A NON-LINEAR STATISTICAL MODEL FOR KALU RIVER DAILY STREAM FLOW PREDICTION

H.A.N.S. Hewavitharana[1], K.Perera[2] and K.D.W. Nandalal[3]

[1]*Postgraduate Institute of Science, University of Peradeniya*
[2] *Department of Engineering Mathematics,* [3]*Department of Civil Engineering, University of Peradeniya*

Introduction

Hydrological modeling for the prediction of daily stream flow is a major task in water resources related studies, such as water resources assessments, water resources system management and operation, flood prediction and mitigation, river and reservoir water quality management, etc. Various attempts have been made in the past towards this goal using diverse techniques. The work reported in this paper is an attempt made in the development of a statistics based model to predict daily stream flow of *Kalu* River at *Ratnapura*. It is a non-linear regression model based on Marquardt's procedure, which predicts daily stream flow (m^3/s) based on previous stream flow (m^3/s) in the river and average rainfall (mm) in the catchment (Amisigo *et al*, 2008). The model was validated using Nash-Sutcliffe efficiency (Krause *et al.*, 2005), Root mean square error (RMSE) and pseudo R^2.

Methodology

Flow in the *Kalu* River upper catchment at *Ratnapura* and rainfall from eight rainfall stations above *Ratnapura* were used for the research. Thiessen polygon method was carried out to calculate average rainfall over the catchment. Marquardt's procedure in non-linear regression is used to fit the model. Overall significance of the fitted model was established using the analysis of variance technique, Nash-Sutcliffe efficiency, RMSE and the Pseudo R^2. The significance of model parameters were tested using confidence intervals constructed for the parameters. Shapiro-Wilk test, White's General test and Durbin Watson test were used to test for normality, heteroskedastisity and serial correlation of the residuals respectively. Several models were tested and the above model was observed to be the most suitable one.

Data Analysis

The null hypotheses that the variables are non stationary (excitant of unit root) was carried out using the Augmented Dicky Fuller test and it shows that both variables are stationary (p-value < 0.01). By observing the Box-Plot of the daily stream flow for each year and by considering the outliers and identified the minimum value of the outlier as 100 m^3/sec. Analysis was carried out for the daily stream flow data by removing the outliers. Daily stream flow and daily rainfall for the period 1987-1994 were used for the calibration of the model while data for the period 1995-1997 were used for the validation of the model.

Results and Discussion

Since there is seasonal variation, cosine and sine terms with time trend were used to catch the seasonality of the data set. Also, lag 1 and lag 2 average daily rainfalls (r_{t-1} and r_{t-2}) and lag 1 daily stream flow (s_{t-1}) were used as predictor variables and daily stream flow S_t as the response variable. The developed model is given below.

$$St = 0.028 * \cos(3*\pi*t/950) + 0.028*\sin(2*\pi*t/3) + 0.916*S(t-1) + 0.007*r(t-1) - 0.002*r(t-2) \quad (1)$$

According to the analysis of variance, since the p-value <0.05, it can be concluded that the model is significant. The predicted and the observed daily stream flow of the *Kalu* River at *Ratnapura* for the calibration period 1987 to 1994 are given in Figure 1.

Further, the Pseudo R^2=0.83, the Nash-Sutcliffe efficiency = 0.71 and the RMSE= 14.44, confirms the validity of the model. Table 1 presents the significance of parameters. In table 2 residuals are normally distributed (P-values >0.05), constant variance (P-values >0.05) and the Durbin Watson test statistic closer to 2 confirms that the underline assumptions of the residuals are satisfied by the fitted model.

Figure 2 shows the plot of daily stream flow for the validation period 1995 to 1997.

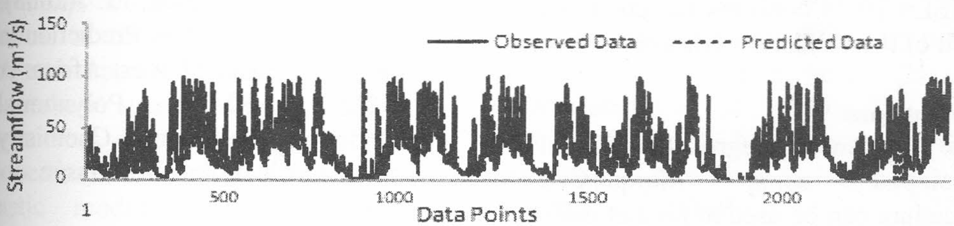

Fig. 1. The plot of daily stream flow for the calibration period

Table 1. Significance of the parameters

Parameter Estimates	Approximate Standard error	95% Confidence Interval
0.028	0.005	(0.013 0.034)
0.028	0.005	(0.013,0.034)
0.916	0.005	(0.902 0.942)
0.007	0.001	(0.006 0.008)
-0.002	0.001	(-0.002,-0.001)

Table 2. Residual Analysis

Hypothesis on residuals	Test	Test statistics	P-value
Ho: Normally distributed	Shapiro-Wilk Test	0.9878	0.496
Ho: Constant variance	White's Test	2.5	0.457
Ho: No serial correlation	Durbin Watson Test Statistic	1.95	-

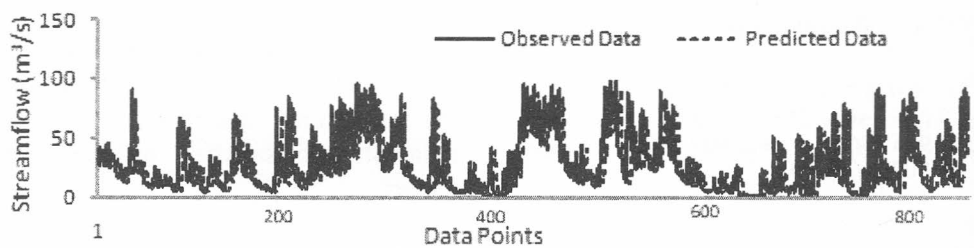

Fig. 2. The plot of daily stream flow for the validation period

The Pseudo R^2=0.85, the Nash-Sutcliffe efficiency = 0.95 and the RMSE= 10.74 confirms the goodness of fit of the validation data set.

Conclusions

The non-linear regression model developed using Marquardt's procedure can be used to predict daily stream flow of the *Kalu* River upper catchment in operational level.

References

Amisigo, B.A., Rogers, C., Andah, W.E.I. and Friesen, J. (2008). Monthly Stream flow Prediction in the Volta Basin of West Africa: A SISO NARMAX Polynomial Modeling, Physics and Chemistry of the Earth, Volume-33, Issue 1-2, 141-140.

Krause, P., Boyle, D.P. and Base, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. Advances in Geosciences, Vol-5, 89-97.