Modeling and Forecasting Rainfall in Catchment Area of Kotmale Reservoir, Sri Lanka

A.M.P.D. Abeykoon[*] and B.L. Peiris[1]

Postgraduate Institute of Agriculture
University of Peradeniya
Sri Lanka

ABSTRACT: The Mahaweli development programme contributes to 13% of the national hydroelectric power production and also the major water supply for Dry Zone of Sri Lanka. When considering Mahaweli cascade, the Kotmale reservoir affords a major contribution. The water level of the reservoir directly depends on rainfall pattern of the catchment area of the reservoir. Early prediction of rainfall pattern helps for efficient water management by predicting the water level of the reservoir. The time series analysis is a main statistical tool for predicting metrological phenomena. Therefore, in this study, the rainfall pattern of Kotmale station was modeled and forecasted using time series analysis with R software. The monthly rainfall data from 2003 to 2017 in the Kotmale station was considered. The study revealed that, SARIMA $(0, 1, 1) (2, 1, 2)_{12}$ as the best fitted model for forecasting the rainfall pattern at the Kotmale station. Using the best fitted model, the rainfall pattern in the Kotmale station was forecasted.

Keywords: Rainfall, Kotmale reservoir, time series, SARIMA and forecasting

INTRODUCTION

The Mahaweli river is the longest river which consists of four cascades Kothmale, Randenigala, Rantembe and Victoria. The main objectives of the Mahaweli development programme are to provide water to dry zone of the country and generating hydroelectric power (Lareef, 2003). Water supply for a major part of agriculture fields in Sri Lanka depends on the Mahaweli system. Thirteen percent of the electricity requirement is fulfilled by hydropower in Sri Lanka. One third of hydropower depends on the Mahaweli system. The Kothmale reservoir has a major contribution for the Mahaweli development programme. Climatic changes happened over many years in the recent past. The rainfall pattern will vary in the future. The water levels of reservoirs depend on rainfall of the reservoir's catchment areas. Changes in rainfall pattern directly affect the agriculture production and hydroelectric power generation. Prediction of the rainfall pattern leads to prediction of the water levels of the reservoir. The predicted water levels of reservoir can be used for the efficient water management of agriculture and hydropower.

Time series analysis is the main tool for analyzing and predicting metrological phenomena such as rainfall, temperature and humidity (Wei, 2005). In this research, the rainfall pattern of the Kothmale reservoir's catchment was analyzed. Nine stations over the Kotmale

---

[1] Department of Crop Science, Faculty of Agriculture, University of Peradeniya, Sri Lanka
[*] Corresponding author: prabhadinithi@gmail.com

catchment (544 km^2) collect daily rainfall data. The data were analyzed for finding the best time series model for rainfall pattern in the Kothmale catchment and the model was used for forecasting the rainfall for future months.

METHODOLOGY

Monthly rainfall data of the Kothmale catchment from 2003 to 2017 was used for analysis. R software was used for analysis and graphics (Cowpertwait & Metcalfe, 2009; Shumway & Stoffer, 2010). The data analysis was done in three steps (Bisgaard & Kulahci, 2011).

1) Identify whether the rainfall series is stationary or not. If it is not stationary, transform into a stationary series if possible.
2) Identify the best fitting model and evaluate the goodness of fit of the model.
3) Forecast the future rainfall using the best fitting model.

RESULTS AND DISSCUSION

The time series plot of monthly average rainfall in millimeters of the Kothmale catchment from 2003 to 2017 is given in Figure 1.
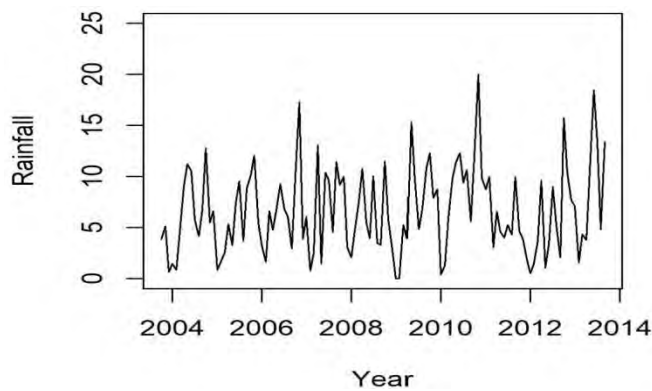

Figure 1. The time series plot of the average monthly rainfall in mm of the Kothmale catchment from 2003 to 2017

The plots of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) are given in Figure 2.
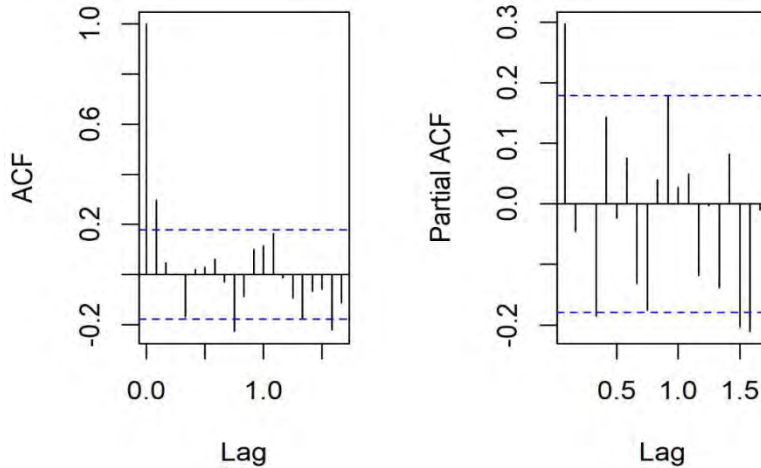
Figure 2. The plot of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the monthly rainfall data of the Kothmale catchment from 2003 to 2017

The original rainfall series was decomposed to trend, seasonal and random components using the additive model. In the decomposed rainfall series, there is a seasonal pattern in monthly rainfall. The series was adjusted for the seasonal variation. The stationary condition is necessarily required to fit most of the time series models. The series was adjusted for seasonality. However, the seasonally adjusted series was not stationary. Using a proper degree of differencing a non-stationary time series may be reduced to a stationary time series (Montgomery *et al*., 2008). By a unit root test the stationary condition can be checked (Dickey *et al*., 2008). Therefore, difference series of adjacent series was obtained. Then the plots of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) were used to identify a suitable autoregressive integrated moving average (ARIMA) model. The plot of the seasonally adjusted differencing rainfall series is given in Figure 3.
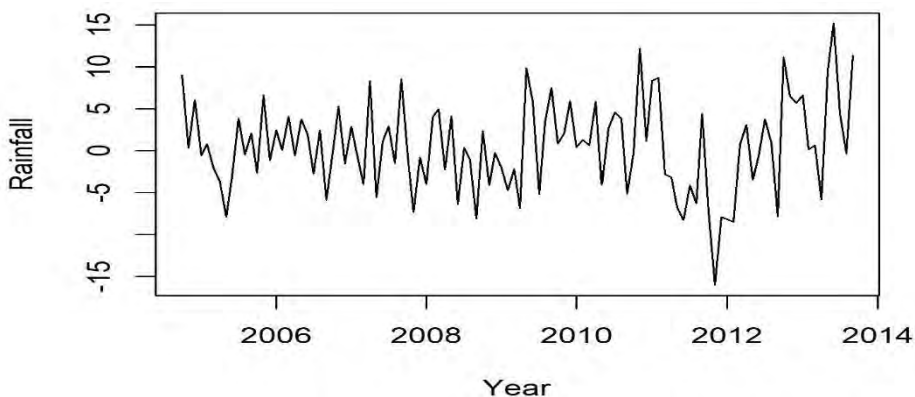

Figure 3. The plot of the seasonally adjusted differencing rainfall series

The plot of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for the seasonally adjusted differencing rainfall series is given in Figure 4.
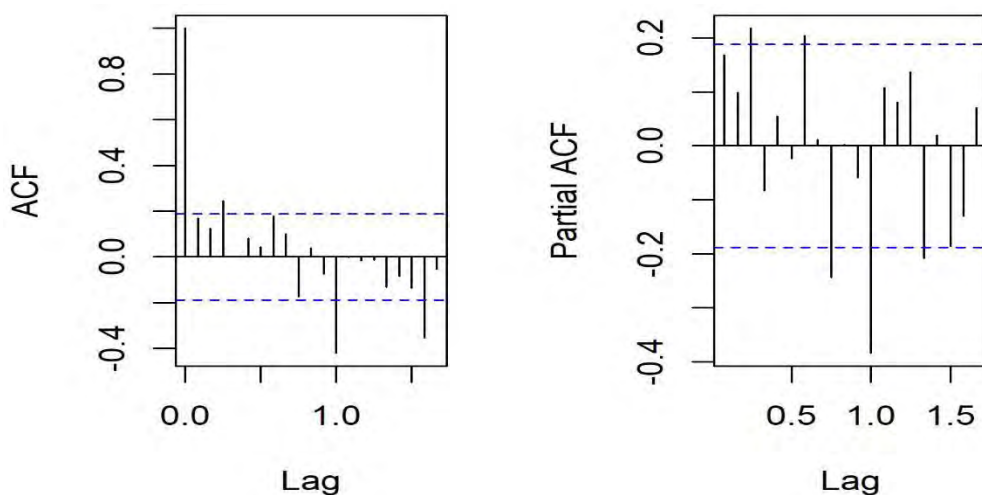
Figure. 4. The plot of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for the seasonally adjusted differencing rainfall series

According to the correlogram, the autocorrelation (ACF) at the lag 1 exceeds the significance bounds and also it was negative. However, the other autocorrelations between the lags 1 to 20 do not exceed the significance bounds. The partial correlogram shows that the partial autocorrelations (PACF) at the lag 1 and 10 exceed the significance bounds and PACF is slowly decreasing in magnitude with increasing lag. According to ACF and PACF plots ARIMA (0, 1, 1) was identified as a candidate best fitting model. Further, the data was fitted to several ARIMA models (Table 1). ARIMA (0, 1, 1) had the minimum AIC (Akaike Information Criterion) value.

Table 1. Autoregressive integrated moving average (ARIMA) models and their corresponding AIC (Akaike Information Criterion) values

ARIMA	AIC
(1, 0, 0)	895.34
(0, 0, 1)	895.55
(1, 0, 1)	895.56
(1, 1, 0)	933.67
(0, 1, 1)	895.10
(1, 1, 1)	895.88

Several seasonal autoregressive integrated moving average (SARIMA) models were introduced for ARIMA (0,1,1) Model. According to the AIC values of SARIMA (0, 1, 1) (2, 1, 2)$_{12}$ was identified as the best fitting model (Table 2).

Table 2. **Seasonal autoregressive integrated moving average (SARIMA) models and their corresponding AIC values**

SARIMA	AIC
$(0,1,1)(1, 0, 0)_{12}$	954.22
$(0,1,1) (0, 0, 1)_{12}$	955.20
$(0,1,1)(1, 0, 1)_{12}$	930.11
$(0,1,1)(1, 1, 0)_{12}$	927.12
$(0,1,1)(0, 1, 1)_{12}$	876.67
$(0,1,1)(1, 1, 1)_{12}$	874.66
$(0,1,1)(2, 0, 0)_{12}$	951.87
$(0,1,1)(0, 0, 2)_{12}$	954.76
$(0,1,1)(2, 0, 2)_{12}$	923.34
$(0,1,1)(2, 2, 0)_{12}$	956.62
$(0,1,1)(0, 2, 2)_{12}$	885.80
$(0,1,1)(2, 2, 2)_{12}$	885.98
$(0,1,1)(1, 2, 2)_{12}$	886.59
$(0,1,1)(2, 2, 1)_{12}$	905.97
$(0,1,1)(1, 2, 1)_{12}$	916.53
$(0,1,1)(1, 1, 2)_{12}$	868.12
$(0,1,1)(2, 1, 1)_{12}$	872.48
$(0,1,1)(2, 1, 2)_{12}$	867.71
$(0,1,1)(0, 1, 2)_{12}$	872.47
$(0,1,1)(1, 0, 2)_{12}$	926.27
$(0,1,1)(2, 0, 1)_{12}$	928.36
$(0,1,1)(2, 1, 0)_{12}$	909.46
$(0,1,1)(0, 2, 1)_{12}$	948.08
$(0,1,1)(1, 2, 0)_{12}$	987.92

The parameter estimates, their standard errors, 95% confidence interval for the parameters, estimated residual variance, log likelihood and AIC for the best fitting model SARIMA $(0, 1, 1) (2, 1, 2)_{12}$ are given in Table 3.

Table 3. **The parameter estimates, their standard errors, 95% confidence interval for the parameters, estimated residual variance, log likelihood and AIC for the best fitting model SARIMA $(0, 1, 1) (2, 1, 2)_{12}$**

	ma1	sar1	sar2	sma1	sma2
Coefficients	-0.9323	0.4747	-0.1826	-1.8401	0.9993
SE	0.0435	0.1008	0.1107	0.6421	0.6958
2.5 %	-1.0174	0.2771	-0.3996	-3.0987	-0.3644
97.5 %	-0.8471	0.6724	0.0344	-0.5816	2.3630

$\hat{\sigma}^2$ = 11.01: Log Likelihood = -427.86: AIC = 867.71

The plot of autocorrelation function for the estimated residuals of the best fitting model SARIMA $(0, 1, 1) (2, 1, 2)_{12}$ is given in Figure 5.
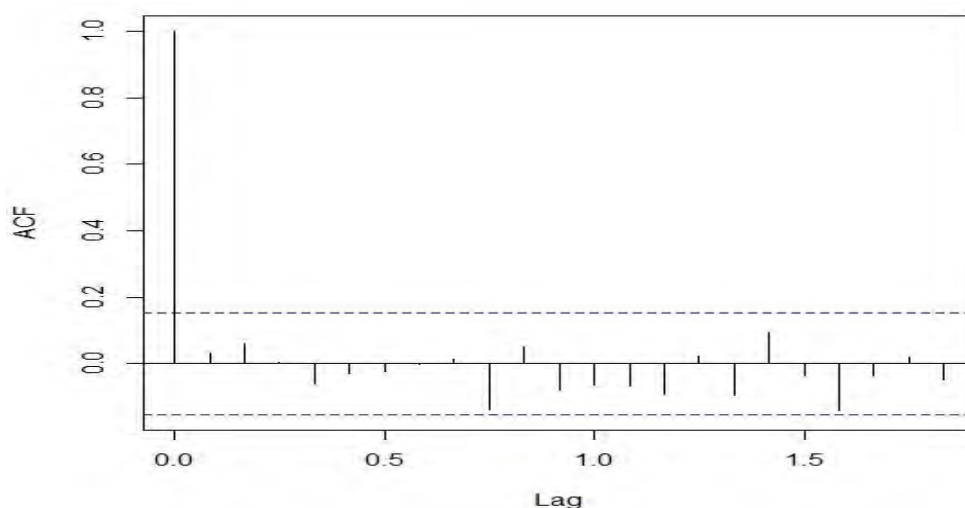
Figure 5. The plot of autocorrelation function for the residuals of the best fitting model

Box-Ljung test (Box *et al*., 2008) was performed on residuals from SARIMA models (0, 1, 1) $(2, 1, 2)_{12}$. The Ljung-Box test statistic was 17.4437 (df = 20) and the *P*-value was 0.624. This indicates that there was no evidence of autocorrelations of residuals of the SARIMA model (0, 1, 1) $(2, 1, 2)_{12}$. Therefore, the SARIMA model (0, 1, 1) $(2, 1, 2)_{12}$ was identified as the best fitting model. Using the SARIMA model (0, 1, 1) $(2, 1, 2)_{12}$ rainfall pattern for 2017 July to 2017 December was forecasted (Table 4).

Table 4. Forecasted rainfall for 2017 July to 2017 December with the SARIMA model (0, 1, 1) $(2, 1, 2)_{12}$

Month	Predicted Rainfall (mm)	Standard Error (SE)
July	8.53	3.56
August	6.38	3.57
September	7.85	3.58
October	9.56	3.56
November	7.88	3.57
December	6.22	3.58

CONCLUSIONS

SARIMA (0, 1, 1) $(2, 1, 2)_{12}$ was identified as the best fitting time series model for forecasting the rainfall pattern in Kothmale catchment area. According to that model, the rainfall pattern from 2017 July to 2017 December was forecasted. The water level of Kothmale reservoir can be predicted using the predicted rainfall. The predicted water level of the reservoir can be used for managing water efficiently for future agriculture production and hydropower generation.

REFERENCES

Bisgaard, S. and Kulahci, M. (2011). Time Series Analysis and Forecasting by Example. John Wiley & Sons, Hoboken, New Jersey, United States.

Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015). Time Series Analysis: Forecasting and Control. John Wiley & Sons, Hoboken, New Jersey, United States.

Cowpertwait, P.S. and Metcalfe, A.V. (2009). Time Series Data. In Introductory Time Series with R. Springer, New York, United States.

Dickey, D., Hasza, D. and Fuller, W. (1984). Testing for unit roots in seasonal time series. J. Am. Stat. Assoc. *79*, 355-367.

Lareef, Z. (2003). El Nino-southern oscillation influences on the Mahaweli stream flow in ri   an a Int. J. Climatol. *23*, 91-102.

Montgomery, D.C., Jennings, C.L. and Kulahci, M. (2015). Introduction to Time Series Analysis and Forecasting. John Wiley & Sons, Hoboken, New Jersey, United States.

Shumway, R.H, and D. Stoffer. (2010). Time Series Analysis and its Applications: with R examples, Springer, New York, United States.

Shumway, R.H., & Stoffer, D.S. (2011). Time Series Regression and Exploratory Data Analysis. In Time Series Analysis and its Applications. Springer, New York, United States.

Wei, W.W.S. (2005). Time Series Analysis - univariate and Multivariate Methods, Pearson, London, United Kingdom.