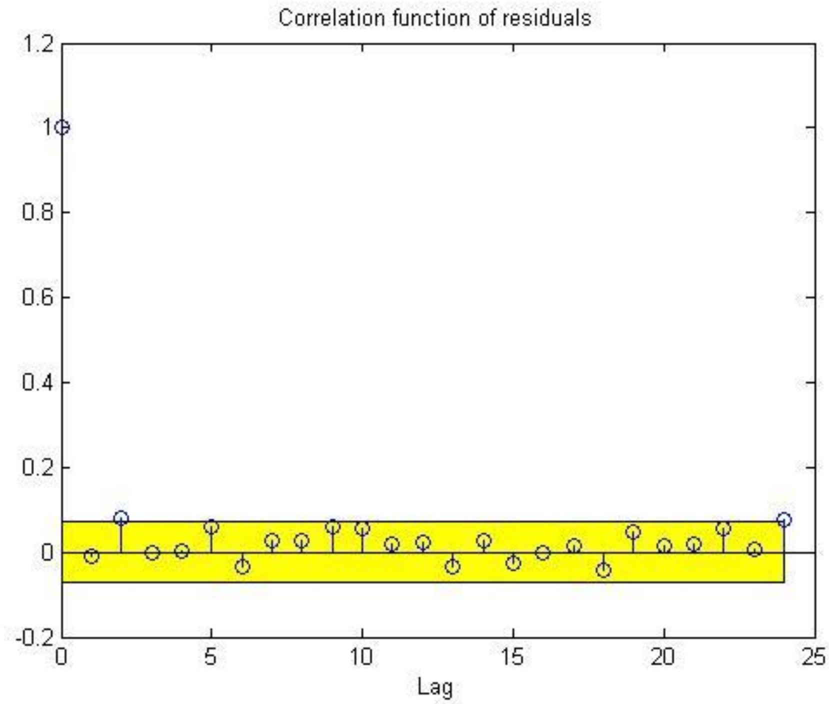


Figure 3 : Correlation of residuals from stationary model ARMA(10,9)

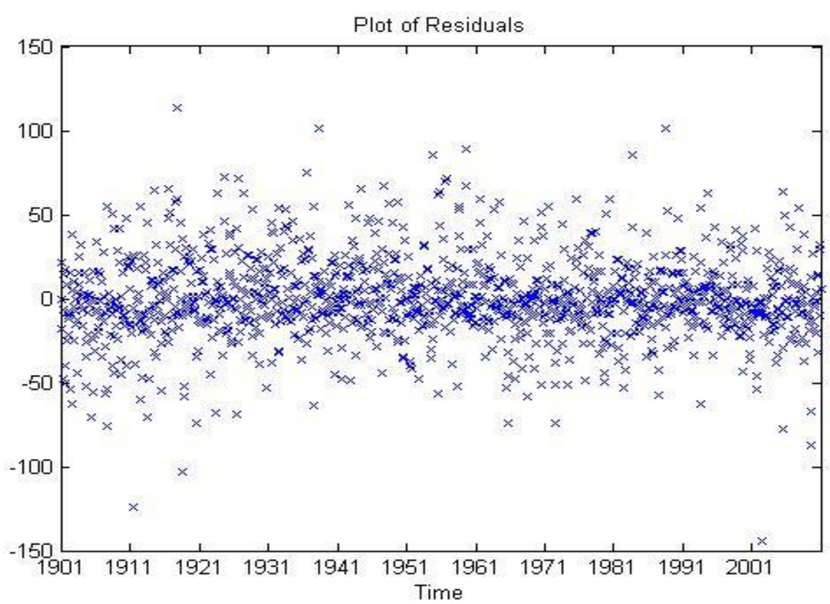


Figure 4 : Plot of residuals versus time

3.2 Model Analysis

For the ARMA(10,9) model the Autoregressive roots obtained are :

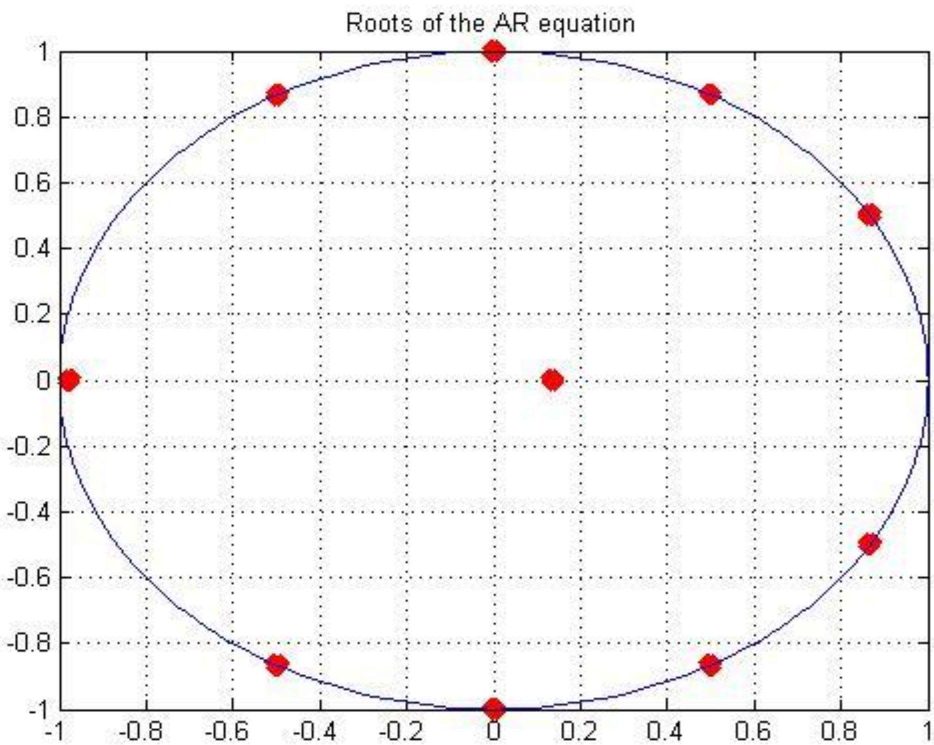


Figure 5 : Plot of the AR roots

Table 2 : AutoRegressive roots and their Corresponding period

Roots	Absolute Value	Angle	Period of Seasonality
-0.9780 + 0.0000i	0.9780	3.1416	2.0000
-0.4998 + 0.8658i	0.9998	2.0944	3.0001
-0.4998 - 0.8658i	0.9998	-2.0944	-3.0001
0.0000 + 1.0000i	1.0000	1.5708	4.0001
0.0000 - 1.0000i	1.0000	-1.5708	-4.0001
0.5000 + 0.8660i	1.0000	1.0472	5.9999
0.5000 - 0.8660i	1.0000	-1.0472	-5.9999
0.8660 + 0.5000i	1.0000	0.5237	11.9985
0.8660 - 0.5000i	1.0000	-0.5237	-11.9985
0.1357 + 0.0000i	0.1357	0	Inf

Seasonality: As can be seen from the table above, the roots 0.8660 ± 0.5i contributes to the seasonality of 12. There are higher harmonics of this seasonality period = 6(=12/2), period = 4(=12/3), period = 3(=12/4), period =2(=12/6) in the model. Parsimonious model for seasonality of 12, 6 and combination of 12 and 6 is tried and compared with the original model, so that we can reduce the number of parameters in the autoregressive equation.

Trend: There is no root at 1. Hence, there is no trend observed in the model.

3.3 Parsimonious Model

Check for seasonality of 12

The operator $1 - 1.7321B + B^2$ is applied to the time series X_t to check for seasonality of 12. Then the resultant time series is fitted to an ARMA(8,9) model.

RSS of restricted model ARMA(8,9), RSS_1 = 8.1430e+05

RSS of unrestricted model ARMA(10,9), RSS_0 = 8.1381e+05

Number of restricted parameters, s = 2

Number of observations = 1320

Number of parameters in unrestricted model, r = 20

$F = ((RSS_1 - RSS_0)/s)/(RSS_0/(N-r)) = 0.3906$

$F^{95}_{2,inf} = 3.0026$

$F < F^{95}_{2,inf}$

So, from the F-test we see that that the change in RSS of ARMA(8,9) model is insignificant as compared to the ARMA(10,9) model. Hence, the ARMA(8,9) parsimonious model is adequate and the seasonality of 12 is confirmed in the model. The parameters of the ARMA(8,9) model is mentioned in table 3.

Check for seasonality of 6

The operator $1 - B + B^2$ is applied to the time series X_t to check for seasonality of 6. Then the resultant time series is fitted to an ARMA(8,9) model.

RSS for ARMA(8,9) = 8.1430e+05

RSS of restricted model ARMA(8,9), RSS_1 = 8.1391e+05

RSS of unrestricted model ARMA(10,9), RSS_0 = 8.1381e+05

Number of restricted parameters, s = 2

Number of observations = 1320

Number of parameters in unrestricted model, r = 20

F = ((RSS$_1$ - RSS$_0$)/s)/(RSS$_0$/(N-r)) = 0.0819

$F^{95}_{2,inf}$ = 3.0026

F < $F^{95}_{2,inf}$

So, from the F-test we see that that the change in RSS of ARMA(8,9) model is insignificant as compared to the ARMA(10,9) model. Hence, the ARMA(8,9) parsimonious model is adequate and the seasonality of 6 is confirmed in the model.

Combined model of seasonality 6 and 12

Now, we combine the seasonality of 12 and 6 and check whether this model is adequate. The operator $1 - 2.73*B + 3.732*B^2 - 2.732*B^3 + B^4$ is applied to the time series X_t. Then the resultant time series is fitted to an ARMA(6,9) model.

RSS of restricted model ARMA(8,9), RSS$_1$ = 8.8546e+05

RSS of unrestricted model ARMA(10,9), RSS$_0$ = 8.1381e+05

Number of restricted parameters, s = 4

Number of observations = 1320

Number of parameters in unrestricted model, r = 20

F = ((RSS$_1$ - RSS$_0$)/s)/(RSS$_0$/(N-r)) = 28.6139

$F^{95}_{4,inf}$ = 2.3788

F > $F^{95}_{4,inf}$

So, from the F-test we see that that the change in RSS of ARMA(6,9) model is significant as compared to the ARMA(10,9) model. Hence, the ARMA(6,9) parsimonious model is inadequate and we can't use this model.

3.4 Discussion

We have confirmed that the seasonality of 12 and 6 exists in our model, but the combined model was found to be inadequate. Hence, the parsimonious model with seasonality of 12 is used ahead as seasonality of 12 coincides the rainfall pattern in India. The monsoon season comes once every year from May-August and thus it makes more sense to use the model with seasonality of 12. The parameters of the parsimonious model ARMA(8,9) model is tabulated below along with the original ARMA(10,9) model.

Table 3 : Parsimonious model

Parameter	Order of the Model	
	Original Model (10,9)	$1 - 1.7321B + B^2$ (8,9)
ϕ_1	0.89 ± 0.0718	-0.8403 ± 0.0715
ϕ_2	-1.4085 ± 0.0854	-1.8650 ± 0.0550
ϕ_3	0.707 ± 0.1480	-1.6810 ± 0.1431
ϕ_4	-0.6833 ± 0.1481	-1.7301 ± 0.1100
ϕ_5	-0.3656 ± 0.1710	-1.6807 ± 0.1431
ϕ_6	0.4489 ± 0.1439	-0.7306 ± 0.1099
ϕ_7	-1.2542 ± 0.1246	-0.8401 ± 0.0716
ϕ_8	0.8564 ± 0.0953	0.1344 ± 0.0550
ϕ_9	-1.0714 ± 0.0550	–
ϕ_{10}	0.1327 ± 0.0549	–
θ_1	0.7633 ± 0.0490	0.7693 ± 0.0487
θ_2	-1.3336 ± 0.0836	-1.3313 ± 0.0830
θ_3	0.5676 ± 0.1453	0.5706 ± 0.1440
θ_4	-0.6703 ± 0.1670	-0.6655 ± 0.1656
θ_5	-0.3826 ± 0.1936	-0.3784 ± 0.1919
θ_6	0.3196 ± 0.1681	0.3247 ± 0.1667
θ_7	-1.1439 ± 0.1460	-1.1398 ± 0.1449
θ_8	0.6532 ± 0.0847	0.6554 ± 0.0840
θ_9	-0.9473 ± 0.0490	-0.9454 ± 0.0488
RSS	8.1381e+05	8.1430e+05

3.5 Forecasting

3.5.1 Forecasting Data

Rainfall is forecasted for the next five years from the period Jan 2011 to Dec 2015. Forecast of every month is made. Forecast of rainfall up to 60 months ahead (Dec 2015) is made using the parsimonious model. So, it forecasts up to 60 steps ahead. The forecast is then compared with the actual data. Confidence interval is calculated using the Green's function. The forecast data along with 95% confidence interval and actual data is shown in figure below.

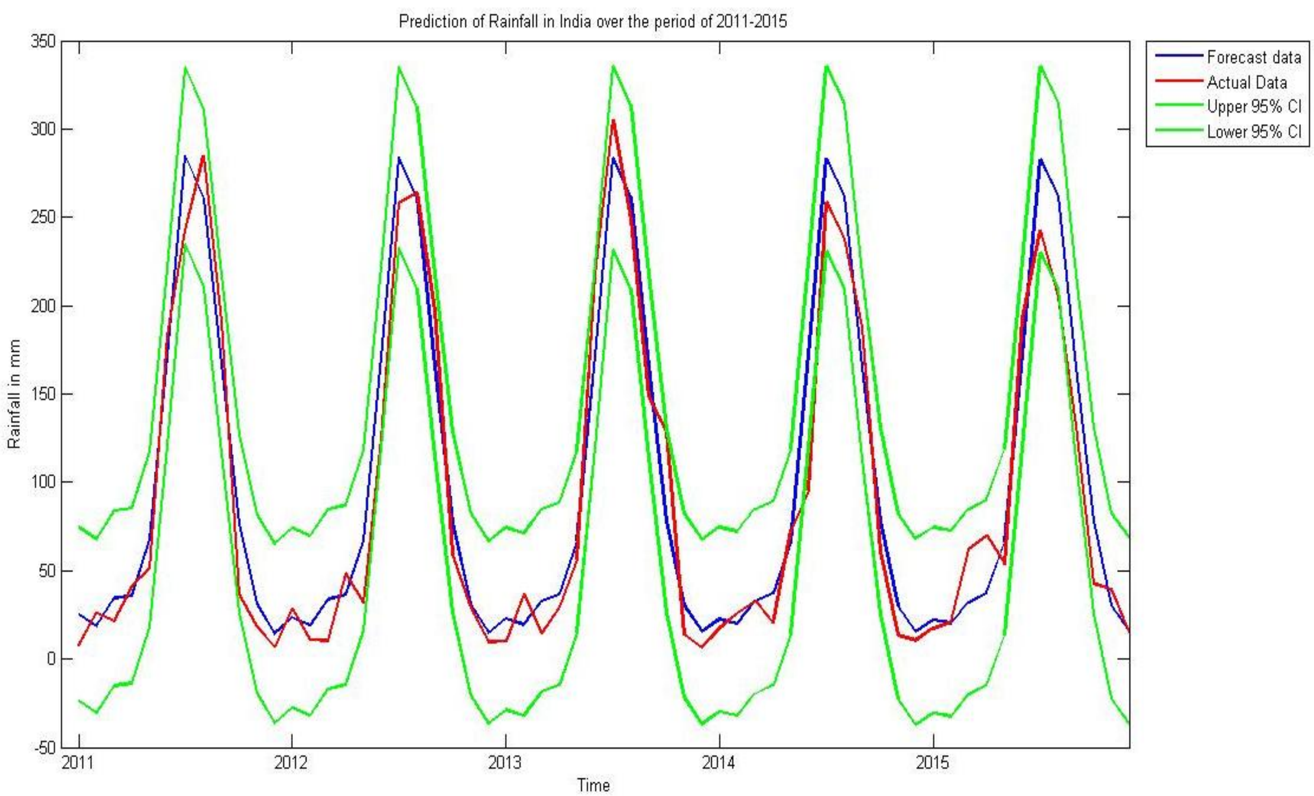


Figure 6 : Plot of the Forecast data using stationary model

3.5.2 Analysis of Forecast Data

The prediction of rainfall is quite accurate. The prediction is accurate even up to 60 steps ahead. The model is able to predict the rainfall up to five years with reasonable accuracy. The total prediction error is 3.7309×10^4. The prediction is not accurate when the peak rainfall is lower than normal. The variance of the prediction is stable and it does not diverge as we predict ahead in time, even up to 60 steps ahead (5 years ahead). Actual data always lies within the confidence interval, thus the model's prediction is quite good and can be used to predict the rainfall. The 95% confidence interval of the prediction was calculated using the Green's function value. The predicted value and its confidence interval is provided in the Appendix B. Please see below figure for the plot of standard deviation of the prediction. As we can see from the plot, it is very stable.

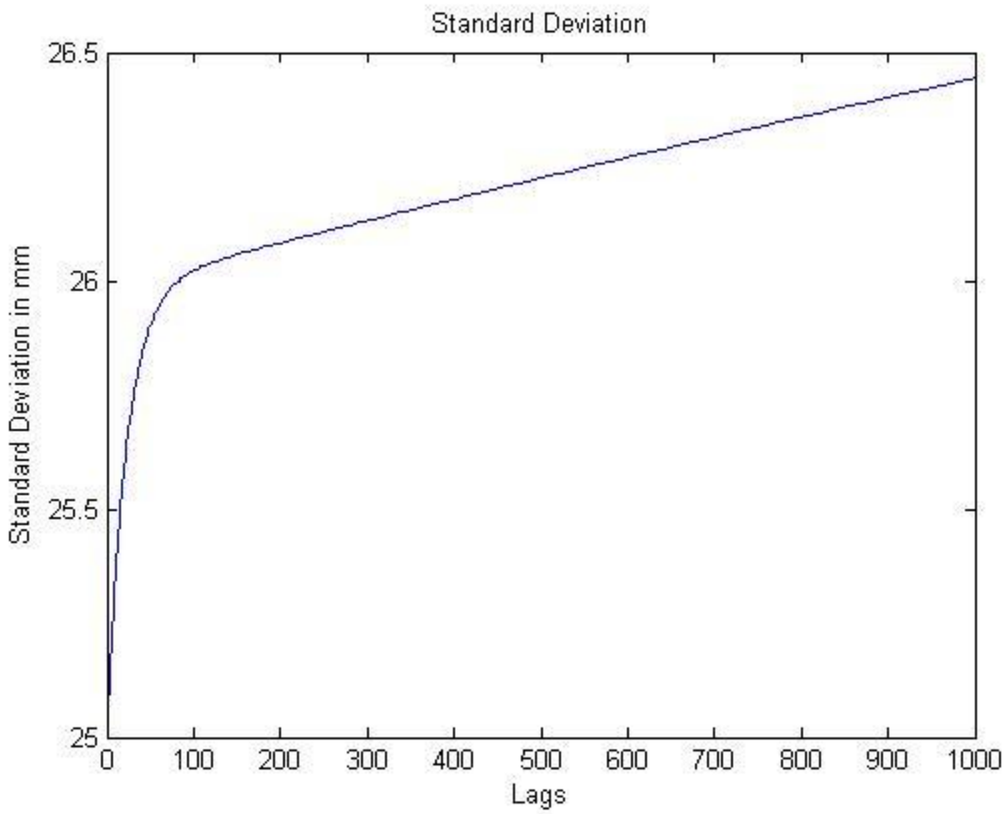


Figure 7 : Plot of the Standard Deviation of the Prediction

4. Non-Stationary Time Series Modeling

4.1 Modeling Process

Rainfall time series is modeled as non-stationary time series and the model is obtained. The model is broken down in two parts, deterministic and stochastic. Non-stationary part of the series is modeled using the deterministic part and is then removed from the series. Increasing order of deterministic model is fit in the series, till the F-criterion says it is adequate. Then the deterministic part is removed from the series. The resulting series is stationary, and ARMA model is fitted using the modeling strategy mentioned in section 3.1.

4.2 Modeling of the Deterministic and Stochastic Parts

There is periodic trend in the data. However, there is no linear trend in the data as can be seen from the Fig1 and the AR roots that were obtained for stationary model had no roots at 1. Thus, periodic trend is fit to the model. The following equation is used to fit the periodic model.

$$Y_t = a_0 + \sum_{i=1}^{n}(a_i \cos wt + b_i \sin wt) + X_t$$

We start by fitting the periodic trend of 12, calculate the RSS. Then, we increase the order and add periodic trend of 6 and compare the RSS. Similarly, we increase the order and use the F-test to determine the adequacy of the model. Based on the F-test, adequate model was the model in which all periodic trends of period 12, 6, 4, 3, 2 were fitted to the model. Table below shows the parameters of the deterministic model.

W_1 = 2*pi/12, W_2 = 2*pi/6, W_3 = 2*pi/4, W_4 = 2*pi/3, W_5 = 2*pi/2

Table 4 : Deterministic model parameters

Parameters	W_1	W_1 and W_2	W_1, W_2, and W_3	W_1, W_2, W_3, and W_4	W_1, W_2, W_3, W_4, and W_5 (adequate model)
A0	98.79 ± 2.63	98.79 ± 1.53	98.79 ± 1.38	98.79 ± 1.35	98.79 ± 1.35
A1	-116.31 ± 3.72	-116.31 ± 2.17	-116.31 ± 1.96	-116.31 ± 1.92	-116.31 ± 1.91
A2		47.44 ± 2.17	47.44 ± 1.96	47.44 ± 1.92	47.44 ± 1.91
A3			-15.84 ± 1.96	-15.84 ± 1.92	-15.84 ± 1.91
A4				7.83 ± 1.92	7.83 ± 1.91
A5					2.06 ± 1.35
B1	-29.23 ± 3.72	-29.23 ± 2.17	-29.23 ± 1.96	-29.23 ± 1.92	-29.23 ± 1.91
B2		29.89 ± 2.17	29.89 ± 1.96	29.89 ± 1.92	29.89 ± 1.91
B3			-7.33 ± 1.96	-7.33 ± 1.92	-7.33 ± 1.91
B4				-1.48 ± 1.92	-1.48 ± 1.91
B5					0
RSS	3.1468e+06	1.0712e+06	8.7001e+05	8.2802e+05	8.2238e+05

Deterministic part is then removed from the series. The resultant series is stationary and no trend is seen in it (Fig 7).

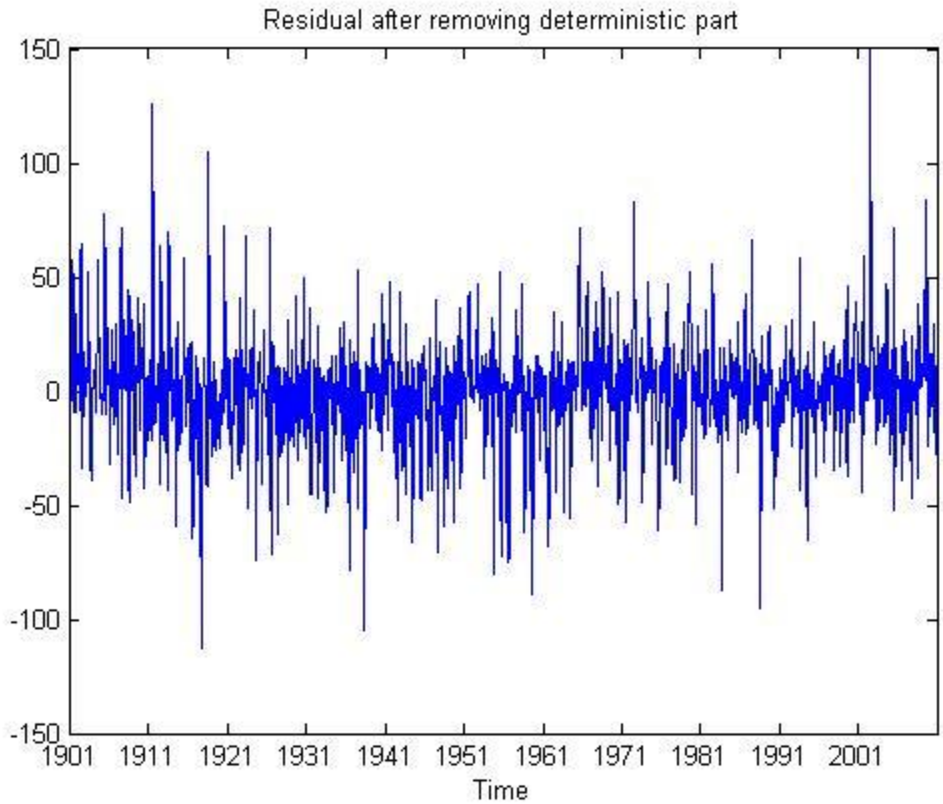


Fig 8 : Series after removing deterministic part

ARMA model is fitted to the series obtained after removing the deterministic part, using the strategy mentioned in Section 3.1. The ARMA model obtained in ARMA(3,2). Model parameters are mentioned below in the table.

Table 5 : ARMA(3,2) Model Parameters

Parameter	Order of the Model
	(3,2)
ϕ_1	0.8340 ± 0.5893
ϕ_2	0.2667 ± 0.6270
ϕ_3	-0.1120 ± 0.0725
θ_1	0.7193 ± 0.5925
θ_2	0.2516 ± 0.5708
RSS	7.8978e+05

The Autoregressive roots of the ARMA(3,2) model are 0.9892, -0.4229, 0.2677. These roots do not lie on unit circle; all of them lie inside the unit circle. Thus, ARMA(3,2) model has no stochastic seasonality and trend. The autocorrelation of the residuals is very low and ensures that the model is adequate and correct.

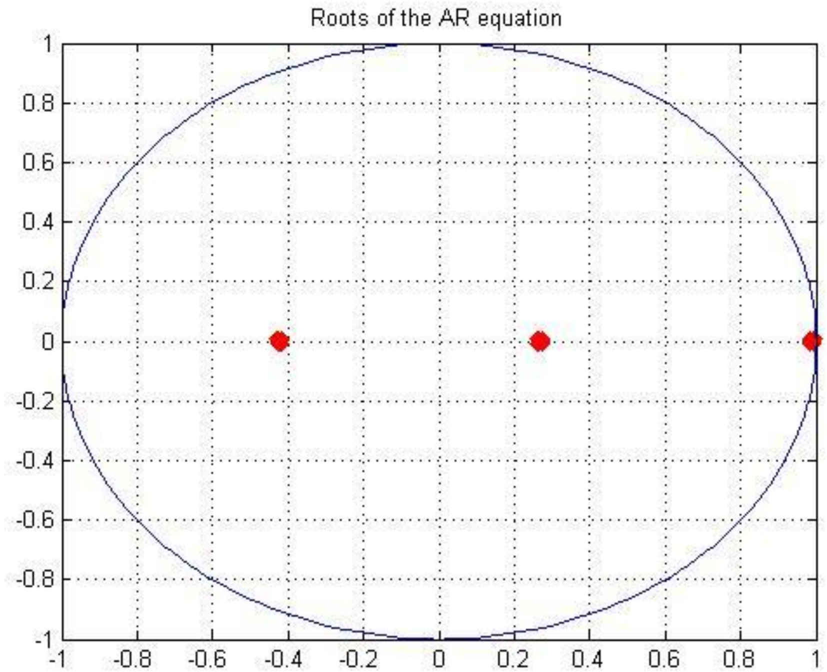


Fig 9 : AR Roots of the ARMA(3,2) model

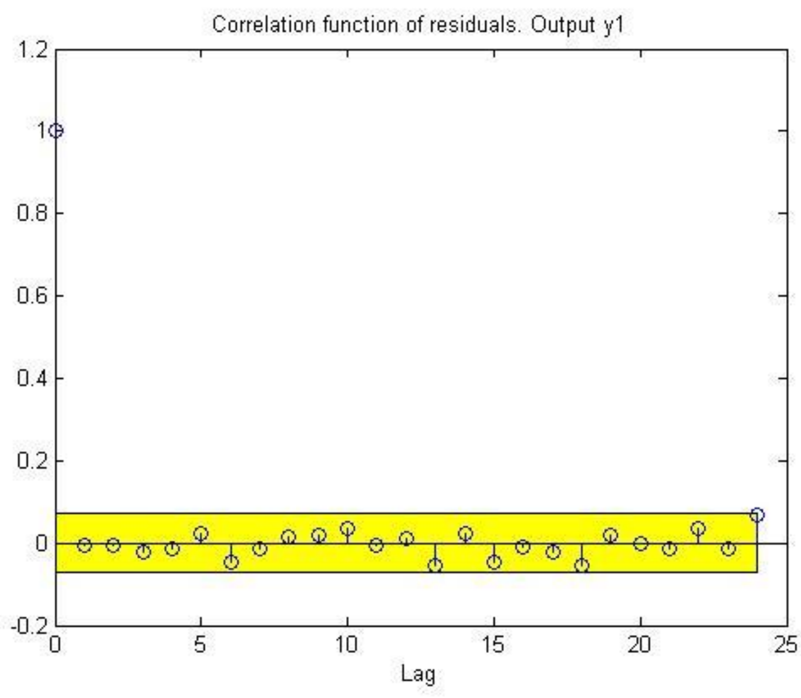


Fig 10: Autocorrelation of the residuals from Non-stationary model

4.3 Model Analysis

RSS of the non-series stationary model is less than the RSS of the stationary model. The largest coefficient in the deterministic part corresponds to period with seasonality 12. This means that the yearly seasonality dominate the series. All the autoregressive roots of the stochastic part are real and lies inside the unit circle. This means that there is no seasonality in the stochastic part.

4.4 Forecasting

Rainfall is forecasted for the next five years from the period Jan 2011 to Dec 2015. Forecast of every month is made. Forecast of rainfall up to 60 months ahead (Dec 2015) is made using the non-stationary model. So, it forecasts up to 60 steps ahead. The forecast is then compared with the actual data. Confidence interval is calculated using the Green's function. The forecast data along with 95% confidence interval and actual data is shown in figure below.

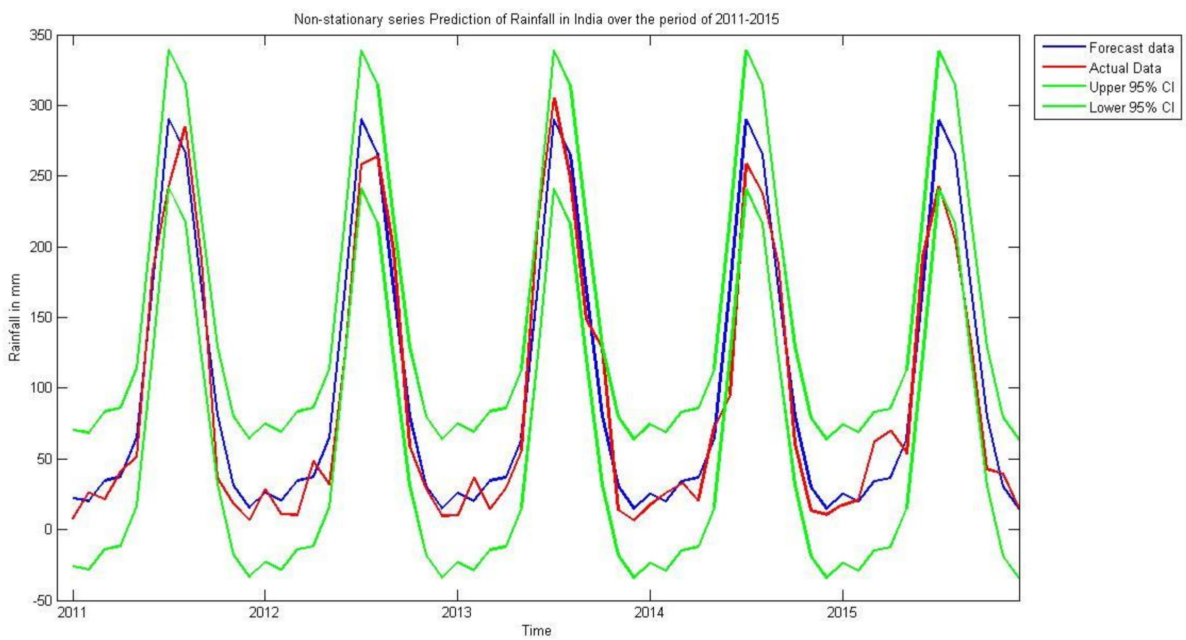


Figure 11 : Plot of the Forecast data using non-stationary model

The total prediction error is 3.9413 X 10^4. Non-stationary model does not perform better than stationary model. Prediction accuracy dropped by 5.6%. The variance of the prediction is stable and it does not diverge as we predict ahead in time, even up to 60 steps ahead (5 years ahead. The 95% confidence interval of the prediction was calculated using the Green's function value. The predicted value and its confidence interval are provided in the Appendix C.

5. Multiple Time Series Model

5.1 Modeling Process

Amount of rainfall in a particular month is related to the temperature of that month. Average Temperature for every month in India from Jan 1901 to Dec 2015 in Celsius is used along with the rainfall data to construct two time series. ARMAV modeling method is used to construct the model.

Two ARMAV model is constructed. First one with rainfall as output and temperature as the input. Second one with temperature as the output and rainfall as the input. The adequate model for the first case is ARMAV(4,4,3). For the second case the adequate model obtained is ARMAV(7,7,6). The adequate model is obtained using the F-test, by repeatedly fitting increasing order model ARMAV(n,n,n-1) to the data till the F-test says that the change is insignificant. Autocorrelation and cross-correlation of the residuals are then checked to confirm the adequacy of the model. Model parameters are mentioned in the table below.

Table 6 : ARMAV(4,4,3) model with rainfall as output and temperature as input

Parameter	Order of the Model
	(4,4,3)
ϕ_{1_11}	0.5304 ± 0.0766
ϕ_{2_11}	0.3051 ± 0.0322
ϕ_{3_11}	-1.0211 ± 0.0302
ϕ_{4_11}	0.2087 ± 0.0694
ϕ_{1_12}	-7.9190 ± 2.3122
ϕ_{2_12}	9.4155 ± 1.5094
ϕ_{3_12}	14.0274 ± 1.5806
ϕ_{4_12}	-4.195 ± 2.526
θ_{1_11}	0.2325 ± 0.0500
θ_{2_11}	0.4452 ± 0.0356
θ_{3_11}	-0.7873 ± 0.0402
RSS	1.3272e+06

Table 6 : ARMAV(7,7,6) model with temperature as output and rainfall as input

Parameter	Order of the Model
	(7,7,6)
ϕ_{1_21}	-0.0050 ± 0.0009
ϕ_{2_21}	0.0152 ± 0.0046
ϕ_{3_21}	-0.0207 ± 0.0129
ϕ_{4_21}	0.0132 ± 0.0204
ϕ_{5_21}	-0.0007 ± 0.0191
ϕ_{6_21}	-0.0046 ± 0.0105
ϕ_{7_21}	0.0022 ± 0.0026
ϕ_{1_22}	3.577± 0.6601
ϕ_{2_22}	-5.8984 ± 2.5605
ϕ_{3_22}	5.4694 ± 4.6511
ϕ_{4_22}	-2.7081 ± 4.9670
ϕ_{5_22}	0.3659 ± 3.1988
ϕ_{6_22}	0.2095 ± 1.1137
ϕ_{7_22}	-0.0266 ± 0.1395
θ_{1_22}	3.2790 ± 0.6634
θ_{2_22}	-4.9861 ± 2.3672
θ_{3_22}	4.1255 ± 3.9842
θ_{4_22}	-1.6285 ± 3.8666
θ_{5_22}	-0.0706 ± 2.1654
θ_{6_22}	0.2241 ± 0.5534
RSS	265.6788

RSS of the combined models is 1.3275 X 10^6. Autocorrelation of the residuals and cross-correlation between the residuals are plotted to check for the adequacy of the model. As can be seen from the figures below, the model is adequate.

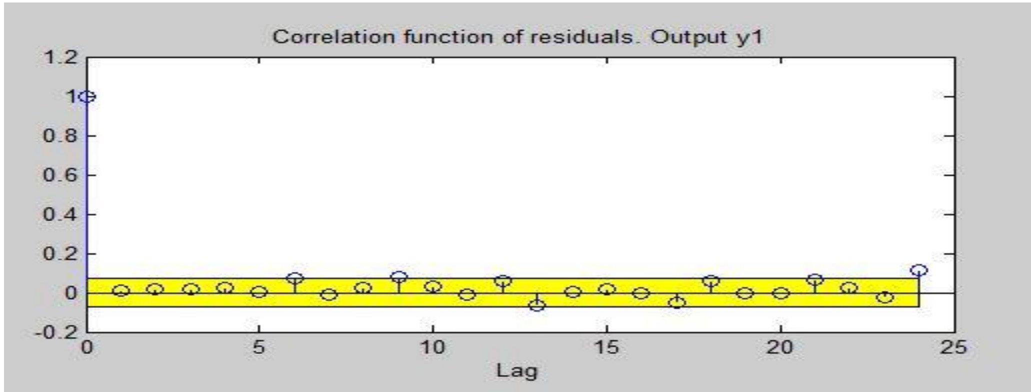


Figure 12 : Autocorrelation of residuals of ARMAV(4,4,3)

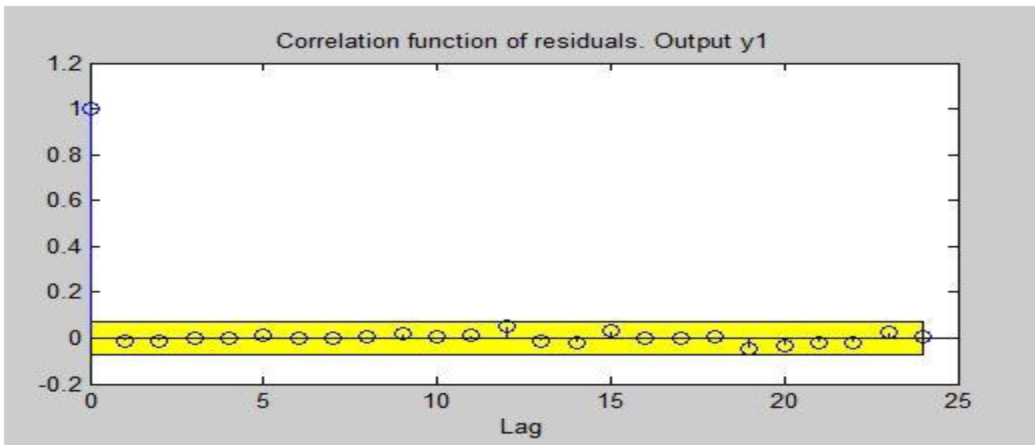


Figure 13 : Autocorrelation of residuals of ARMAV(7,7,6)

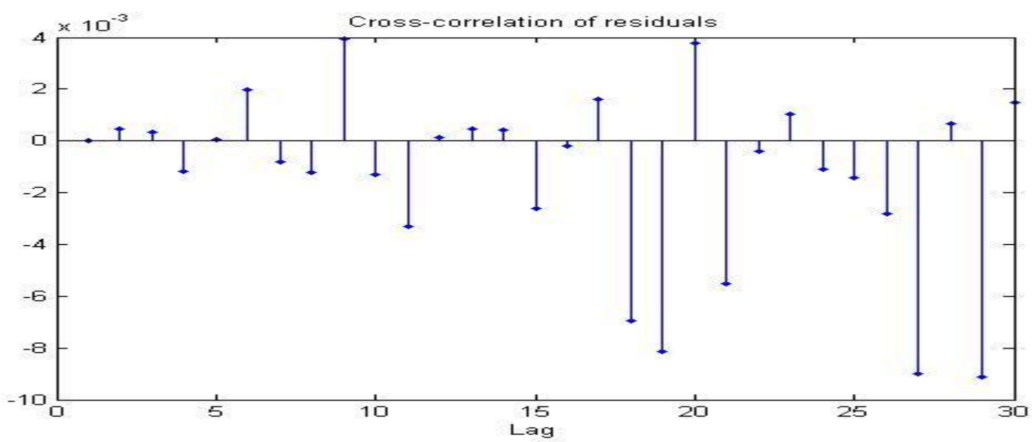


Figure 14 : Cross-correlation of residuals

5.2 Forecasting

ARMAV model is used to predict the rainfall for the five year period from Jan 2011 to Dec 2015. The total prediction error is 1.6687 X 10^5. Prediction error is higher than the stationary model. But if we observe closely, we can see that this model seems to more closely follow the actual rainfall pattern over this period. It also predicts better in cases when the rainfall is higher than normal (year 2013), and when the rainfall is lower than normal (year 2014). The prediction is worse for the 5^{th} year, suggesting that this model is good for prediction up to 4 years ahead. The predicted value is provided in Appendix D.

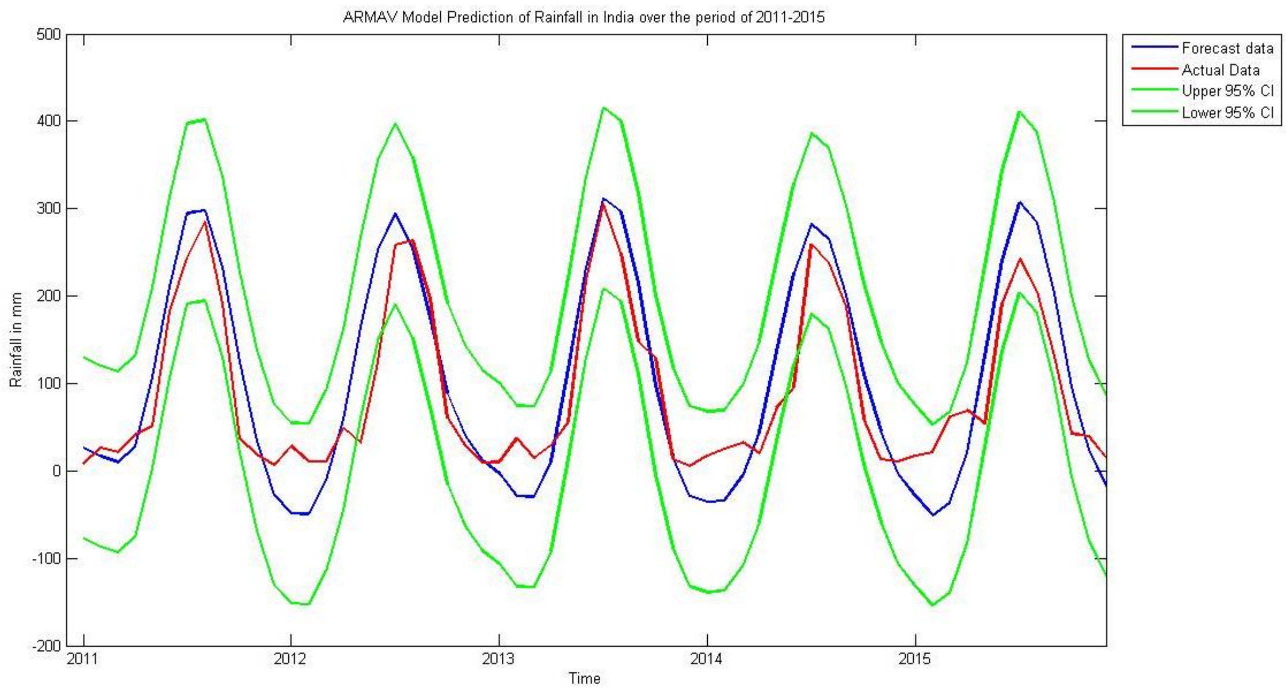


Figure 15 : Forecast using ARMAV model

6. Conclusion

Out of the three models discusses in this report, stationary model is giving the best prediction of rainfall. The predicted data overall is near the actual data. The confidence interval remains fairly stable and uniform as we predict ahead in time. The confidence interval does not diverge even up to 60 time steps ahead. Thus, we see that the models variance is low and stable, and the model could be used for prediction.

Non-stationary model has the lowest RSS among all the models. However, its prediction is not better than the stationary model. While modeling the deterministic part, an exponential decaying term could have been multiplied with the sinusoidal terms to account for any decaying effect over the time. This might have improved the model and its prediction accuracy.

Prediction error in ARMAV model is higher. However, it can be seen that this model is better in predicting rainfall in cases when it is higher or lower than usual, which the other two models were not able to do. Rainfall is dependent on a number of factors like wind patterns, weather conditions, global climate. Prediction accuracy can be improved by accounting for these factors in the model.

7. Appendix

Appendix A

Data location:

Rainfall: https://data.gov.in/catalog/rainfall-india

Temperature: https://data.gov.in/catalog/all-india-seasonal-and-annual-mean-temperature-series

Appendix B

Table 7 : Forecast of the parsimonious stationary model ARMA(8,9) from Jan 2011 to Dec 2015

Time	Actual Value	Forecast Value	Upper 95% CI	Lower 95% CI
Jan-11	7.7	25.35157898	74.40581	-23.7026
Feb-11	26.3	18.55874621	67.68971	-30.5722
Mar-11	21.4	34.394415	83.6331	-14.8443
Apr-11	41	35.83163664	85.14661	-13.4833
May-11	51.6	67.83081899	117.2372	18.42445
Jun-11	182.8	170.7453548	220.2187	121.2721
Jul-11	243	284.6168801	334.1432	235.0906
Aug-11	284.6	260.5911058	310.1891	210.9931
Sep-11	190.5	166.6444605	216.3	116.989
Oct-11	36.5	77.16603199	126.8558	27.47628
Nov-11	18.4	31.28618256	81.03803	-18.4657
Dec-11	6.5	14.25344902	64.09044	-35.5835
Jan-12	28.5	23.44381214	73.32731	-26.4397
Feb-12	10.8	19.06095874	68.97346	-30.8515
Mar-12	10.6	33.68619678	83.63529	-16.2629
Apr-12	48.5	36.48280608	86.48075	-13.5151
May-12	32.1	67.24750083	117.3018	17.1932
Jun-12	125	171.384523	221.4776	121.2914
Jul-12	258.1	283.9578576	334.0799	233.8358
Aug-12	263.9	261.1785356	311.3442	211.0129
Sep-12	197.7	166.0594257	216.2583	115.8606
Oct-12	59.8	77.73116217	127.9467	27.51563
Nov-12	28.9	30.76481302	81.01803	-19.4884
Dec-12	9.6	14.82290346	65.13264	-35.4868
Jan-13	10	22.8786855	73.21508	-27.4577

Feb-13	36.9	19.56611086	69.91675	-30.7845
Mar-13	14.5	33.15601019	83.52665	-17.2146
Apr-13	29.4	36.98477745	87.38503	-13.4155
May-13	56.5	66.81351207	117.2496	16.37738
Jun-13	217.9	171.8780082	222.3367	121.4193
Jul-13	305.4	283.4414541	333.9157	232.9672
Aug-13	248.4	261.626493	312.1277	211.1253
Sep-13	148.4	165.610774	216.1312	115.0903
Oct-13	128.9	78.16296703	128.6909	27.635
Nov-13	13.7	30.37384191	80.92503	-20.1773
Dec-13	6.2	15.26471431	65.85463	-35.3252
Jan-14	17.3	22.43845469	73.04362	-28.1667
Feb-14	25.9	19.94912494	70.5608	-30.6626
Mar-14	32.6	32.7452641	83.36767	-17.8771
Apr-14	20.2	37.36998772	88.0106	-13.2706
May-14	72.8	66.49370609	117.1577	15.8297
Jun-14	95.5	172.2597186	222.9369	121.5826
Jul-14	258.8	283.0344385	333.7197	232.3492
Aug-14	237.9	261.9674735	312.6697	211.2653
Sep-14	187.9	165.2667196	215.9801	114.5533
Oct-14	60.9	78.49253696	129.2089	27.77619
Nov-14	13.5	30.08285887	80.81375	-20.648
Dec-14	10.4	15.60862705	66.36695	-35.1497
Jan-15	17.4	22.09402464	72.86103	-28.673
Feb-15	21	20.23845977	71.00808	-30.5312
Mar-15	62	32.42612235	83.20133	-18.3491
Apr-15	69.4	37.66566401	88.45223	-13.1209
May-15	53.8	66.26145128	117.0637	15.45923
Jun-15	192.8	172.5557035	223.3656	121.7458
Jul-15	242.4	282.7113308	333.5252	231.8974
Aug-15	205.2	262.2263998	313.0512	211.4016
Sep-15	131.8	165.002885	215.8342	114.1716
Oct-15	42.9	78.74371584	129.5759	27.91155
Nov-15	39.4	29.86854153	80.70999	-20.9729
Dec-15	15	15.87745414	66.739	-34.9841

Appendix C

Table 8 : Forecast of the non-stationary model from Jan 2011 to Dec 2015

Time	Actual Value	Forecast Value	Upper 95% CI	Lower 95% CI
Jan-11	7.7	22.15487438	70.3907535602831	-26.0810
Feb-11	26.3	20.08495112	68.6370416908338	-28.4671
Mar-11	21.4	34.64103099	83.4866068298620	-14.204
Apr-11	41	37.42779085	86.2762812982183	-11.4206
May-11	51.6	65.19952936	114.064055226741	16.3350034983410
Jun-11	182.8	172.8752653	221.743324489786	124.007206165084
Jul-11	243	290.4454574	339.319492136070	241.571422697263
Aug-11	284.6	265.9478984	314.826308132691	217.069488633656
Sep-11	190.5	170.1215826	219.004770566306	121.238394664752
Oct-11	36.5	81.09342648	129.981054614238	32.2057983513278
Nov-11	18.4	31.10706768	79.9991316959307	-17.7849963422
Dec-11	6.5	15.1766318	64.0729962800477	-33.71973
Jan-12	28.5	25.99605389	74.8966425355619	-22.90453486255
Feb-12	10.8	20.4452225	69.3499375081392	-28.4594923032
Mar-12	10.6	34.73508885	83.6438443110208	-14.1736666220
Apr-12	48.5	37.17223471	86.0849424948954	-11.74047303531
May-12	32.1	64.97114673	113.887722319508	16.0545711402833
Jun-12	125	172.6060942	221.526454203516	123.685734218734
Jul-12	258.1	290.1886642	339.112727284231	241.264601104663
Aug-12	263.9	265.6875053	314.615191811454	216.759818788132
Sep-12	197.7	169.8660562	218.797288234444	120.934824175479
Oct-12	59.8	80.83961262	129.774313922853	31.9049113185155
Nov-12	28.9	30.85638333	79.7944793176678	-18.0817126497
Dec-12	9.6	14.92846936	63.8698870361089	-34.01294214339
Jan-13	10	25.75063773	74.6953056935790	-23.194030230275
Feb-13	36.9	20.20241903	69.1502674035524	-28.7454491264
Mar-13	14.5	34.49491454	83.4458749632127	-14.45604859738
Apr-13	29.4	36.93464253	85.8886481120741	-12.019363513438
May-13	56.5	64.73611681	113.693102100851	15.7791315217538
Jun-13	217.9	172.3735956	221.333496556417	123.413694639109
Jul-13	305.4	289.958671	338.921425002559	240.995917064497
Aug-13	248.4	265.45999	314.425535667511	216.494444325955
Sep-13	148.4	169.6409923	218.609269701025	120.672714931514
Oct-13	128.9	80.61697363	129.587924029763	31.6460232233210
Nov-13	13.7	30.63614315	79.6097091435788	-18.33748373348
Dec-13	6.2	14.71060212	63.6867275079991	-34.265523208

Jan-14	17.3	25.53511789	74.5137476789886	-23.443519052930
Feb-14	25.9	19.98922128	68.9703016786506	-28.99185917647
Mar-14	32.6	34.28401387	83.2674922332511	-14.6994645252
Apr-14	20.2	36.72601418	85.7118390094456	-12.25981418317
May-14	72.8	64.52973631	113.517857192420	15.5416154241987
Jun-14	95.5	172.1694387	221.159806346772	123.179071091404
Jul-14	258.8	289.7567138	338.749279938757	240.764147703272
Aug-14	237.9	265.2602088	314.254926143034	216.265491357572
Sep-14	187.9	169.4433636	218.440186059861	120.446541122546
Oct-14	60.9	80.42147423	129.420356570342	31.4225918911799
Nov-14	13.5	30.44275014	79.4436481224687	-18.55878337135
Dec-14	10.4	14.51929281	63.5221631436466	-34.483577278960
Jan-15	17.4	25.34586981	74.3506701552456	-23.658930320737
Feb-15	21	19.80201224	68.8087011482444	-29.20467771435
Mar-15	62	34.09882189	83.1073588210900	-14.9097172082
Apr-15	69.4	36.54281754	85.5531628158907	-12.4675272378
May-15	53.8	64.34851349	113.360628296406	15.3363986890841
Jun-15	192.8	171.9901685	221.004014808844	122.976322129340
Jul-15	242.4	289.5793751	338.594915804037	240.563834393755
Aug-15	205.2	265.0847807	314.101979444035	216.067582046009
Sep-15	131.8	169.2698257	218.288646819254	120.251004612619
Oct-15	42.9	80.24980612	129.270214803638	31.2293974373534
Nov-15	39.4	30.27293165	79.2948938406435	-18.74903335158
Dec-15	15	14.35130401	63.3747863555538	-34.67217896107

Appendix D

Table 9 : Forecast of the ARMAV model from Jan 2011 to Dec 2015

Time	Actual Value	Forecast Value	Upper 95% CI	Lower 95% CI
Jan-11	7.7	26.45354537	129.8172809	-76.91019018
Feb-11	26.3	16.36471986	119.7284554	-86.99901569
Mar-11	21.4	9.849646021	113.2133816	-93.51408953
Apr-11	41	28.08580506	131.4495406	-75.27793048
May-11	51.6	107.6214407	210.9851763	4.257705163
Jun-11	182.8	210.2944778	313.6582134	106.9307423
Jul-11	243	294.4141057	397.7778412	191.0503701
Aug-11	284.6	297.9970667	401.3608023	194.6333312
Sep-11	190.5	233.3194843	336.6832199	129.9557488
Oct-11	36.5	124.6382247	228.0019603	21.27448917
Nov-11	18.4	34.43290586	137.7966414	-68.93082968
Dec-11	6.5	-27.82268268	75.54105286	-131.1864182
Jan-12	28.5	-48.51011765	54.85361789	-151.8738532
Feb-12	10.8	-49.38629952	53.97743603	-152.7500351
Mar-12	10.6	-11.19571338	92.16802217	-114.5594489
Apr-12	48.5	58.53818926	161.9019248	-44.82554629
May-12	32.1	164.0300484	267.393784	60.66631288
Jun-12	125	252.3161315	355.6798671	148.952396
Jul-12	258.1	293.8730498	397.2367854	190.5093143
Aug-12	263.9	255.6083903	358.9721258	152.2446547
Sep-12	197.7	176.8323365	280.1960721	73.46860097
Oct-12	59.8	89.12463171	192.4883673	-14.23910383
Nov-12	28.9	40.02900125	143.3927368	-63.33473429
Dec-12	9.6	11.72621257	115.0899481	-91.63752298
Jan-13	10	-3.141999278	100.2217363	-106.5057348
Feb-13	36.9	-29.09282328	74.27091227	-132.4565588
Mar-13	14.5	-29.75819893	73.60553662	-133.1219345
Apr-13	29.4	11.93318822	115.2969238	-91.43054733
May-13	56.5	117.611105	220.9748406	14.24736948
Jun-13	217.9	233.1956733	336.5594089	129.8319378
Jul-13	305.4	311.7288355	415.0925711	208.3651
Aug-13	248.4	296.8163939	400.1801294	193.4526583
Sep-13	148.4	214.1333251	317.4970606	110.7695895
Oct-13	128.9	98.64037241	202.004108	-4.723363141
Nov-13	13.7	16.52973128	119.8934668	-86.83400426
Dec-13	6.2	-29.23252863	74.13120691	-132.5962642

Jan-14	17.3	-35.51495401	67.84878154	-138.8786896
Feb-14	25.9	-33.90714832	69.45658723	-137.2708839
Mar-14	32.6	-6.60994992	96.75378563	-109.9736855
Apr-14	20.2	45.02717811	148.3909137	-58.33655744
May-14	72.8	137.2919563	240.6556919	33.92822079
Jun-14	95.5	225.9476048	329.3113403	122.5838692
Jul-14	258.8	282.7954847	386.1592203	179.4317492
Aug-14	237.9	265.4700735	368.8338091	162.106338
Sep-14	187.9	200.7988945	304.16263	97.43515892
Oct-14	60.9	111.0228172	214.3865527	7.659081628
Nov-14	13.5	45.62281093	148.9865465	-57.74092462
Dec-14	10.4	-3.002289816	100.3614457	-106.3660254
Jan-15	17.4	-28.88331437	74.48042118	-132.2470499
Feb-15	21	-50.88695399	52.47678156	-154.2506895
Mar-15	62	-36.32620551	67.03753003	-139.6899411
Apr-15	69.4	20.30459229	123.6683278	-83.05914325
May-15	53.8	131.4205183	234.7842538	28.05678271
Jun-15	192.8	240.4413845	343.8051201	137.077649
Jul-15	242.4	307.2537483	410.6174838	203.8900127
Aug-15	205.2	284.1763101	387.5400456	180.8125745
Sep-15	131.8	202.9236376	306.2873732	99.55990207
Oct-15	42.9	96.37519433	199.7389299	-6.988541212
Nov-15	39.4	24.58205452	127.9457901	-78.78168103
Dec-15	15	-16.90529386	86.45844168	-120.2690294