Journal of Business Economics
ISSN 2682- 6933 | E-ISSN 2705-9944
2022 Volume 04 Issue 02:29-36

ANALYSIS OF MONTHLY RAINFALL IN AMPARA DISTRICT USING SEASONAL ARIMA MODEL

G.M.U.S. Karunarathna[1], A. Jahufer[2]

[1,2]Department of Mathematical Sciences, Faculty of Applied Sciences,

South Eastern University of Sri Lanka
[1]Usajini95@gmail.com, [2]jahufer@seu.ac.lk

ABSTRACT

Rainfall is one of the foremost important phenomena of the natural system. The purpose of this research study is to fit a time series model to analyze and forecast future rainfall in Ampara district. Accordingly, this paper presents seasonal autoregressive integrated moving average model to analyze the monthly rainfall data in Ampara district. For the study purpose, monthly rainfall data from January 1991 to December 2020 were used. Using the Auto Correlation Function and Partial Auto Correlation Function plots the seasonality was determined. The seasonal autoregressive integrated moving average model was chosen based on the Akaike Information Criterion, Schwarz Information Criterion, Hannan-Quinn Criterion, Log Likelihood and Adjusted R-squared values. Based on the results obtained for the seasonal autoregressive integrated moving average models, SARIMA$(0,0,2)(0,1,1)_{12}$ is selected as a best model for monthly rainfall for Ampara district.

Keywords: Rainfall, SARIMA, Ampara District, AIC, SIC, ACF.

1. Introduction

Ampara district plays an important role in the paddy production of Sri Lanka and majority of the people are cultivating paddy for their livelihood. This district produces about 20% of the national requirement of the paddy. In addition, other crops such as sugar cane, vegetables, chili and maize are also cultivated. Therefore, it is of great value to study the rainfall patterns in the Ampara District and to understand how to extend the cultivation period with the monthly rainfall.

Prediction of rainfall is tough because of its non-linear pattern and a large variation in intensity. Till today, varied ways that are used to predict rainfall (Mahmud, et. al, 2017). Among them, Autoregressive Integrated Moving Average (ARIMA) modeling, introduced by Box and Jenkins is an effective method (Mahmud, et. al., 2017).

The purpose of this article is to identify appropriate Box-Jenkins time series model for monthly rainfall in Ampara district, Sri Lanka. This paper is composed into five sections. Section two describes literature review; methodology is given in section three. In section four results and discussion are given. In the last section five conclusions are given.

2. Literature Review

Jayawardene, et, al. (2005) studied the trends of rainfall in Sri Lanka using over 100 years of data and found that A statistically significant increasing trend of rate 3.15 mm/year was observed at Colombo and decreasing trends were observed at Nuwara Eliya and Kandy with rates of 4.87 mm/year and 2.88 mm/year respectively. Since no coherent increase or decrease of rainfall in any group of stations in the wet or dry zones was observed, the possibility of large-scale change over the past century was ruled out.

Alahacoon and Edirisinghe (2021) studied spatial variability of rainfall trends in Sri Lanka from 1989 to 2019 as an indication of climate change. The most important thing reflected in this study is that there has been a significant increase in annual rainfall from 1989 to 2019 in all climatic zones (wet, dry, intermediate and Semi-arid) of Sri Lanka. The maximum increase is recorded in the wet zone and the minimum increase is in the semi-arid zone. There could be an increased risk of floods in the southern and western provinces in the future, whereas areas in the eastern and southeastern districts may face severe droughts during the northeastern monsoon. It is advisable to introduce effective drought and flood management and preparedness measures to reduce the respective hazard risk level.

Zahir (2013) studied the Rainfall Variability on Climate Changes in Eastern Province in Sri Lanka. The objective of this paper is to seek evidence spatiotemporal trends for rainfall variability on climate change in Eastern Province of Sri Lanka by analyzing long-term monthly data of rainfall received during the four rainy seasons, ie. the Northeast monsoon, the first inter-monsoon and the second inter-monsoon during the period 1980-2010, from meteorological stations of the Department of Meteorology. Five stations of Eastern Province have observed either flooding in rapid sequence in recent years.

Burt and Weerasinghe (2014) analyzed the rainfall distributions in Sri Lanka in time and space: an analysis based on daily rainfall data. Rainfall across Sri Lanka over three decades is investigated in relation to the main atmospheric drivers known to affect climate in the region: sea surface temperatures in the Pacific and Indian Oceans, of which the former are shown to be more important. The strong influence of El Niño and La Niña phases on various aspects of the daily rainfall distribution in Sri Lanka is confirmed: positive correlations with Pacific sea-surface temperatures during the north east monsoon and negative correlations at other times. It is emphasized in the discussion that Sri Lanka must be placed in its regional context and it is important to draw on regional-scale research across the Indian subcontinent and the Bay of Bengal.

Abeysekera et, al., (2015) studied the recent trends of extreme positive rainfall anomalies in the dry zone of Sri Lanka. A trend analysis was carried out, by considering the 95th and 99th percentile of daily time series of annual and seasonal rainfall viz, First Inter Monsoon (FIM), South West Monsoon (SWM), Second Inter Monsoon (SIM) and North East Monsoon (NEM) as the cut-off values to designate Heavy and Very Heavy rainfall events of positive anomalies, respectively. The base period to calculate the cut-off values was taken as the 30-year period from 1960-1989. Even though no significant trend in occurrence of Heavy Rainfall (HRF) events or Very Heavy Rainfall (VHRF) events has been observed during last 25-year period from 1990- 2014, annually or seasonally, it is revealed that an apparent trend of these extreme positive rainfall anomalies is discernible, especially during SWM and NEM rainfall seasons. This trend has been clearly evident during last 5-year period from 2010-2014. It may likely to inflict significant implications on the agricultural production in the region in terms of both quantity and quality, as SWM and NEM seasons coincide with the reproductive phase of the crops grown in both Yala and Maha cultivation seasons in the Dry zone. As the climate change is likely to exacerbate extreme rainfall events in future, these apparent trends may even become obvious thus, excess soil moisture stress in rainfed upland crops, flood damages in lowland paddy fields and rapid drying out in cascade of tanks in long-run will be a serious threat to agricultural productivity in the Dry zone of Sri Lanka.

3. Methodology

3.1 Study area and data collection

Ampara is located in the Eastern Province of Sri Lanka. It extends 7^0 17’ 51’’ north latitude and 81^0 40’ 55’’ east longitude. The average monthly rainfall data of Ampara district during the period 1991-2020 was used for this study and the data set contains 357 observations. Log transformation data were used for the analysis. The data set clearly shows that the highest density of the rainfall in period October to February and reaches it speaks in December.

3.2 Model building and selection procedures

Model building purpose the time series models Seasonal Autoregressive Integrated Moving Average (SARIMA) model is used. Model selection purpose, the minimum Akaike Information Criteria (AIC), Schwarz Information Criterion (SIC), Hannan-Quinn Criteria (HQC) and maximum Log-likelihood and R-square values are used.

3.3 Model diagnostic checking

The estimated model should be tested for suitability. This shall be done by some analysis of the residuals of the estimated model. The diagnostic check in which the residuals from the fitted model ought to be examined against adequacy which is done be correlation analysis along the residual ACF plots. If the residuals are correlated, then the model ought to be refined. Otherwise, the autocorrelations are white noise and therefore the model is adequate to represent our time series.

4. Results and Discussion

4.1 Preliminary analysis

The descriptive statistics value for the natural log transformation data is given in table 1.

Table 1: Descriptive statistics

Variable	Mean	SD	Minimum	Maximum
Values	4.473518	1.199733	-2.302585	6.738745

According to the Table 1 the minimum and the maximum rainfall received for the Ampara district were -2.302585 mm and 6.738745 mm respectively. The average rainfall received was 4.473518 mm and the standard deviation for the rainfall was 1.199733. Monthly wise the descriptive statistics values are given in Table 2.

Table 2: Monthly wise descriptive statistics values

Month	**Mean**	**SD**	**Minimum**	**Maximum**
Jan	5.102032	0.915112	2.197225	6.122559
Feb	4.556689	1.28605	-0.22314	5.784748
Mar	3.801764	1.528759	-2.30259	5.591733
Apr	4.122025	0.685397	1.902108	5.095038
May	3.902333	0.901425	1.740466	6.011267
Jun	3.174338	0.806636	1.163151	4.637637
Jul	3.639904	0.716429	2.066863	5.357529
Aug	4.1413	0.656465	1.94591	5.02913
Sep	4.313985	0.863379	1.62941	5.545568
Oct	5.149815	0.54849	4.225373	6.376931
Nov	5.912534	0.417167	5.083266	6.640268
Dec	5.834798	0.655592	3.440418	6.738745

According to the above Table 2 the minimum rainfall received for the Ampara district was in March and the maximum rainfall received for the Ampara district was in December.

The Logarithm transformation time series data plot was given in figure 1.

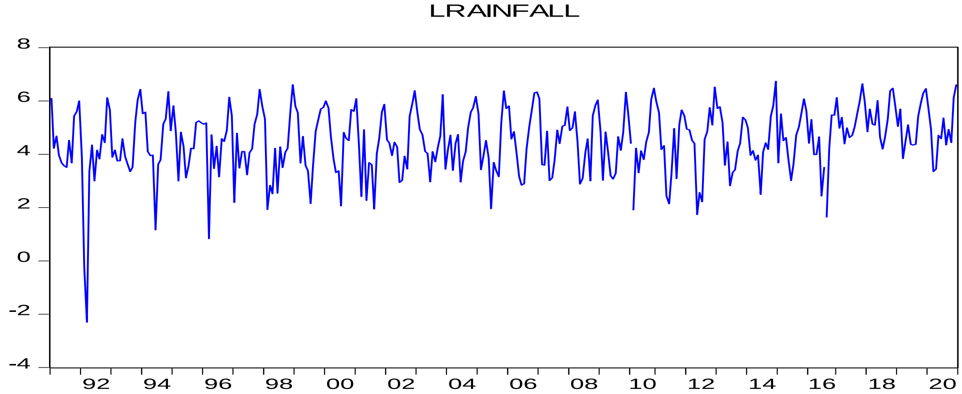

Figure 1: Time series plot of logarithm transform

From the above Figure 1 it is often seen that the data set is very fluctuating and there is no visible trend. The data patterns indicated that series is stationary and also the mean of the time series is relatively constant with respect to the time.

4.2 Unit root test

Stationary test was carried out using the unit root test. The test results are given below Table 3
Table 3: Unit root test results

Null Hypothesis: LRAINFALL has a unit root
Exogenous: Constant
Lag Length: 11 (Automatic - based on SIC, maxlag=16)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.351176	0.0134
Test critical values:	1% level	-3.449108	
	5% level	-2.869701	
	10% level	-2.571187	

From the above table 3 it can be concluded that natural log transformation data set is stationary at 5% significance level.
Moreover, it was confirmed by the ACF and PACF plots seasonal pattern in twelve months' period. Since data ought to be seasonal in twelve months' period of time. To remove the seasonal pattern first different of the data set was considered. First difference of the data plot is given in Figure 2.

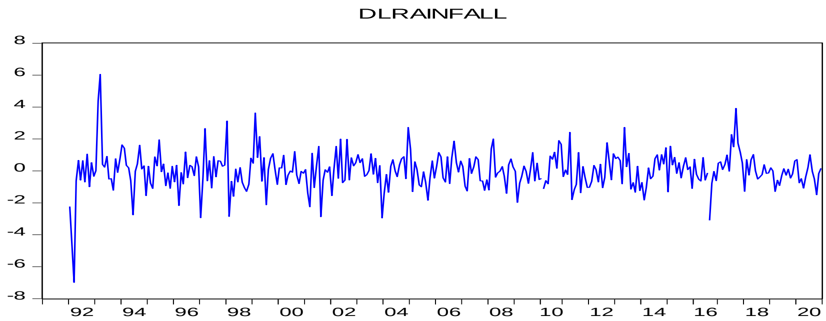

Figure 2: plot of the first difference data

From the above plot it can be said the seasonal pattern is not present for first difference data. Furthermore, the result is confirmed by the ACF and PACF plots respectively.

4.3 Model selection

Towel SARIMA models were selected based on their model selection statistics values and given in below Table 4. The best model is selected using the AIC, SIC, HQC, Adj. R^2 and Log Likelihood values.

Table 4: SARIMA models

Model	AIC	SIC	HQC	Adj. R^2	Log Likelihood
SARIMA(1,0,1)(0,1,1)$_{12}$	2.707494	2.763197	2.729677	0.404963	-462.0427
SARIMA(2,0,1)(1,1,1)$_{12}$	2.709061	2.775905	2.735681	0.405542	-461.3130
SARIMA(2,0,1)(0,1,1)$_{12}$	2.704084	2.770928	2.730705	0.408558	-460.4545
SARIMA(2,0,2)(0,1,1)$_{12}$	2.708612	2.786597	2.739669	0.407770	-460.2356
SARIMA(1,0,0)(0,1,1)$_{12}$	2.706898	2.751461	2.724645	0.403806	-462.9399
SARIMA(0,0,2)(0,1,1)$_{12}$	**2.697346**	**2.753050**	**2.719530**	**0.411139**	**-460.2922**
SARIMA(2,0,0)(0,1,1)$_{12}$	2.703321	2.759024	2.725504	0.407303	-461.3228
SARIMA(2,0,0)(3,1,0)$_{12}$	0.736266	2.814250	2.767323	0.386730	-465.0058
SARIMA(2,0,0)(2,1,0)$_{12}$	2.788150	2.854994	2.814771	0.347936	-474.9559
SARIMA(0,0,2)(1,0,0)$_{12}$	2.907102	2.962806	2.929286	0.256380	-496.4751
SARIMA(0,0,1)(3,1,0)$_{12}$	2.754493	2.821337	2.781113	0.373486	-469.1500
SARIMA(0,0,1)(2,1,0)$_{12}$	2.808074	2.863778	2.830258	0.332600	-479.3928

From Table 4, in terms of minimum AIC, SIC and HQC, the seasonal ARIMA (0,0,2)(0,1,1)$_{12}$ model is performed best for monthly rainfall in Ampara district.

4.4 Parameter Estimation

Table 5: Parameter estimation summary for SARIMA(0,0,2) (0,1,1)$_{12}$ model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.028042	0.020207	1.387764	0.1661
MA(1)	0.237845	0.046185	5.149856	0.0000
MA(2)	0.170554	0.054966	3.102918	0.0021
SMA(12)	-0.784264	0.038639	-20.29699	0.0000
SIGMASQ	0.816265	0.057331	14.23765	0.0000

The coefficient values, standard error, t-statistics and probability values of seasonal ARIMA (0,0,2)(0,1,1)$_{12}$ model reported in Table 5.

From the above Table 5, the P-values showed that every coefficient are extremely significant at 5% level of significance.

The plots of actual, fitted and residuals are shown in Figure 3. From these plots fitted values agree closely with the actual data.

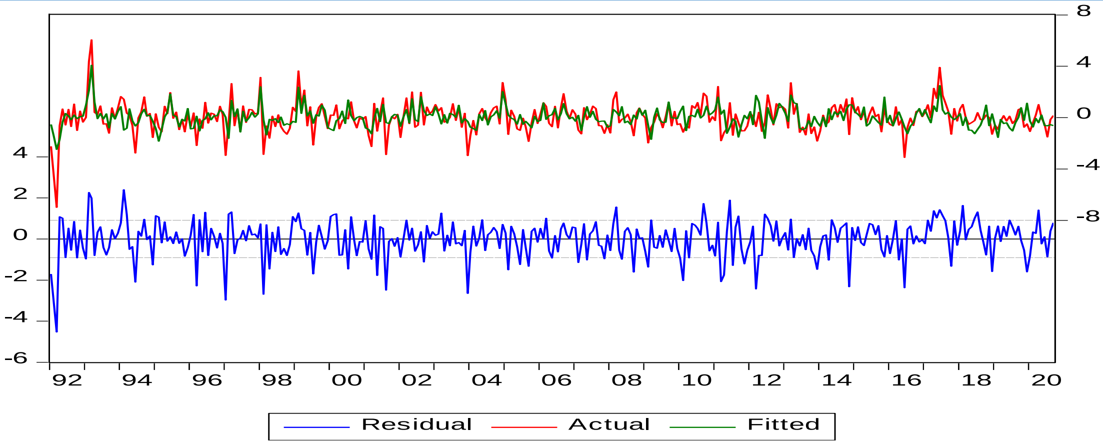

Figure 3: Actual, fitted and residual graph for the selected model

4.5 Model Diagnostics

The majority of the ACF and PACF values of residuals of the selected model lies within the confidence limit which indicates no significant correlation among them.

5. Conclusion

This study explored the presenting time series analyses for average monthly rainfalls in Ampara district from January 1991 to December 2020. The minimum AIC, SIC and HQ criterion results proposed SARIMA (0,0,2) $(0,1,1)_{12}$ is a best model. It may be used as the basis for forecasting, designing and management of the rainfall during this region.

References

1. Abeysekera, A.B., Punyawardena, B.V.R. and Premalal, K.H.M.S., (2015). Recent Trends of Extreme Positive Rainfall Anomalies in The Dry Zone of Sri Lanka. *Tropical Agriculture*, 163, pp. 1-108.
2. Alahacoon, N. and Edirisinghe, M., (2021). Spatial Variability of Rainfall Trends in Sri Lanka from 1989 to 2019 as an Indication of Climate Change. *International Journal of Geo-Information*. 10(84), pp. 2-18.
3. Burt, T.P. and Weerasinghe, K.D.N., (2014). Rainfall Distributions in Sri Lanka in Time and Space: An Analysis Based on Daily Rainfall Data. *Climate*, 2, pp. 242-263.
4. Jayawardene, H.K.W.I., Sonnadara, D.U.J. and Jayawardene, D.R., (2005). Trends of Rainfall in Sri Lanka over the Last Century. *Sri Lankan Journal of Physics*, 6, pp.7–17.
5. Mahmud, I., Bari, S.H., Rahman, M.T., (2017). Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method. *Environmental Engineering Research.* 22(2): 162-168.
6. Zahir, I.L.M., (2013). Rainfall Variability on Climate Changes in Eastern Province in Sri Lanka. *Proceeding of the Third International Symposium*. Proceedings of the Third International Symposium, SEUSL.