Annual Academic Sessions 2011
ISSN 2012-9912; ©2011 Open University of Sri Lanka

A STATISTICAL APPROACH TO FIND THE FACTORS AFFECTING THE DAILY STREAM FLOW OF *KALU GANGA* AT *ELLAGAWA* AND THE DELAY EFFECT OF RAINFALL IN THE *ELLAGAWA* CATCHMENT ON THE DOWNSTREAM FLOW OF THE RIVER

K. D. V. F. Siriwardana*

Department of Mathematics and Computer Science, Open University of Sri Lanka

INTRODUCTION

Kalu Ganga is the second largest river in Sri Lanka in terms of annual volume of runoff to the sea. Being situated entirely in the wet zone of Sri Lanka, it has a high rainfall to runoff response and the water is discharged as floods. The floods in *Kalu Ganga* are a regular feature. Floods cause less damage to plantations in the upper reaches, but cause more damage in lower reaches below *Ellagawa*. Studying the flow at *Ellagawa* will be useful for the disaster mitigation programmes for pre-preparation. Objective of this study was to find the factors affecting the daily flows at *Ellagawa* and find the delay effect of rainfall in the *Ellagawa* catchment on the downstream flow of the river. A statistical approach was used in this study.

METHODOLOGY

Over the period of January 2001 to December 2006 a random sample of 100 days were selected and the following data were collected on these days. Daily average discharge in cubic meters per second at the *Ellagawa* gauging station was obtained from the Hydrology Division of Irrigation Department. Daily rainfall in mm in the selected eleven rainfall stations within the catchment area (*Galathura, Wellandura, Eheliyagoda, Keragala, Balangoda, Alupola, Hapugastenna, Rathnapura, Landsdown, Halwathura* and *Deepedena*), Maximum daily temperature in Celsius and Relative Humidity measured daily at 8.30a.m. at the *Rathnapura* Meteorology station were obtained from the Department of Meteorology.

Regression analysis was used to find the significant factors affecting the daily stream flow at *Ellagawa.* Natural logarithm of Daily flows at *Ellagawa* in cubic meters per second (m^3/s) was selected as the response variable. On the day, previous day, two days before and three days before measurements of Average Rainfall, Temperature, Humidity and previous day, two days before and three days before measurements of natural logarithm of Flow were selected as the predictor variables. To satisfy the model assumptions in Regression analysis, Natural logarithm of the Flow was used as the response variable in model building instead of the Flow. *Thiessen Polygon* Technique$_4$ was used to find the Average rainfall of the study location. Out of many predictor variables, the inclusions of the selected variables were decided using partial F-test in forward selection procedure, backward elimination procedure and best subsets. Analysis of variance was used to check the statistical significance of the selected model. R^2 – Coefficient of determination was used to check the proportion of variation explained by the model. Hypothesis testing (T tests) was used to check the statistical significance of the contribution of variables. Residual analysis was used to check the model assumptions and whether the variables were specified correctly (Plots of residuals against the fitted values and the Normal probability plots were used). Variance inflation factor (VIF) was used to detect the multicollinearity in the predictor variables.

* All correspondence should be addressed to K. D. V. F. Siriwardana, Department of Mathematics and Computer Science, Open University of Sri Lanka (email: kdsir@ou.ac.lk)

RESULTS AND DISCUSSION

Out of the selected random sample of 100 days, 99 days were used in model selecting procedure since one observation was found with missing Rainfall data. The selected sample was well distributed over the years and months within the sample population.

Following formula was found from the Thiessen Polygon Technique$_4$ to calculate the Average rainfall of the study location.

Average Rainfall = (439 Daily Rainfall of Alupola +250 Daily Rainfall of Balangoda +
152 Daily Rainfall of Deepadena +232 Daily Rainfall of Eheliyagoda
+642 Daily Rainfall of Galathura +201 Daily Rainfall of Halwathura
+534 Daily Rainfall of Hapugastenna +675 Daily Rainfall of Kera+
941 Daily Rainfall of Wellandura +679 Daily Rainfall of Ratnapura +
696 Daily Rainfall of Landsdown)/5441

Result obtained from the Best subsets technique is given in Table-1, the result obtained from Stepwise regression analysis procedure is given in Table-2 and the selected model is given in Table-3.

The selected model was obtained by after careful analysis of Best subsets technique and the Stepwise(forward selection and backward elimination at 5% level of significance). multicollinearity did not exist in the predictor variables (VIF < 3.5). Selected model is significant at 5% level of significance and 90% of the total variation is explained by the fitted model. Flow of the previos day (p=0.00), Rainfall of the two days before (p=0.001), Relative Humidity of the day (p=0.009), previous day (p=0.034) and two days before (p=0.002) were selected as the most important factors affecting the daily stream flow at *Ellagawa* station among the predictor variables considered in this study. Rainfall of the day and previous day were not selected as significant factors for the flow at the *Ellagawa* station when the selected variables are present in the selected model. This result led to the conclusion that Runoff at the *Ellagawa* station will be increased after two days on average after rainfall in upper catchment. Therefore, there is almost of two days delay on average effect of rainfall in the *Ellagawa* catchment on the downstream flow of the river.

Best Subsets Regression: Ln(Flow) versus Ln(Flow)-pre, Ln(Flow)-pre, ...

Response is Ln(Flow)

99 cases used, 1 cases contain missing values

Vars	R-Sq	R-Sq(adj)	Mallows C-p	S	Ln(Flow)-pre1	Ln(Flow)-pre2	Ln(Flow)-pre3	Rainfall Day	Rainfall-pre1	Rainfall-pre2	Rainfall-pre3	Humidity(Day)	Humi-pre1	Humi-pre2	Humi-pre3	Temp	Temp-pre1	Temp-pre2	Temp-pre3
1	86.1	85.9	42.1	0.38950	X														
1	70.6	70.3	193.5	0.56502		X													
2	87.8	87.5	27.2	0.36657	X							X							
2	87.2	87.0	32.6	0.37472	X								X						
3	88.8	88.4	19.4	0.35318	X					X		X							
3	88.7	88.3	20.4	0.35483	X						X	X							
4	89.4	88.9	15.3	0.34517	X					X		X		X					
4	89.3	88.8	16.3	0.34684	X	X					X	X							
5	89.9	89.3	12.4	0.33870	X					X		X	X	X					
5	89.7	89.2	14.0	0.34138	X	X				X		X		X					
6	90.2	89.5	11.7	0.33593	X	X	X			X		X		X					
6	90.1	89.5	12.2	0.33675	X	X				X		X	X	X					
7	90.5	89.8	9.9	0.33106	X	X	X			X		X	X	X					
7	90.4	89.7	11.0	0.33304	X	X	X				X	X	X	X					
8	90.7	89.9	10.4	0.33018	X	X	X			X	X	X	X	X					
8	90.7	89.9	10.5	0.33036	X	X	X			X		X	X	X		X			
9	90.9	90.0	10.0	0.32760	X	X	X				X	X	X	X				X	X
9	90.9	90.0	10.3	0.32825	X	X	X			X		X	X	X		X	X		
10	91.1	90.1	10.7	0.32703	X	X	X			X		X	X	X		X	X		X
10	91.1	90.1	10.8	0.32718	X	X	X				X	X	X	X		X		X	X
11	91.3	90.2	10.4	0.32454	X	X	X			X	X	X	X	X		X	X		X
11	91.2	90.1	11.1	0.32592	X	X	X				X	X	X	X		X	X	X	X
12	91.4	90.2	11.4	0.32457	X	X	X			X	X	X	X	X	X	X	X		X
12	91.4	90.2	11.7	0.32509	X	X	X			X	X	X	X	X		X	X	X	X
13	91.5	90.2	12.5	0.32476	X	X	X			X	X	X	X	X	X	X	X	X	X
13	91.4	90.1	13.1	0.32587	X	X	X		X		X	X	X	X	X	X	X	X	X
14	91.5	90.1	14.1	0.32580	X	X	X		X	X	X	X	X	X	X	X	X	X	X
14	91.5	90.1	14.5	0.32656	X	X	X	X		X	X	X	X	X	X	X	X	X	X
15	91.6	90.0	16.0	0.32763	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table-1: Best subsets

Stepwise Regression: Ln(Flow) versus Ln(Flow)-pre1, Ln(Flow)-pre2, ...

```
  Alpha-to-Enter: 0.05  Alpha-to-Remove: 0.05


Response is Ln(Flow) on 15 predictors, with N = 99
N(cases with missing observations) = 1 N(all cases) = 100
```

Step	1	2	3	4	5
Constant	0.2061	-0.8731	-0.6653	-0.1824	-0.2791
Ln(Flow)-pre1	0.957	0.923	0.832	0.871	0.864
T-Value	24.46	24.29	17.22	17.37	17.52
P-Value	0.000	0.000	0.000	0.000	0.000
Humidity(Day)		0.0154	0.0156	0.0198	0.0137
T-Value		3.68	3.87	4.57	2.67
P-Value		0.000	0.000	0.000	0.009
Rainfall-pre2			0.0108	0.0131	0.0132
T-Value			2.90	3.49	3.57
P-Value			0.005	0.001	0.001
Humi-pre2				-0.0126	-0.0202
T-Value				-2.34	-3.17
P-Value				0.022	0.002
Humi-pre1					0.0152
T-Value					2.15
P-Value					0.034
S	0.389	0.367	0.353	0.345	0.339
R-Sq	86.05	87.77	88.77	89.38	89.89
R-Sq(adj)	85.91	87.52	88.41	88.93	89.34
Mallows C-p	42.1	27.2	19.4	15.3	12.4

Table-2 : Stepwise regression analysis

Regression Analysis: Ln(Flow) versus Ln(Flow)-pre1, Humidity(Day), ...

```
The regression equation is
Ln(Flow) = - 0.289 + 0.869 Ln(Flow)-pre1 + 0.0145 Humidity(Day)
           + 0.0133 Rainfall-pre2 - 0.0205 Humi-pre2 + 0.0147 Humi-pre1
```

Predictor	Coef	SE Coef	T	P
Constant	-0.2892	0.3727	-0.78	0.440
Ln(Flow)-pre1	0.86852	0.04915	17.67	0.000
Humidity(Day)	0.014465	0.005068	2.85	0.005
Rainfall-pre2	0.013274	0.003697	3.59	0.001
Humi-pre2	-0.020477	0.006346	-3.23	0.002
Humi-pre1	0.014719	0.007062	2.08	0.040

S = 0.338784   R-Sq = 89.9%   R-Sq(adj) = 89.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	96.559	19.312	168.26	0.000
Residual Error	94	10.789	0.115		
Total	99	107.347			

Table-3 Selected model

CONCLUSIONS/RECOMMENDATIONS

Prediction of *Kalu Ganga* floods is essential as the damages it causes are high. These damages are being affecting to the economy of the country as well as the smooth lifestyle of the people living close to *Kalu Ganga*. According to this study, by observing out the Rainfall within the *Ellagawa* catchment area and the Relative humidity at *Rathnapura* meteorology station regularly, prediction of floods below the *Ellagawa* can be done average of two days earlier. Floods will not come bellow *Ellagawa* before the second day or after the fourth day with respect to rainfall in the upper catchment. Therefore flood alert would be applicable below the *Ellagawa* is on the second day to the fourth day after rainfall in the upper catchment.

In this study, delay effect was given in days since daily data was used. Hourly data are not available for the selected predictor variables which were considered in the study. If hourly data have been available, using the technique used in the study, the delay effect can be found in hours. Further more models for the stream flows of rivers can be built separately by using both fields of Statistics and Hydrology. New topic on hybrid from the two fields for better prediction capabilities is a distinct possibility.

REFERENCES

Department of Engineering Hydrology, University College Galway, Ireland,1995, Hand Book, International Advance Course/Workshop on River Flow Forecasting.

De Silva M.A.P.(2006),A time series model to predict the runoff ratio of catchment of the Kalu Ganga basin, Natn. Sci. Foundation Sri Lanka,volume 34(2),103-105

Dharmasena G.T.(2003),Investigations of Trends in Hydrological Time Series, Sri Lanka Water Heritage, History of Water conservation, volume 2,143-151.

Hydrology for Engineers(1975), Ray K. Linsley J.R., Max A. Kohler,Joshep L.H. Paulhus, McGRAW-HILL KOGAKUSHA, LTD