JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

FORECASTING ANNUAL RAINFALL DATA IN KARNATAKA USING ARMA MODEL AND CALCULATIONS OF EXCEEDANCE PROBABILITIES

Satish R. Huddar1, Prof. K. Sivasubramaniyan2
[1]Department of Minor Irrigation, Govt. of Karnataka, Bengaluru-560001, India.
[2]Department of Economics, St. Peter's Institute of Higher Education, Avadi, Chennai-600054, India

ABSTRACT: The Karnataka gets significant amount of rainfall, which is influenced by the diversified topographical and climatic zones. There are mainly four climatic regions in Karnataka; one, Costal region, which receives highest average annual rainfall, o more than 3500 millimetres. Secondly, it is *Malnad* region, which gets an average annual rainfall of about 2000 millimetres. However, both the North Interior Karnataka and South Interior Karnataka Regions receive an average precipitation of about 700 millimetres respectively. This study analyses the annual rainfall trends in Karnataka from 1989 to 2023 and forecasts the expected rainfall for 2024-2028 using an Auto Regressive Moving Average Model (ARMA) - based on time series model. The analysis identifies historical fluctuations, seasonal variations, and the impact of extreme events on annual precipitation levels. The forecast suggests a gradual recovery in rainfall levels following a significant drop in 2023. The findings provide valuable insights for policymakers, agricultural planners, and water resource management authorities. The Statistical analysis of rainfall data using time series techniques is essential for understanding long-term trends, variability and predicting future patterns. This paper employs statistical and basic probability models to examine historical rainfall data, identify trends, and provide insights into rainfall variability in Karnataka.

KEY WORDS: Auto Regressive Moving Average Model, Exceedance Probabilities, Return Period, Root Mean Squared Error.

INTRODUCTION: Rainfall is a crucial climatic factor influencing agriculture, water resources, and overall ecosystem stability in Karnataka. Understanding past trends and forecasting future rainfall can help mitigate risks associated with droughts and floods. This study examines Karnataka's historical rainfall data and employs a statistical forecasting model to predict future trends.

LITERATURE REVIEW: It is found that there is less research on rainfall data is found both by Government and private agencies in Karnataka, as data is in only raw form or very few analytical studies were taken. The main raw source of rainfall data is Directorate of Economics and Statistics, Govt. of Karnataka. Very recently, the Karnataka State Natural Disaster Monitoring Centre also providing insights into the rainfall data besides India Metrological Department. The dataset includes monthly and seasonal rainfall records from various districts of Karnataka over two decades. Pre-processing steps for data include handling missing values, normalizing data, and ensuring consistency across different time periods.

The two methods are adopted for analysing the following rainfall data.

1. Calculations of Exceedance Probabilities.
2. Ascertaining whether data is stationary by Dicky-Fuller test
3. Application of ARMA Model for Forecasting time series data of rainfall.

DATA COLLECTION: The Secondary data is collected from the report of the *Karnataka at a Glance*, the yearly publication from the Directorate of Economics and Statistics, Government of Karnataka however, the few annual data are collected from the annual reports of the Karnataka State Natural Disaster Monitoring Centre (KSNDMC), Revenue Department, Government of Karnataka.

METHODOLOGY:

1. Calculations of Exceedance Probabilities:

A rather simple, graphical method to determine the probability or frequency of occurrence of yearly rainfall will be described here.The following data representsannual rainfall pattern for 35 years in Karnataka from 1989 to 2023.

Table-1: Annual Rainfall in Karnataka from 1989-2023 (in mm)

Year	Annual	Year	Annual	Year	Annual	Year	Annual	Year	Annual	Year	Annual
1989	1130	1995	1219	2001	1194	2007	1578	2013	1409	2019	1635
1990	1227	1996	1327	2002	1033	2008	1340	2014	1414	2020	1567
1991	1252	1997	1489	2003	1036	2009	1544	2015	1160	2021	1474
1992	1095	1998	1479	2004	1271	2010	1500	2016	996	2022	1603
1993	1046	1999	1430	2005	1584	2011	1375	2017	1063	2023	959
1994	1389	2000	1420	2006	1403	2012	1094	2018	1366		

The data is initially ranked for calculation of the exceedance probabilities by using the below formula for each of the ranked observations.

$$(P\%) = \frac{m - 0.375}{N + 0.25} X100$$

Where, P=probability in % of the observation of the rank m, m=the rank of the observation, N = total number of observations used=35, T=Return Period=100/P. There are several other, but similar, equations known to compute experimental probabilities.

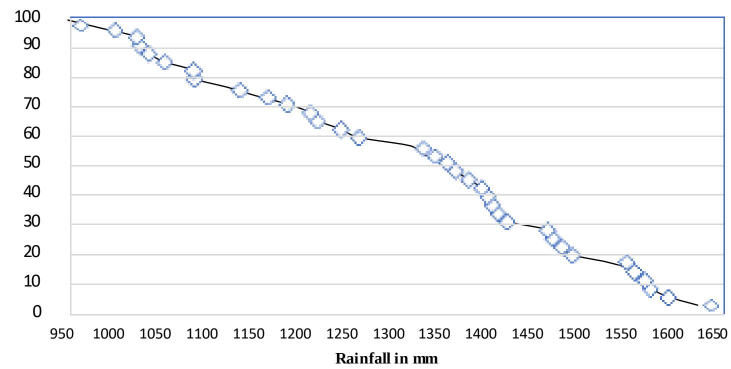
Thus, the following Probabilities of exceedance and the Return Periods are obtained.

Table-2: Ranked Annual Rainfall in Karnataka from 1989-2023 (35 Years) (in mm)

Rank (m)	Year	Annual Rainfall in mm	P(%)	T=100/P (years)	Rank (m)	Year	Annual Rainfall in mm	P(%)	T=100/P (years)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	2019	1635	2	56	19	2008	1340	53	2
2	2022	1603	5	22	20	1996	1327	56	2
3	2005	1584	7	13	21	2004	1271	59	2
4	2007	1578	10	10	22	1991	1252	61	2
5	2020	1567	13	8	23	1990	1227	64	2
6	2009	1544	16	6	24	1995	1219	67	1
7	2010	1500	19	5	25	2001	1194	70	1
8	1997	1489	22	5	26	2015	**1160**	73	1
9	1998	1479	24	4	27	1989	1130	76	1
10	2021	**1474**	27	4	28	1992	1095	78	1
11	1999	1430	30	3	29	2012	1094	81	1
12	2000	1420	33	3	30	2017	1063	84	1
13	2014	1414	36	3	31	1993	1046	87	1
14	2013	1409	39	3	32	2003	1036	90	1
15	2006	1403	41	2	33	2002	1033	93	1
16	1994	1389	44	2	34	2016	996	95	1
17	2011	1375	47	2	35	2023	959	98	1
18	2018	1366	50	2	-	-	-	-	-

The next step is to plot the ranked observations (columns 3, 8 Table-2) against the corresponding probabilities (columns 4, 9 Table-2).

Figure-1 : Cummulative Distribution of Rainfall in Karnataka from 1989 to 2023



RESULTS-1: From this curve it is now possible to obtain the probability of occurrence or exceedance of a rainfall value of a specific magnitude. Inversely, it is also possible to obtain the magnitude of the rain corresponding to a given probability.

In the above example, the annual rainfall with a probability level of 73 percent of exceedance is 1160 mm (Figure 7), i.e. on average in 73 percent of time annual rain of 1160 mm would
be equalled or exceeded. For the probability value of 27 percent of exceedance is 1474 mm. (Table-2 and Figure-1).

The return period T (in years) can easily be derived once the exceedance probability P (%) is known from the equations. T=100/P years.

From the above, it can be seen that, the return period for the 73 percent and the 27 percent exceedance probability events would thus be:

$T_{73\%}$=100/73=1.4 years.
i.e. on average an annual rainfall of 1160 mm or higher can be expected in 2 years out of 3;

$T_{33\%}$=100/27=3.7 years.
i.e. on average an annual rainfall of 1474 mm or more can only be expected in 1 year out of 3.

METHODOLOGY:

2. Calculations of Exceedance Probabilities:

Ascertaining whether data is stationary by Dicky-Fuller test and then application of ARMA Model for forecastingabove time series data of rainfall. (Table-1)

To check whether annual rainfall data is stationary/non-stationary, we can just see the simple graph of the data in Figure-2 drawn for the rainfall in mm versus the years. It seems to be stationary, however, statistically, one can apply the Dickey-Fuller test. It is a statistical method used to determine whether a time series is stationary or contains a unit root, indicating non-stationarity.

Let us consider a stochastic process, such that, $Y_t=B*Y_{t-1}+e_t$ Subtracting Y_{t-1} from both sides, we get,

$Y_t - Y_{t-1} = (B-1)*Y_{t-1}+ e_t$

So, if, B — 1 = 0 or B = 1 (Unit root); the process is said to be stationary. That is, we test the following hypothesis,

Null Hypothesis: $H_0 : B - 1 = 0$ Versus Alternative Hypothesis: $H_1 : B - 1 \neq 0$ Accordingly, we determine the stationary status.

So, when B — 1=0 that is null hypothesis is accepted stop differencing that means $Y_t - Y_{t-1}$ = Constant. That is data is stationary. To interpret test results, compare the p-value to a chosen significance level, typically 0.05. Here, the value of p is less than 0.05. Hence, there exists a unit root that is the data is stationary.

RESULTS-2: The excel calculations of Regression analysis is given below:

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	1325.350453	28.32877373	46.78460372	4.86619E-31
X Variable 1	0.531668994	0.134137049	3.963625248	0.00038788

METHODOLOGY:

3. Fitting Auto Regressive Moving Average Model(ARMA) for Rainfall data:

So, the data is stationary and we can apply **ARMA Model** for calculation of forecasts. Here, we have used ARMA(1,1) as below:

The **ARMA(p,q)** is the linear combination of past actual periods and the error terms and is given by;

$$y_t = \phi_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + - \ldots + \theta_q e_{t-q} + \epsilon$$

However, as the above data in Table-1 is Stationary and there exists a positive Auto Correlation and Partial Auto Correlation, we can fit the model for ARMA(1,1) as below.

Steps:

i. Centre each data by ($Y_t - Y_{mean}$) as in column 3.

ii. Consider initial residual as zero.

iii. Choose auto-regressive coefficient(β_1) and error coefficient (θ_1) as 0.5 each arbitrarily.

iv. Calculate forecasts as the linear combination of centred data and residual of the previous period.

v. Calculate, Root Mean Squared Error (RMSE) = Square root of [Sum of Squared residuals÷No. of observations](RMSE=196.38)

vi. Now, find the minimum value of RMSE, for the optimum vales of (β_1) and (θ_1) by using excel solver. (They are RMSE=175.79, for the optimum vales of (β_1)=0.3 and (θ_1)=0.2)

vii. Hence, the final forecast for these optimised RMSE, (β_1) and (θ_1) are obtained.

Results-3:

Final Forecast= $Y_{t+1} = Y_{mean}(1-\beta_1) + \beta_1 y_t + \theta_1 e_{t-1}$

Year	Annual Rainfall in mm (Yt)	Centred Data (Yt-Ymean)	Residue (initially it is 0) Centered data minus forecast	Initial Forecast ARMA	Residue or error Square
1989	1130	-187.17	0	-	0.00
1990	1227	-90.17	-42.91	-47.26	1841.41
1991	1252	-65.17	-32.69	-32.48	1068.67
1992	1095	-222.17	-198.32	-23.85	39329.41
1993	1046	-271.17	-170.18	-100.99	28962.90
1994	1389	71.83	178.82	-106.99	31976.48
1995	1219	-98.17	-156.78	58.61	24581.25
1996	1327	9.83	70.10	-60.28	4914.69
1997	1489	171.83	153.48	18.35	23555.64
1998	1479	161.83	83.70	78.13	7006.11
1999	1430	112.83	53.02	59.81	2811.28
2000	1420	102.83	62.34	40.49	3886.08
2001	1194	-123.17	-163.25	40.07	26649.11
2002	1033	-284.17	-216.12	-68.05	46707.95
2003	1036	-281.17	-160.50	-120.67	25760.35
2004	1271	-46.17	61.15	-107.32	3739.63
2005	1584	266.83	264.64	2.18	70036.74
2006	1403	85.83	-41.45	127.28	1717.87
2007	1578	260.83	248.54	12.29	61771.66
2008	1340	22.83	-99.29	122.12	9857.85
2009	1544	226.83	243.54	-16.71	59310.92
2010	1500	182.83	70.43	112.40	4960.39
2011	1375	57.83	-4.28	62.11	18.29
2012	1094	-223.17	-236.80	13.63	56076.50
2013	1409	91.83	201.78	-109.95	40714.99
2014	1414	96.83	27.97	68.86	782.26
2015	1160	-157.17	-187.95	30.78	35325.56
2016	996	-321.17	-238.94	-82.23	57093.89
2017	1063	-254.17	-118.99	-135.18	14159.08
2018	1366	48.83	139.94	-91.11	19583.09
2019	1635	317.83	273.82	44.00	74979.52
2020	1567	249.83	107.60	142.23	11577.29
2021	1474	156.83	69.39	87.44	4815.42
2022	1603	285.83	230.52	55.31	53140.80
2023	959	-358.17	-482.52	124.35	232826.57

Thus, the Forecasts for the years **2024:1118, 2025:1267, 2026:1304, 2027:1314, 2028: 1316, 2029:1317, 2030:1317** and so on.

CONCLUSION: The calculations reveals that the annual rainfall data of Karnataka for the years 1989 to 2023 is Stationary, proved by Dicky-Fuller Test. The basic model for the rainfall forecasts is identified. The study reveals that Karnataka's rainfall patterns exhibit periodic fluctuations with occasional extreme years. The forecast indicates a **recovery from the 2023** and a stabilization around 1317mm from 2030 in the coming years. These findings are useful for water resource planning, agricultural decision-making, and climate adaptation strategies. There is further scope for analysing and finding forecasts for the seasonal Rainfall data in Karnataka on the same lines.

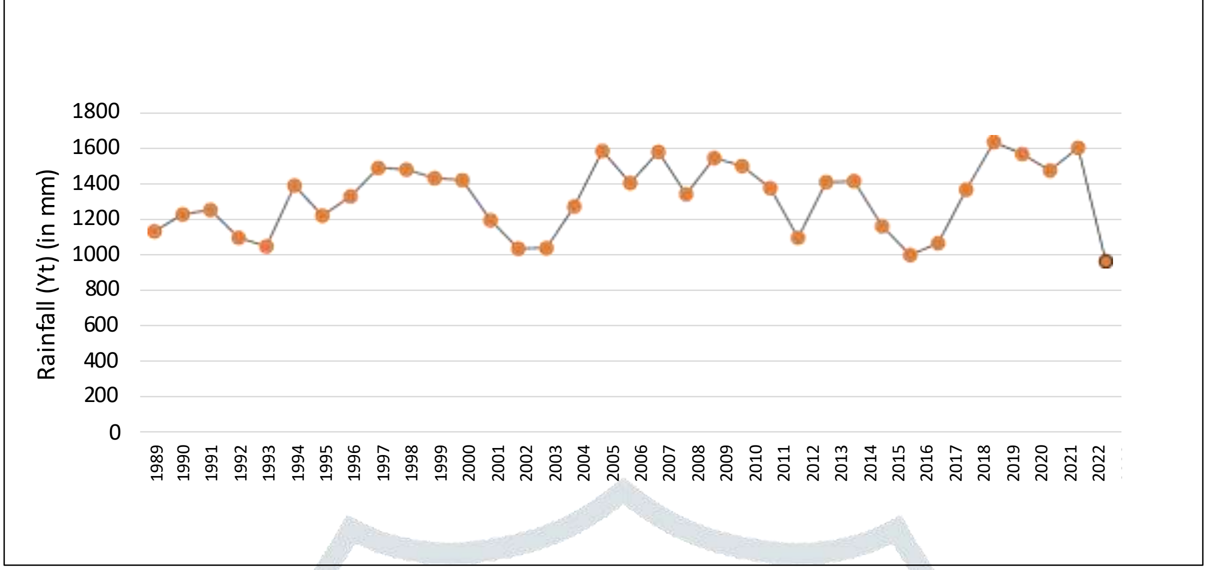


Figure-2: Pattern of Rainfall in Karnataka

REFERENCES:

1) Directorate of Economics and Statistics, Karnataka (Rainfall Data, 1989-2023)

2) Karnataka State Natural Disaster Monitoring Centre (KSNDMC), Revenue Department, Govt. of Karnataka.

3) National Programme on Technology Enhanced Learning (NPTEL), A course on Business Forecasting by Prof. Pankaj Dutta, IIT Bombay.

4) The Food and Agriculture Organization (FAO) is a specialized agency of the United Nations, a study of rainfall run-off analysis.(https://www.fao.org/4/u3160e/u3160e05.htm)

5) Box, G. E., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control.

6) Wilks, D. S. (2011). Statistical Methods in the Atmospheric Sciences.

7) Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427–431.

8) Statistical Analysis Applications to Business and Economics, Clark A. Hawkins and Jean E. Weber,

9) Box, G.E.P., Jenkins, G.M., Reinsel, G.C., & Ljung, G.M. (2015), Time Series Analysis: Forecasting and Control (5th Edition), The classic and foundational text that introduced ARMA models. Explains ARMA(p, q) theory, model fitting, and diagnostics in detail.

10) Hamilton, J.D. (1994), Time Series Analysis, Advanced graduate-level textbook with in-depth coverage of ARMA models in both theoretical and applied contexts.

11) Shumway, R.H., & Stoffer, D.S. (2017), Time Series Analysis and Its Applications: With R Examples (4th Edition).

12) Fundamentals of Applied Statistics, S. C. Gupta and V.K.Kapoor, Sultan Chand & Sons Educational Publishers, New Delhi.

13) Statistical Forecasting by Warren Gilchrist, John Wiley & Sons Publications, A Wiley-Interscience Publication.

14) Treatise on Doing Basic Research in Economics, S. Manickam, Formerly, Tamil Nadu Government Research Fellow, Deputy Director (retd.)

15) Economic survey of Karnataka, Planning, Programme Monitoring and Statistics Department, Government of Karnataka.