A REGRESSION MODEL FOR THE STREAM FLOW AT CALIDONIA IN UPPER KOTMALE CATCHMENT

A.M.C.C. PREMATILAKE*, K. PERERA* AND S.B. WEERAKOON**

**Department of Engineering Mathematics and **Department of Civil Engineering, Faculty of Engineering, University of Peradeniya*

Upper Kotmale Catchment(UKC), which covers an area of 301km^2 in the central hills of Sri Lanka and having an annual rainfall of about 2270mm, is the uppermost catchment of the Mahaweli River. Many water resource projects have been proposed utilizing the stream flow generated by the catchment. Though there are several hydrological models developed for the catchment, no statistical modelling has been carried out yet.

In the present statistical modelling approach, a multiple linear regression (MLR) model is developed for the response variable, which is the stream flow generated by the upper catchment of UKC above the stream flow measuring station at Calidonia. This sub catchment covers an area of 184 km^2 and MLR model was fitted regressing daily rainfall (R_t) and evaporation (E_t) on the stream flow at Calidonia (S_t). All the other parameters affecting the stream flow (slope of the catchment, vegetation cover, soil composition, land use, etc.) were considered to remain unchanged within the catchment during the modelling period.

Since the stream flow depends not only on the rainfall and evaporation on the same day but also the rainfall R_{t-j} and evaporation E_{t-j} on preceding days, previous days' rainfall and evaporation data were also considered as variables of the model. In addition, according to the seven trends observed in the time series plot of average rainfall year for the calibration period (1987-1990) six dummy variables (Z_j; j=1 to 6) were added to the model in order to increase the predictability by removing the effect of time trends.

Box-Cox transformation on the response variable at λ=-0.15 was used as the remedial measure for the heteroskedasticity of residuals. Extreme events were removed using standardized residuals with ± 1.96. Comparing predictability using R square value of several models starting with nine rainfall and five evaporation variables, the optimized model is obtained as follows.

$$W_t = 1.03 + 0.00801R_t + 0.0171R_{t-1} + 0.00794R_{t-2} + 0.00614R_{t-3} + 0.00773R_{t-4} - 0.0249E_t + 0.474Z_1 + 0.415Z_2 + 0.587Z_3 + 0.509Z_4 + 0.713Z_5 + 0.400Z_6$$

where $W_t = \dfrac{\left(S_t^{\lambda} - 1\right)}{\lambda}$ and $\lambda = -0.15$

Predictability (R^2) of the above model for the calibration period (1987-1990) is 0.788 where that for the validation period (1991-1993) is 0.777. As required for valid regression model, residuals follow a normal distribution and no multicollinearity among predictors. Hypothesis test (F-test) of the ANOVA provided enough evidence for a better fit of the model for the data.