

MEASURING FLOOD RISK IN RATNAPURA TOWN AREA IN SRI LANKA

H. W. Y. J. Hettiwaththa^{1*} and R. A. B. Abeygunawardana¹

¹Department of Statistics, University of Colombo, Sri Lanka

Abstract: Flood is a common chaotic natural problem frequently occurring in Ratnapura district. Recent chain of flood events that occurred in Ratnapura district has raised the question regarding the capability of defending civilian lives and property from this natural disaster. Ratnapura town area mainly face flood risks due to the over flow of Kalu river during the South West monsoon season. Flood risk cannot be directly measured. But flood risk is directly proportional to the river water level. When river water level is increasing, flood risk also increases. Since river water level can be taken as a measurement for flood risk, this analysis is based on forecasting river water level of Kalu River. Data were collected from meteorological stations at Ratnapura, Galabada, Guruluwana and Lellopitiya which are located in upper catchment area of Kalu River. ARIMAX-GARCH model was fitted to forecast water level using the rainfall data and the water discharge rate. The accuracy of the fitted model to forecast water level was high when comparing the estimated values with the actual values. Therefore the ARIMAX-GARCH model can be used to measure the flood risk in Ratnapura town area.

Keywords: Flood; Forecasting; Risk; ARIMAX-GARCH

Introduction

Flood is a natural event where usually a dry land area suddenly sinks under water. Some flooding can arise suddenly and discharge rapidly. Some flooding take days or months to build up and unload. In Sri Lanka, it was reported that there were major floods in many years with incisive loss of life, property damage and environmental damages. There are 103 major river basins and 17 of them are in flood prone areas (Ratnayake et al., 2010). Among them Kalu river basin, Kelani river basin, Gin river basin and Nilwala river basin are four main river basins which have high flood risks. Flood events in Sri Lanka between 1990 and 2011 are given in the following figure,

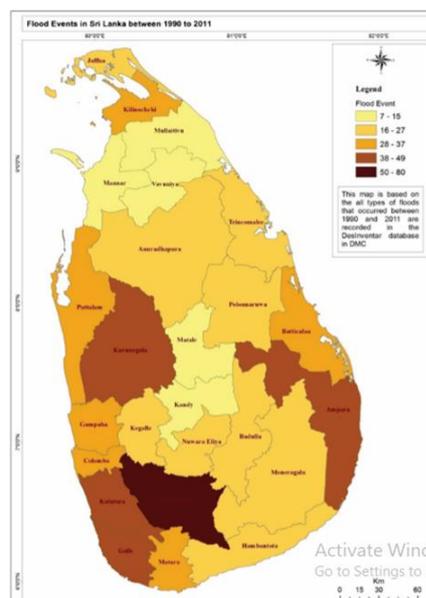


Figure 1 Flood events in Sri Lanka

According to Figure 1 Ratnapura district faced the largest number of flood events during last 22 years. Ratnapura district faced more than 50 flood events in that time period. This map was derived from the data available on all types of floods at Disaster Management Centre.

Floods are one of the major natural problems occurring in Ratnapura District. Ratnapura town is situated on the right bank of the Kalu River. Damage caused by flooding is almost yearly occurrence in between the main rainy seasons in the Kalu river Basin. In some years when floods are of large magnitude, the damages are excessive and widespread. This entails heavy social and economic costs. In some occasions loss of human lives had been reported in Ratnapura district. If there is an early warning to people, most of the damages due to flood can be minimized. Forecasting flood is very important in most of the applications in disaster management, agriculture and ecology.

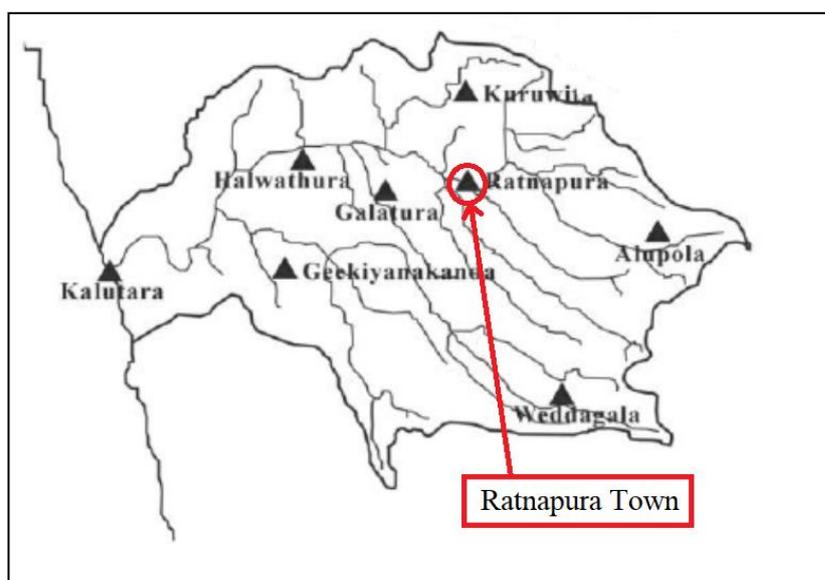


Figure 2 Location of Kalu River basin

In the Figure 2, the black line mark the Kalu River basin and the situation of the Ratnapura town is marked by the red circle.

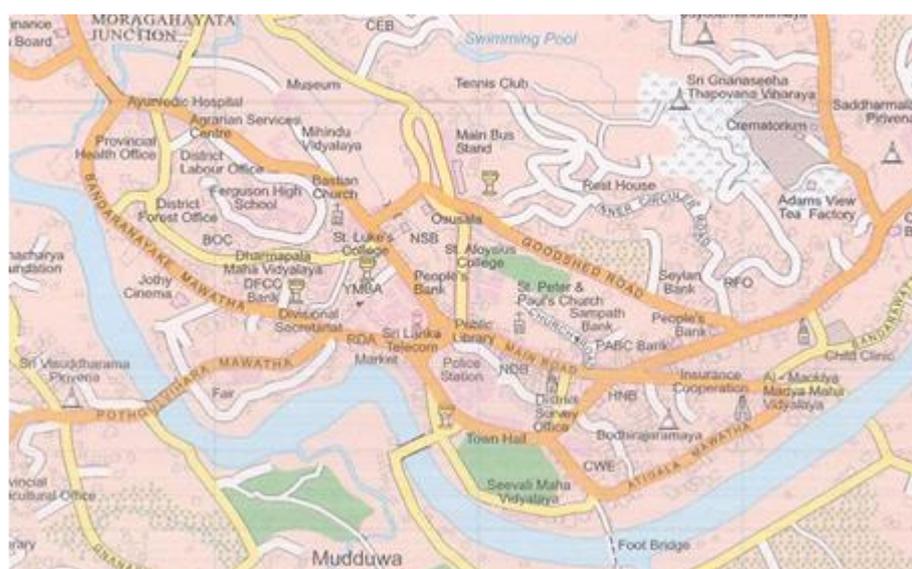


Figure 3 Location of the Ratnapura city on the right bank of the Kalu River

Figure 3 is an enhanced map depicting the situation of Ratnapura town area. The map illustrates how Kalu River flows through Ratnapura town area. As shown in the map, during a flood event, Ratnapura town area is at immediate risk of flooding. The very high annual rainfall falling into the catchment area of 604 km² above Ratnapura is the main cause of flooding in Ratnapura town. Floods occur during the South West monsoon period from May to September and during the inter-monsoonal period from September to October. Since Ratnapura town is situated on the right bank of the Kalu River, flooding of Ratnapura town had been a frequent occurrence. It is reported that “Approximately one-half of the population lives on the flood plain, and the majority of flood plain residents are concentrated in the most hazardous areas” (Karunanayake & Katupotha, 1990). Ratnapura town is at a flood risk when the river water level rises to 18.3 m (61.0 ft.) MSL (Mean Sea Level). When considering the previously occurred floods, Department of Irrigation has classified four flood levels at Ratnapura town based on the water level of Kalu River as shown in the Table 1.

Table 1 Classification of flood levels

| Classification | Water Level (m MSL) |
|----------------|---------------------|
| Normal | 18.3 |
| Minor | 20.1 |
| Major | 21.3 |
| Critical | 24.4 |

Source: Department Of Irrigation

Flood risks cannot be completely eliminated or controlled. Flood forecasting models can be used to issue timely flood warnings prior incident to the residents in the flood prone area to minimize flood damages and risk.

Materials and Methods

Daily rainfall data at four gauging stations which are located in the upper catchment area of Kalu River were collected from Department of Meteorology over a period of ten years (2007-2016). The selected meteorological stations are Ratnapura, Galabada, Guruluwana and Lellopitiya. Daily water level and daily discharge rate of Kalu River at Ratnapura water level gauging station were collected from Department of Irrigation from years 2007 to 2016. Daily rainfall data and river discharge rate at Ratnapura from January 2007 to April 2016 were taken as the exogenous variables to develop a time series model for river water level at Ratnapura. Regression model and ARIMA model were combined to get regression with ARIMA errors. The models proposed in this study were able to handle volatility associated with the data and the models work well with skewness and volatility clustering.

ARIMAX model

The ARIMAX model is an extended version of the ARIMA model which contain the exogenous variables in the model. Forecasting performance of the model can be improved by adding other exogenous variables into ARIMA model. This model is also called as dynamic regression model. For time series data Y_t and explanatory data X_t , the ARIMAX(p,d,q) model can be defined as,

$$\nabla^d Y_t = \sum_{i=1}^p \phi_i \nabla^d Y_{t-i} + \sum_{j=1}^q \theta_j e_{t-j} + \sum_{m=1}^M \beta_m X_{m,t} + e_t \quad [1]$$

Where ∇ is non-seasonal differencing operator, p is the order of the autoregressive part, d is the order of the differencing, q is the order of the moving average process, M is the number of explanatory variables, ϕ_i , θ_j and β_m are model parameters and e_t is a white noise series. It is important in modelling the ARIMAX model, that all

variables (Y_t & X_t) in the model should be stationary. So if there are any non-stationary series, use difference operator to make the series stationary.

ARCH/GARCH models

ARIMA model does not capture the heteroscedasticity (violation of constant variance) of a time series. Heteroscedasticity of a time series can be observed in a form of leverage effect, clustering of volatilities and high kurtosis. When such problems arise, ARCH/GARCH models can be used to model time series. The variance equation of the GARCH (p,q) model can be written as,

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad [2]$$

Where ω is intercept and $\omega > 0$, ε_t is the error terms and $\varepsilon_t = \sigma_t z_t$. The random variable z_t is a white noise process. α_i and β_j are model parameters and $\alpha_i \geq 0$ for $i = 1, \dots, p$ and $\beta_j \geq 0$ for $j = 1, \dots, q$

ARIMAX-GARCH models

ARMAX model with GARCH error term is called as ARIMAX (p,d,q) – GARCH (p,q) model. The procedure of the ARIMAX - GARCH modelling methodology is shown in figure 4.

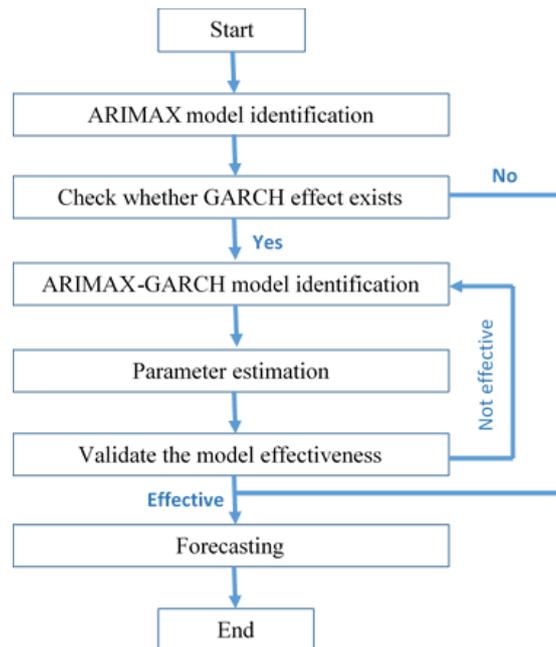


Figure 4 Procedure of ARIMAX-GARCH forecasting

Akaike's Information Criterion (AIC)

The Akaike's information criterion is an estimator commonly used in model selecting. It estimate the goodness of fit of the fitted model. When there are collection of models for particular dataset, AIC compares the quality of each model relative to other models. The model which have the minimum AIC value, is selected as the best fitted model. Although AIC value gives the best model, it does not say anything about the quality of the model.

$$AIC = -2 * \ln(\text{likelihood}) + 2 * K \tag{3}$$

Where $K = p + q + P + Q$ (number of terms estimated in the model)

Ljung-Box test

Ljung-Box test called as Q-test was introduced by Ljung and Box (1978) which can be used in ARIMA modelling. It is applied to the residuals of the fitted model to check whether the residuals have significant autocorrelations. The hypothesis of the Ljung-Box test is,

H0 : There is no serial correlation

H1 : There is a serial correlation

The test statistic is,

$$Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k} \tag{4}$$

Where n is the sample size, $\hat{\rho}_k$ is the sample autocorrelation at lag k , and h is the number of lags being tested.

Results

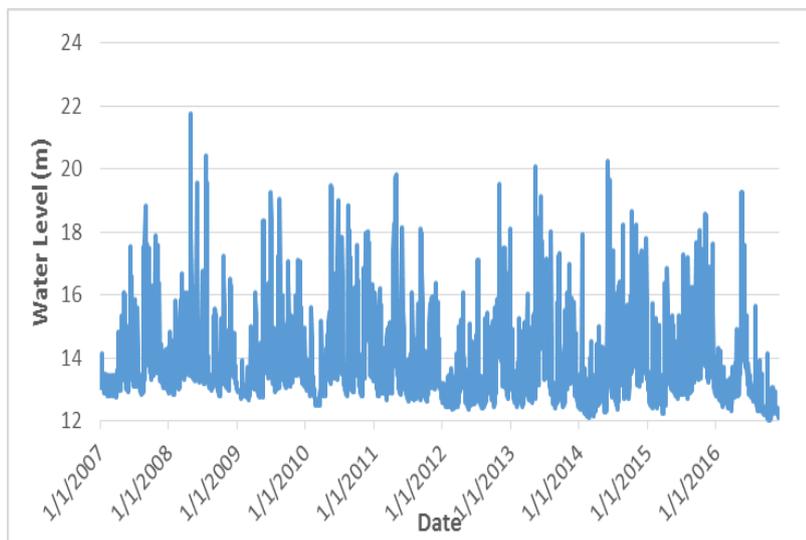


Figure 5 Maximum daily water level at Ratnapura

Analysis of daily river water level

Figure 5 is a time series plot of daily maximum water level of Kalu River. It shows that there are occasions where water level of Kalu river exceeding 18.3 m which is a normal flood level in every year (from 2007-2016). Therefore in every year, there is a flood risk in Ratnapura town area. Vertical axis of above figure start from 12 as the minimum water level of Kalu River is 12 m (MSL).

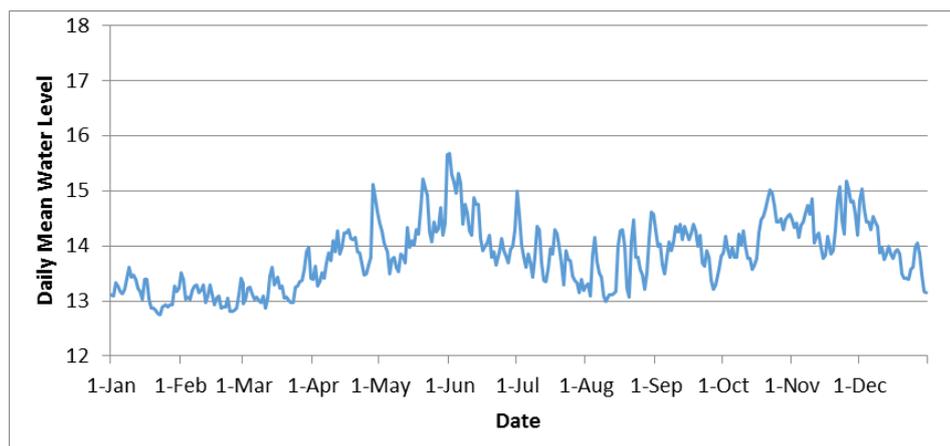


Figure 6 Daily mean water level over ten year period of Kalu River at Ratnapura

Figure 6 shows the daily mean water level of Kalu River at Ratnapura. Since high daily mean water level was recorded from May to July and September to December, there is a flood risk during those period.

Forecasting the river water level to measure the flood risk

Rainfall data at four gauging stations and river discharge rate at Ratnapura were taken as the exogenous variables to develop a time series model for river water level at Ratnapura. When estimating the regression with ARIMA errors, all the variables in the model must be stationary. Therefore the first difference of all variables in the model were taken to make the series stationary. The most suitable ARIMAX model was identified based on the Akaike's information criterion (AIC) values to find the best model to forecast river water level.

Table 2 AIC and BIC values of the fitted ARIMAX models

| | Model | AIC value | BIC value |
|---|----------------|-----------|-----------|
| 1 | ARIMAX (0,1,4) | 4661.58 | 4722.91 |
| 2 | ARIMAX (0,1,5) | 4663.37 | 4730.84 |
| 3 | ARIMAX (1,1,1) | 4657.03 | 4706.10 |
| 4 | ARIMAX (1,1,4) | 4634.51 | 4701.98 |
| 5 | ARIMAX (2,1,3) | 4661.37 | 4722.01 |

Since ARIMAX (1,1,4) has the lowest AIC value and the lowest Bayesian information criterion (BIC) value, model 4 was selected as the best model and significance of the parameters and residual diagnostics were checked. Autocorrelation function (ACF) and p-values for Ljung-Box statistic suggested the adequacy of the selected model. However, significant autocorrelation coefficients can be seen in ACF of squared residuals plot (Figure 7) and p-values for Ljung-Box statistics of squared residuals are less than 0.05 after the 7th lag (Figure 8). It illustrates that there still remains serial correlation in the series.

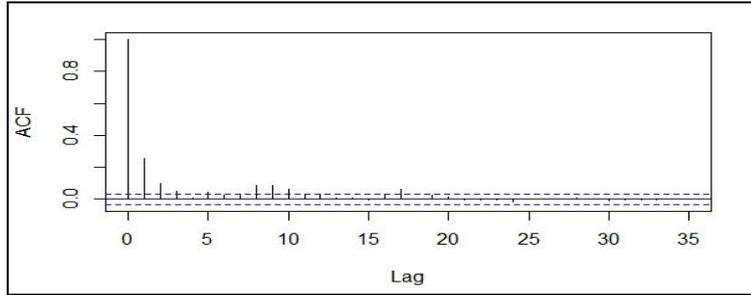


Figure 7 ACF of squared residuals for ARIMAX (1,1,4) model

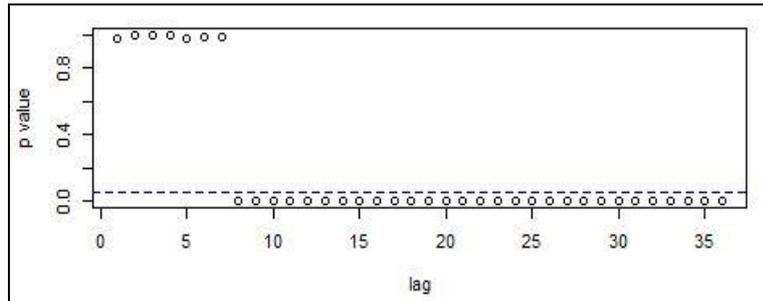


Figure 8 P values for Ljung-Box statistic of squared residuals

Therefore ARCH/GARCH should be used to model the volatility of the series. By adding and removing all the terms of the model, the best ARIMAX-GARCH model for the series was identified using AIC values.

Table 3 AIC values of the fitted ARIMAX-GARCH models

| Model | AIC value |
|--------------------------------|-----------|
| 1 ARIMAX (1,1,4) – GARCH (1,1) | 1.1857 |
| 2 ARIMAX (1,1,4) – GARCH (1,2) | 1.1769 |
| 3 ARIMAX (1,1,4) – GARCH (2,0) | 1.2091 |
| 4 ARIMAX (2,1,3) – GARCH (1,1) | 1.1835 |
| 5 ARIMAX (2,1,3) – GARCH (1,2) | 1.1744 |
| 6 ARIMAX (2,1,3) – GARCH (2,1) | 1.1849 |
| 7 ARIMAX (1,1,3) – GARCH (1,2) | 1.1776 |

According to the Table 3, ARIMAX (2,1,3) – GARCH (1,2) model has the minimum AIC value. Hence, model 5 was selected as the best model and residual analysis was carried out to check the adequacy of the model.

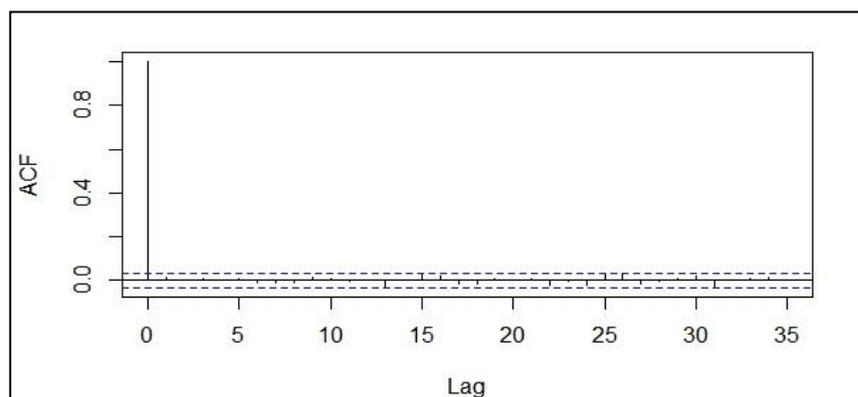


Figure 9 ACF of standardized residuals for ARIMAX (2,1,3) – GARCH (1,2)

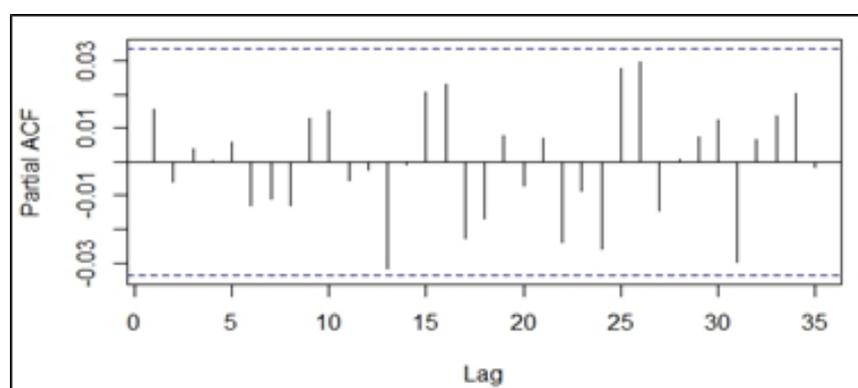


Figure 10 PACF of standardized residuals for ARIMAX (2,1,3) – GARCH (1,2)

ACF plot of standardized residuals (Figure 9) & partial autocorrelation function (PACF) plot of standardized residuals (Figure 10) clearly shows the white noise behavior. There are not any significant lags in the ACF and PACF plots. It confirmed that the selected model is the most appropriate model. As another method to check the adequacy of the model, Ljung-Box statistic was applied and got following results.

Table 4 Q statistics on standardized residuals

| | Statistic | p-value |
|--------|-----------|---------|
| Q (1) | 0.8259 | 0.3635 |
| Q (14) | 3.0079 | 1.0000 |
| Q (24) | 7.6626 | 0.9804 |

Table 5 Q statistics on squared standardized residuals

| | Statistic | p-value |
|--------|-----------|---------|
| Q (1) | 0.005494 | 0.9409 |
| Q (8) | 3.674631 | 0.5629 |
| Q (14) | 6.410823 | 0.5835 |

Since p-values are greater than 0.05, it can be concluded that the mean model and the variance model are correctly specified. Hence, the selected model is the most appropriate model for forecast water level of Kalu River.

Therefore, the most suitable model is,

$$Y_t = 2.42885Y_{t-1} - 1.878754Y_{t-2} + 0.449904Y_{t-3} + e_t - 2.048355e_{t-1} + 1.183733e_{t-2} - 0.134966e_{t-3} + 0.004583Rat + 0.004788Gal + 0.002118Gur + 0.004575Lel + 0.02653Dis$$

$$\sigma_t^2 = 0.034750 + 0.476115e_{t-1}^2 + 0.150403\sigma_{t-1}^2 + 0.339053\sigma_{t-2}^2$$

Model Forecasting

River water level was forecasted for the month May, 2016 using the selected model. Forecasted values were compared with the actual values to check the accuracy of the model.

Table 6 Accuracy of the selected model for river water level forecasting

| Date | Forecasted Value | Actual Value | Error | Whether flooding or not | |
|-----------|------------------|--------------|-------|-------------------------|--------|
| | | | | Forecast | Actual |
| 5/1/2016 | 12.81 | 12.86 | 0.05 | No | No |
| 5/2/2016 | 12.87 | 13.06 | 0.19 | No | No |
| 5/3/2016 | 13.08 | 13.00 | -0.08 | No | No |
| 5/4/2016 | 12.99 | 12.83 | -0.16 | No | No |
| 5/5/2016 | 12.79 | 12.92 | 0.13 | No | No |
| 5/6/2016 | 12.94 | 14.39 | 1.45 | No | No |
| 5/7/2016 | 14.52 | 14.34 | -0.18 | No | No |
| 5/8/2016 | 13.85 | 13.49 | -0.36 | No | No |
| 5/9/2016 | 13.29 | 14.59 | 1.30 | No | No |
| 5/10/2016 | 14.78 | 14.62 | -0.16 | No | No |
| 5/11/2016 | 14.79 | 13.81 | -0.98 | No | No |
| 5/12/2016 | 13.67 | 13.94 | 0.27 | No | No |
| 5/13/2016 | 13.98 | 14.45 | 0.47 | No | No |
| 5/14/2016 | 14.60 | 14.44 | -0.16 | No | No |
| 5/15/2016 | 14.45 | 17.29 | 2.84 | No | No |
| 5/16/2016 | 17.48 | 18.74 | 1.26 | No | Yes |
| 5/17/2016 | 18.80 | 19.25 | 0.45 | Yes | Yes |
| 5/18/2016 | 18.95 | 19.24 | 0.29 | Yes | Yes |
| 5/19/2016 | 19.07 | 18.12 | -0.95 | Yes | No |
| 5/20/2016 | 17.93 | 18.76 | 0.83 | No | Yes |
| 5/21/2016 | 18.84 | 17.70 | -1.14 | Yes | No |
| 5/22/2016 | 17.70 | 16.41 | -1.29 | No | No |
| 5/23/2016 | 16.25 | 14.82 | -1.43 | No | No |
| 5/24/2016 | 14.85 | 14.24 | -0.61 | No | No |

| | | | | | |
|-----------|--------|-------|-------|----|----|
| 5/25/2016 | 14.38 | 16.24 | 1.86 | No | No |
| 5/26/2016 | 16.51 | 15.58 | -0.93 | No | No |
| 5/27/2016 | 15.58 | 15.93 | 0.35 | No | No |
| 5/28/2016 | 15.79 | 17.34 | 1.55 | No | No |
| 5/29/2016 | 17.55 | 15.87 | -1.68 | No | No |
| 5/30/2016 | 15.68 | 15.89 | 0.21 | No | No |
| 5/31/2016 | 15.77 | 17.58 | 1.81 | No | No |
| <hr/> | | | | | |
| MAE | 0.8193 | | | | |
| RMSE | 1.0662 | | | | |
| MAPE | 0.0508 | | | | |

Table 6 illustrates that there are no high differences between the actual values and the forecasted values. MSE, RMSE and MAPE for the model are minimum. These factors emphasize that the forecasted values are very close to the actual values. Therefore, fitted model is the most suitable model to forecast the water level of Kalu River at Ratnapura and the forecasted river water level can be taken as the measurement of future flood risks.

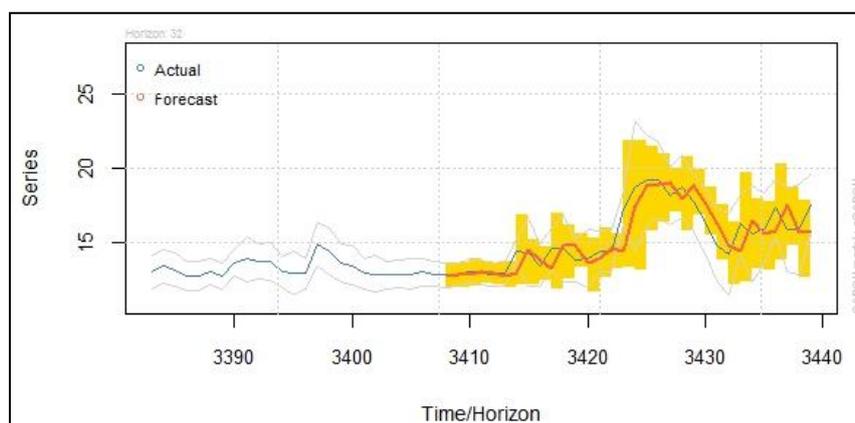


Figure 11 Forecast vs. actual series

Discussion and Conclusion

The location of Ratnapura town in the Kalu River basin expose the area to flood risk. Ratnapura city faces flood risk annually specially during May to July and September to December. As the flood risk is associated with the river water level, flood risk can be classified based on the river water level. According to the classification of flood level from the Department of Irrigation, when river water level is 18.3 m (MSL) the flood risk is normal and when river water level rises to 21.3 m (MSL) there is a very high risk of flood in Ratnapura town area. Since river water level is the most suitable measurement for flood risk, forecasting river water level allows to measure future flood risks. ARIMAX – GARCH model is the most adequate model to forecast river water level accurately in Ratnapura town area. Since river water level can be forecasted accurately using this method, this model can be used to measure flood risk in Ratnapura town area.

References

- Bolshakov, V. (2013). Regression-based Daugava River Flood Forecasting and Monitoring. *Information Technology and Management Science*, 137-142. doi:10.2478/itms-2013-0021
- Campolo, M., Soldati, A., & Andreussi, P. (2003). Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences–Journal–des Sciences Hydrologiques*, 48(3), 381-398.
- De Silva, M. M., Weerakoon, S. B., Herath, S., & Ratnayake, U. (2012). Flood Inundation Mapping along the Lower Reach of Kelani River Basin under the Impact of Climatic Change. *Engineer - Journal of the Institution of Engineers, Sri Lanka*, 45(2), 23-29.
- Karunanayake, M. M., & Katupotha, J. (1990). An environmental profile of Ratnapura district. Coliombo 10: Central environmental authority.
- Ljung, G. M., & Box, G. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297-303.
- Nandalal, K. D. (2009). Use of a hydrodynamic model to forecast Floods of Kalu River in Sri Lanka. *Journal of Flood Risk Management*, 151-158. doi:10.1111/j.1753-318X.2009.01032.x
- Ratnayake, U., Sachindra, D. A., & Nandalal, K. D. (2010). Rainfall forecasting for flood prediction in the Nilwala basin. *Proceedings of the International Conference on Sustainable Built Environment (ICSBE-2010)*, pp. 355-362.
- Shamseldin, A. Y., Abdo, G. M., & Elzein, A. S. (1999). Real-Time Flood Forecasting on the Blue Nile River. *Real-Time Flood Forecasting on the Blue Nile River*, 24(1), 39-45. doi:10.1080/02508069908692132
- Wang, W., Gelder, V. P., Vrijling, J. K., & Ma, J. (2005). Testing and modelling autoregressive conditional heteroskedasticity of streamflow processes. *Nonlinear Processes in Geophysics*, 55-66.
- Yaziz, S. R., Azizan, N. A., Zakaria, R., & Ahmad, M. H. (2013). The performance of hybrid ARIMA-GARCH modeling in forecasting gold price. *20th International Congress on Modelling and Simulation, Adelaide, Australia*, 1201-1207.
- Zhao, J. H., Dong, Z. Y., & Zhao, M. L. (2009). A statistical model for flood forecasting. *Australasian Journal of Water Resources*, 13(1), 43-52.