*Journal of Harbin Engineering University* Vol 44 No. 11
*ISSN: 1006-7043* November 2023

# Modelling And Forecasting Rainfall in East and Western Region of Tamil Nadu using Sarima Model

**Dr. S. Santha,**
Assistant professor and Head, Department of Mathematics,
1036Government Arts and Science College, Nager Coil- 629004.
Email: santhawilliam14@gmail.com

**T. Subasini,**
Reg. No: 18121172092018, Research Scholar, Rani Anna Government Arts and
Science College for Women, Tirunelveli. Affiliated toManonmaniam Sundaranar University, Abishekapatti, Tirunelveli-627012.TamilNadu, India.
Email: subasinit@gmail.com

**Abstract**

Research on modeling and rainfall forecasting is important because the findings will assist policymakers, legislators, and decision-makers in reducing the harm that drought or flooding can do to the ecology. Rainfall data from the eastern and western parts of Tamil Nadu have been fitted to time series stochastic models in this study. SARIMA (seasonal Autoregressive Integrated Moving Average) was used to model and predict the rainfall data.The purpose of this study is to model and forecast rainfall using the Seasonal Autoregressive Integrated Moving Average Model (SARIMA).The statistical departments of Coimbatore and Nagapattinam submitted 504 observations spanning 42 years of monthly rainfall data, starting in January 1981 and ending in December 2022, in order to undertake this study. The Nagapattinam and Coimbatore districts were prepared for the monthly rainfall. It was expected that there will be monthly rain in the districts of nagapattinam and Coimbatore using the seasonal ARIMA model after being applied to the collection of rainfall data. The outcome suggests that the sarima (1, 0, 0) (1, 1, 0)12 model has the greatest in sample and out sample performance when forecasting rainfall in the districts of Coimbatore and Nagapattinam.

**Keywords**: Monthly Rainfall, SARIMA, stationary, ACF plot, PACF plot.

## INTRODUCTION

According to Ikot (The Environment and Socioeconomic growth, 2021) rain is essential for all living forms, socioeconomic growth, and the preservation of thriving ecosystems. A country's agricultural performance is typically influenced by the distribution of rainfall in a particular area.

Around the world, there is a lot of discussion about rainfall. In this article, the monthly rainfall rates will be modeled using Box-Jenkins techniques. The specific strategy will be the SARIMA (seasonal autoregressive integrated moving average) strategy suggested by Box and Jenkins [11]. Recently, numerous authors have started modeling real-world data using the Sarima modeling method. According to Hang et al. [12], sarima modeling outperformed multiple linear regressions; autoregressive adjusted poison regression, and conventional poison regression. The monthly rainfall in Tamil Nadu was fitted with a sarima by Nirmala and Sundaram [13]. Jiang et al. [14] discovered that SARIMA modeling generated better predictions than those generated by Loess-based dynamic harmonic regression, also known as seasonal trend decomposition. The footfalls of Indian international tourists were modeled by Padhan1 [5] using a SARIMA (1, 1, 1) (2, 1, 4). In order to simulate rainfall in Bangladesh's Dhaka division, Mahsin et al. [16] employed a sarima (0,0,1) (0,1,1).A SARIMA (0,1,0) (0,0,1)4 was fitted by Gikungu et al. [14] to the quarterly inflation rates for Kenya. In this study, the Coimbatore and Nagapattinam districts of Tamil Nadu's rainfall were predicted and modelled using SARIMA model approaches.

## 2.Materials and Methods

### 2.1 Area of study

The Western Ghats mountain range covers Tamil Nadu's Coimbatore district on both the west and north. The district is bordered by Idukki in the south, Erode district in the northeast, Nilgiris in the north, and Palakkad in the west. The remainder of the sector has a pleasant climate for the majority of the year because it is located in the Western Ghats' rain-shadow zone. In Coimbatore city, the average high and low temperatures in the range from 35 to 18 degrees Celsius in summer and winter, respectively. In the plains, there is about 700 mm of yearly precipitation. From 1981 until 2022, data were gathered from Coimbatore's statistical department.

One of Tamil Nadu's 38 districts, Nagapattinam is a coastal district. It is located in southern India. Another area of Nagapattinam is located to the south of Karikkal and Tiruvarur districts, while the district is located on the east coast to the south of Cuddalore district. The Nagapatinam statistical department provided the information for this study, which covered the years 1981 to 2022.

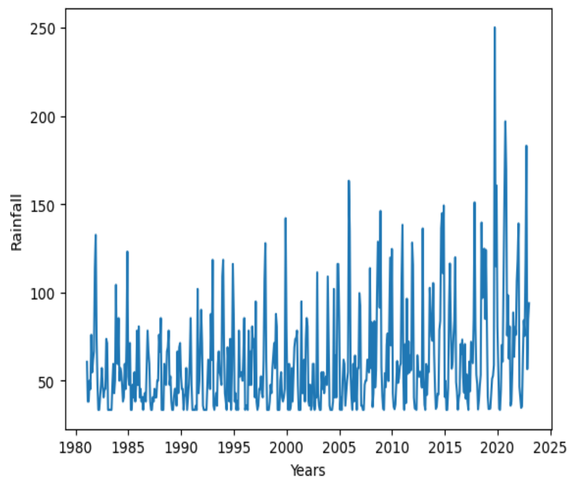


Fig 1: monthly rainfall data in Coimbatore district

## 2.3 SARIMA MODEL

Definition

ARIMA(p,d,q) (P,D,Q) Seven parameters characterise seasonal arima models

ARIMA(p,d,q)(P,D,Q)$_m$

$$\left(1-\varphi_1 A-\varphi_2 A^2-\cdots-\varphi_p A^p\right)\left(1-\beta_1 A^s-\beta_2 A^{2s}-\cdots-\beta_P A^{Ps}\right)(1-A)^d\left(1-A^s\right)^D$$

$$Y_t=\left(1-\psi_1 A-\psi_2 A^2-\cdots-\psi_q A^q\right)\left(1-\theta_1 A^s-\theta_2 A^{2s}-\cdots-\theta_Q A^{Qs}\right)\varepsilon_t$$

Where

$\left(1-\varphi_1 A-\varphi_2 A^2-\cdots-\varphi_p A^p\right)$ *is the Auto regressive part of order* $(p)$

## 2.2 Data collection

The study's primary objective was to develop a method for forecasting rainfall in the Tamil Nadu districts of Coimbatore and Nagapattinam. Monthly time series data for the years 1981 to 2022 were used to conduct the study.There were 504 observations in the data gathered from the statistical offices of the Coimbatore and Nagapattinam districts. The monthly rainfall statistics are displayed in Figs. 1 and 2. A total of 384 observations were utilized to develop the model, and 120 observations were used to carry out out-of-sample tests.

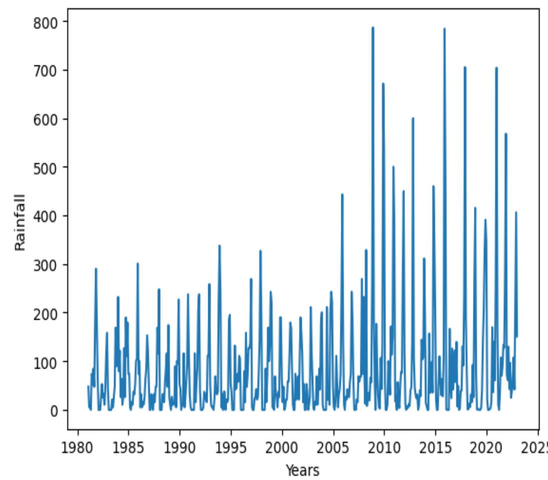


Fig 2: monthly rainfall in Nagapattinam district

$\left(1-\beta_1 A^s-\beta_2 A^{2s}-\cdots-\beta_P A^{Ps}\right)$ *is the seasonal Autoregressive part of order p*

$(1-A)^d$ *differencing order* $d(I(d))$

$\left(1-A^s\right)^D$ *Seasonal differencing of orderD* $(I_s(D))$

$\left(1-\theta_1 A^s-\theta_2 A^{2s}-\cdots-\theta_Q A^{Qs}\right)$ *Seasonal Moving average part of order Q*

$\left(1-\psi_1 A-\psi_2 A^2-\cdots-\psi_q A^q\right)$ *moving average part of order q*

*s is the period of the seasonal pattern i.e s* $=12$ *month*

2.3.1 Working Rule

SARIMA (seasonal auto-regressive integrated moving average)

Step 1: Create the graph.

Step 2: Verify the data's stationary.

Step 3: To find suitable AR and MA models by using ACF and PACF plot for the SARIMA

Step 4: The SARIMA model's best fit is determined .

Step 5: Plot ACF and PACF for the SARIMA model residual in to make sure there is no more data left to be extracted.

Step 6: Predict the series using the SARIMA that fits you the best.

2.3.2 Model performance measure

This study's goal was to provide an assessment of the SARIMA model's numerical performance using precipitation records from Coimbatore and Nagapattinam in the Indian state of Tamil Nadu. The Mean Absolute Error and the Root Mean Square Error were used to measure the accuracy of the forecasts. AIC and BIC, which are explained below, are two additional commonly used metrics for model identification.

AIC = 2p - 2ln(Ĺ)

BIC= p ln(n) – 2ln(Ĺ) where n is the total number of observations, e, the best estimate of a parameter, and L, the likelihood function.

The number of models that reduce either the AIC or the BIC determines the best model order.

2.3.3 MEAN SQUARE ERROR

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_{i)}{}^2$$

Where, n = count of observations

$y_i = observed\ data\ point,$

$$\hat{y}_i = predicted\ values$$

2.3.4 ROOT MEAN SQUARE ERROR

$$RMSE = \sqrt{MSE}$$

3. Result and Discussion

3.1 STATIONARITY TEST

For stationarity, Utilising the modified Dickey-Fuller test. In the context of the dataset, Table 1 displays the stationarity result. According to the table, seasonal variation and seasonality in the series with order 1 are both stationary.

| station | df statistic | Lag order | P value | seasonality | seasonality |
|---|---|---|---|---|---|
| | | | | decision | order |
| Coimbatore | -9.415 | 12 | 0.01 | stationary | 1 |
| Nagapattinam | -7.12 | 12 | 0.01 | stationary | 1 |

Table 1: Test for stationarity

3.2 Model Identification

Models with monthly probability time series have been found to be the most precipitation data have been utilized in Autocorrelation Function (ACF).

The monthly rainfall plots for the Coimbatore and Nagapattinam districts are shown in Figures 3, 4, 5, and 6. The figures indicate that the AR (Autoregressive Model) and MA (Moving Average) should be used in the highest order possible

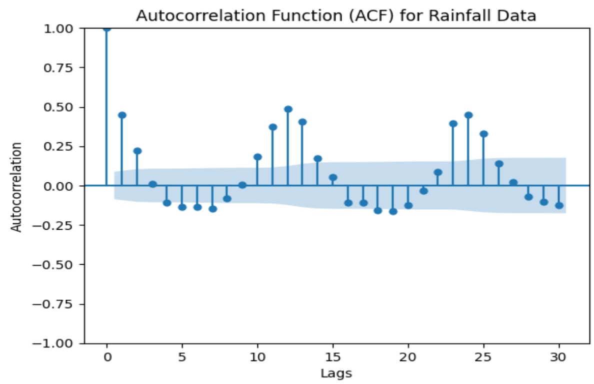


Fig 3: Rainfall series in Coimbatore district as depicted by a ACF plot

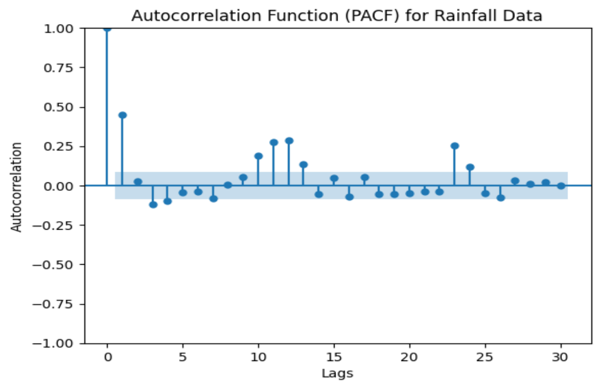


Fig 4: Rainfall series in Coimbatore district as depicted by a PACF plot

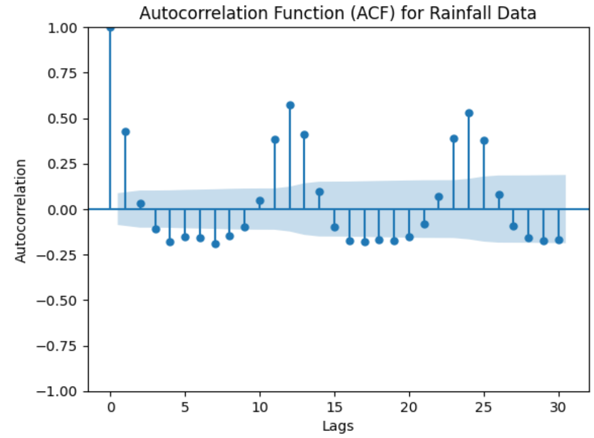


Fig 5: Rainfall series in Nagapattinam district as depicted by a ACF plot

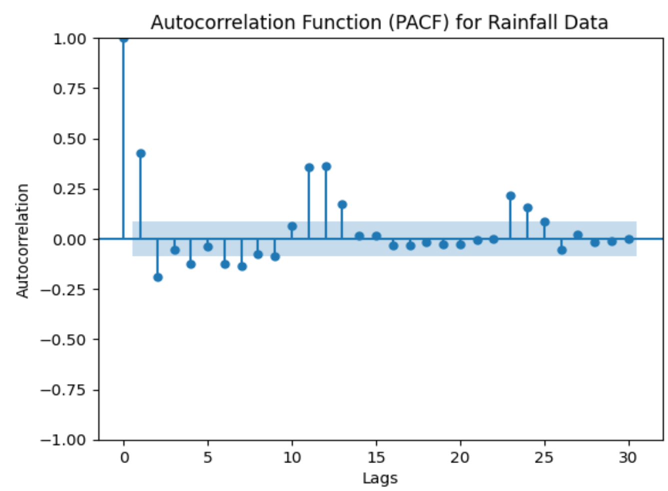


Fig 6 : Rainfall series in Nagapattinam district as depicted by a PACF plot

| Models | AIC | BIC |
|---|---|---|
| (1,0,0)(0,0,1)$_{12}$ | 5611 | 5624 |
| (1,0,0)(0,1,0)$_{12}$ | 5525 | 5534 |
| (1,0,0)(1,1,0)$_{12}$ | 5410 | 5422 |
| (2,0,1)(0,1,0)$_{12}$ | 5514 | 5531 |
| (2,0,0)(0,0,1)$_{12}$ | 5525 | 5538 |

Table 2: Various models for monthly rainfall in the district of Coimbatore

| Models | AIC | BIC |
|---|---|---|
| (1,0,0)(0,0,1)$_{12}$ | 6135.765180563572 | 6148.432909367786 |
| (1,0,0)(0,1,0)$_{12}$ | 6003.117757085063 | 6011.514714518048 |
| (1,0,0)(1,1,0)$_{12}$ | 5882.781342979649 | 5895.376779129126 |
| (2,0,1)(0,1,0)$_{12}$ | 6001.162096317492 | 6017.956011183461 |
| (2,0,0)(0,0,1)$_{12}$ | 5999.343528878705 | 6011.938965028182 |

Table 3: Various models for monthly rainfall in the district of Nagapattinam.

This model (1,0,0) (1,1,0) was selected after the model was found using the AIC and BIC criteria.

3.3 Parameter Estimation

We have determined that the data series adheres to the SARIMA model after obtaining data stationarity and analysing the ACF. Table 4 displays the model parameter estimation table needed to properly depict the monthly rainfall series.

| station | model | parameters | Estimate |
|---|---|---|---|
| Coimbatore | $(1,0,0)(1,1,0)_{12}$ | ar.L1 | 0.109464 |
| | | ar.S.L12 | -0.470825 |
| | | sigma2 | 3431.088116 |
| Nagapattinam | $(1,0,0)(1,1,0)_{12}$ | ar.L1 | 0.060574 |
| | | ar.S.L12 | -0.468291 |
| | | sigma2 | 8964.417996 |

Table 4: SARIMA$(1, 0, 0) (1,1,0)_{12}$ parameter estimation for monthly rainfall in the districts of Coimbatore and Nagapattinam.

The rainfall in the district of Coimbatore is predicted by using this equation to be

$$(1-\varphi_1 A)(1-\beta_1 A^{12})\ (1-\mathrm{A}^{12})Y_t = \varepsilon_t$$

$$(1-0.109464A)(1+(.470825)A^{12})\ (1-\mathrm{A}^{12})Y_t = 3431.088$$

The rainfall in the district of Nagapattinam is predicted by using this equation to be

$$(1-0.060574A)(1+(0.468291)A^{12})\ (1-\mathrm{A}^{12})Y_t = 8964.417996$$

Figures 7 and 8 depict the rainfall forecast for the districts of Coimbatore and Nagapattinam based on the aforementioned equation.

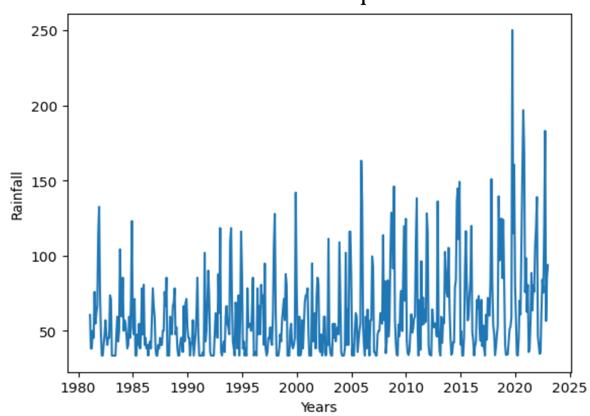


Fig 7: monthly precipitation forecast for the Coimbatore district

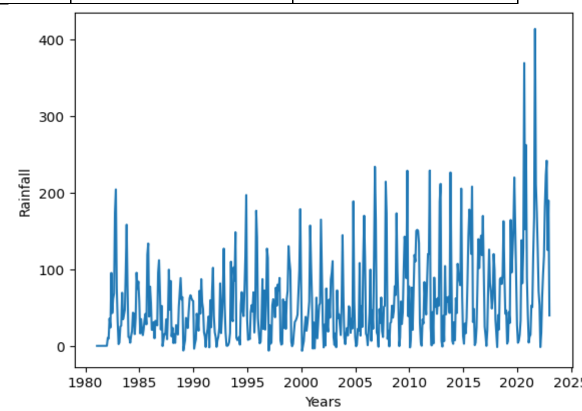


Fig 8: monthly precipitation forecast for the Nagapattinam district

| station | model | MSE | RMSE |
|---|---|---|---|
| Coimbatore | $(1,0,0)(1,1,0)_{12}$ | 0.513142227938152 | 0.07163394641775 |
| Nagapattinam | $(1,0,0)(1,1,0)_{12}$ | 1.3019071689289312 | 0.11410114674835355 |

Table5 : Model performance requirements

**Conclusion**

A powerful tool that is easy to use and very flexible is stochastic simulation. Rainfall is a significant problem for Tamil Nadu's eastern and western regions because it is essential for all life forms. Research on modeling and predicting rainfall is important because it will help policymakers, planners, and decision-makers

mitigate the negative impacts of natural disasters like floods and droughts. Information for this study was provided by the Tamil Nadu Meteorological Agency. Overall, the most rainfall was reported at the weather stations in Nagapattinam and Coimbatore.
The AIC, RMSE,and MSE model performance measures were used to compare the models, and the SARIMA model was employed to model the data set. The most accurate model was picked, and predictions were made with its help. This model had the lowest error. The weather station datasets in Tamil Nadu's Coimbatore and Nagapattinam were best fit by the SARIMA (1,0,0) (1,1,0) model. We used these adjusted models to forecast precipitation throughout the following years.

**References**

1. Akinbobola A, Okogbue EC, Ayansola AK (2018) Statistical modeling of monthly rainfall in selected stations in forest and savannah eco-climatic regions of Nigeria. J Climatol Weather For 2018(6), S1. https://doi.org/10.4172/2332-2594.1000226
2. Bjornlund V, Bjornlund H, Van Rooyen AF (2020) Why agricultural production in sub-Saharan Africa remains low compared to the rest of the world–a historical perspective. Int J Water Resour Dev 36(sup1):S20–S53. https://doi.org/10.1080/07900627.2020.1739512
3. .Box G, Jenkins G (1970) Time series analysis: forecasting and control
4. Darji M (2019) Rainfall forecasting using neural networks
5. Di C, Yang X, Wang X (2014) Hybrid neural network models for hydrologic time series forecasting. Plus One 9:e104663
6. Dwivedi DK, Kelaiya JH, Sharma GR (2019) Forecasting monthly rainfall using autoregressive integrated moving average model (ARIMA) and artificial neural network (ANN) model: a case study of Junagadh, Gujarat, India. J Appl Natl Sci 11(1), 35–41. https://doi.org/10.31018/jans. v11i1.1951
7. Grigonyte E, Butkeviciuye E (2016) Short-term wind speed foresting using ARIMA model. Energetika, 62
8. Ikot (2021) The environment and social and economic development nigerian aluminium smelter company. https://alscon.net/ikot-abasi-smelter/78-the-environment-and-socio-economic-dev
9. Ray S, Das SS, Mishra P, Al Khatib AMG (2021) Time series ARIMA modeling and forecasting of monthly rainfall and temperature in the south Asian countries. Earth Syst Environ 5(3):531–546.
10. Shad M, Sharma YD, Singh A (2022). Forecasting of monthly relative humidity in Delhi, Indian using ARIMA and ANN models. Modeling Earth Systems and Environment. https://doi.org/10. 1007/s.40808-022-01385-8
11. Box , G. E. P. and Jenkins, G. M., 1976. Time series analysisi, forecasting and control. San Francisco: Holden-Day
12. Zhang, Y., Bi, P., and Hiller, J., 2008. "Climate variations and salmonellosis transmission in Adelaide, South Australia: a comparison between regression models." International Journal of Bimeteorology, vol. 52, pp. 179-187.
13. Nirmala, M. and Sundaram, S. M., 2010. "A seasonal Arima Model for forecasting monthly rainfall in Tamilnadu." National Journal on Advances in Building Sciences and Mechanics, vol. 1, pp. 43-47.
14. Jiang, B., Liang, S., Wang, J., and Xiao, Z., 2010. "Modelling MODIS LAI time series using three statistical methods." Remote Sensing of Environment, vol. 114, pp. 1432-1444
15. Gikungu, S. W., Waititu, A. G., and Kihoro, J. M., 2015. "Forecasting inflation rate in Kenya using SARIMA model." American Journal of Theoritical and Applied Statistics, vol. 4, pp. 15 – 18.
16. Padhan, P. C., 2011. "Forecasting International tourists footfalls in India: An assortment of competing models." International Journal of Business and Management, vol. 6, pp. 190-202.
17. Mahsin, M., Akhter, Y., and Begum, M., 2012. "Modelling rainfall in Dhaka division of Bangladesh using time series analysis." Journal of Mathematical Modelling and Application, vol. 1, pp. 67-73.
18. Uchenna Amadi*, Love Cherukei Nnoka**, George Isobeye*** “An Application of Sarima Model on Monthly Naira-Dollar Exchange Rates Innocent”.
19. E. H. Chukwueloka and A. O. Nwos Modelling and Prediction of Rainfall in the North-Central Region of Nigeria Using ARIMA and NNETAR Model