TRANSFER FUNCTION MODEL OF THE RAINFALL-RUNOFF RELATIONSHIP OF A DRAINED CATCHMENT

Ayob Katimon[1], Ahmad Khairi Abd Wahab[1] and Ahmad Perwira Tarigan[2]

[1]*Department of Hydraulics and Hydrology, Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 UTM, Skudai, Johor*

[2]*Department of Civil Engineering, Universitas Sumatera Utara, Jalan Dr Mansyur No. 9, Kode Pos 20155, Padang Bulan, Medan, Sumatera Utara, Indonesia*

ABSTRACT

Transfer function model of the rainfall and runoff relationship with various complexities was developed to investigate the hydrological dynamics of a catchment that undergoes continual drainage process. The structure of the impulse response weight of the dynamic components of the models provides some physical meaning of the responses between the hydrological variables investigated. The study revealed that non-linear transfer function models of order one and noise term of ARIMA (1,0,0) best represents the monthly rainfall-runoff relationships of a drained catchment. The best- fitted transfer function model is capable of illustrating the cumulative hydrological effect of a catchment when subjected to reclamation and drainage

Keywords : *ARIMA, Drained Catchment, Rainfall-runoff, Transfer Function Model*

INTRODUCTION

Land drainage activities can alter the hydrological criteria of a catchment. To evaluate the long-term hydrological impact of drained catchment, such as found in agricultural peatland, a sound understanding of the hydrological functions are required. Traditionally, paired catchment experiments have been used to evaluate the effect of disturbances [1, 2]. However, this approach is not only time consuming but also require a properly planned and instrumented catchment study. The paired catchment approach is also unable to evaluate the relative importance of various factors that might underlie differences in results between the sites [3]. Another approach to evaluate the impact of disturbance and alteration on the hydrologic function of a catchment is by comparing the hydrologic record before and after the catchment being altered [1]. As such, pre and post drainage hydrological data are thus required. When both aforementioned approaches are not applicable either due to the absence of paired catchment or unavailability in pre-drainage hydrological record, as experienced in the present study, another alternative approach is needed to address the issue. One of many methods that can be considered is by applying a deterministic physical-based hydrologic model such as TOPMODEL [4] and DRAINMOD [5]. These models are basically hill-slope and parameter–based rich model that require an intensive quantitative knowledge in the physical characteristics of the catchment at spatial level.

In comparison to physical-based models, the transfer function time-series modelling approach has several advantages. Physical-based hydrological models require parameterisation and are based on the predetermined theory of hydrology, whereas a time series model is essentially a `black-box' [6]. It is purely calibration and requires no theory that links the input and outputs series. Under the situations where the theory of hydrological processes of drained catchment such as in peat areas is yet clearly defined [7,8], time series transfer function modelling approaches is found to be applicable.

The main objective of this paper is to develop a dynamic regression or transfer function (TF) model of the relationship between rainfall and runoff of a drained catchment particularly in agricultural peatland. It is intended to examine the ability of the TF model to understand the dynamics of the hydrologic behavior of the study catchment based on past time series data. Long-term rainfall-streamflows records obtained from a 184 hectare drained and agricultural catchment located in peat areas at Parit Madirono, Benut, Johor, Malaysia (103˚16'15" E, 01˚42'35"N, and called the Madirono catchment) [9] have been used in the present study.

MODELLING METHODOLOGY

The proposed methodology consists of applying the linear transfer function modeling approach with different hydrologic variables to investigate the stream flow dynamics of the study catchment. It is expected that the Transfer Function model makes it possible to provide a greater insight to the dynamic response of the stream flow to rainfall thus the overall hydrologic behavior of the catchment.

THE BASIC MODEL STRUCTURE

A single linear transfer function model representing the relationship between input and output time series data can be expressed as,

$$Y_t = C + v(B)\, X_t + N_t \qquad [1]$$

where Y_t is the output series or exogenous variables, X_t is the input series or endogenous variables, C is the constant term, $v(B)$ is the dynamic component or impulse response function of the model, N_t is the stochastic noise and B is the backshift operator. The stochastic noise N_t may be autocorrelated and is assumed to be independent of X_t .Because the dynamic term $v(B)$ in Equation 1 represents the dynamic behaviour of serial correlation of X_t at different time lags, it can be written in a polynomial form as [10],

$$v(B) = v_0 + v_1B + v_2B^2 + ... + v_kB^K \quad [2]$$

where v_0 through v_k are called the transfer function weight or impulse response weight. Thus, Equation 1 becomes [10],

$$Y_t = C + (v_0 + v_1B + v_2 B^2 + ... + v_k B^k)X_t + N_t \quad [3]$$

where k is the order of the transfer function, that is the longest lag of input series X_t used in the model.

PARSIMONIOUS MODEL STRUCTURE

An important criterion of good models is their simplicity or parsimoniousness. For that reason, the term $v(B)$ in Equation 3 can be rewritten in a simpler form as [11],

$$v(B) = \frac{\omega(B)}{\delta(B)} X_{t-b} + N_t \quad [4]$$

Thus, the parsimonious form of Equation 3 becomes

$$Y_t = C + \frac{\omega(B)}{\delta(B)} X_{t-b} + N_t \quad [5]$$

where,

$$\omega(B) = \omega_0 - \omega_1B - \omega_2B^2 - .... -\omega_sB^s,$$
$$\delta(B) = 1 - \delta_1B - \delta_2B^2 - ... - \delta_rB^r$$

r,s,b are constants. Constant b is called a delay factor that is, a delay of b period before X_t begin to influence Y_t . The constant r is the decaying factor of the impulse response weights and b is dead time factor.

FEEDBACK CHECKING

An important assumption in building a single-equation Transfer Function model is that there is no feedback from earlier values of the Y_t series to the current values of X_t series. Consider again a regression-lag model in Equation 3 of order k

$$Y_t = C + (v^0 + v^1B + v_2B^2 + ... + v_kB^k) X_t + N_t \quad [6]$$

Decomposing the B terms into $BX_t = X_{t-1}$, $B^2X_t = X_{t-2}$, etc.

$$Y_t = C + v_0X_t + v_1X_{t-1} + v_2X_{t-2} + ... + v_kX_{t-k} + N_t \quad [7]$$

To check the feedback effect of Y_t series on X_t series, the following equation is estimated,

$$X_t = C + b_1 X_{t-1} + b_2X_{t-2} + ... + b_kX_{t-k}$$
$$+ c_1Y_{t-1} + c_2Y_{t-2} + ... + c_kY_{t-k} + N_t \quad [8]$$

Using multiple regressions approach, the coefficient of c_1,c_2, c_k can be computed and significant statistical test can be employed.

MODELLING ALGORITHM

The basic modeling algorithm [10] is summarised in the flow chart presented in Figure 1.

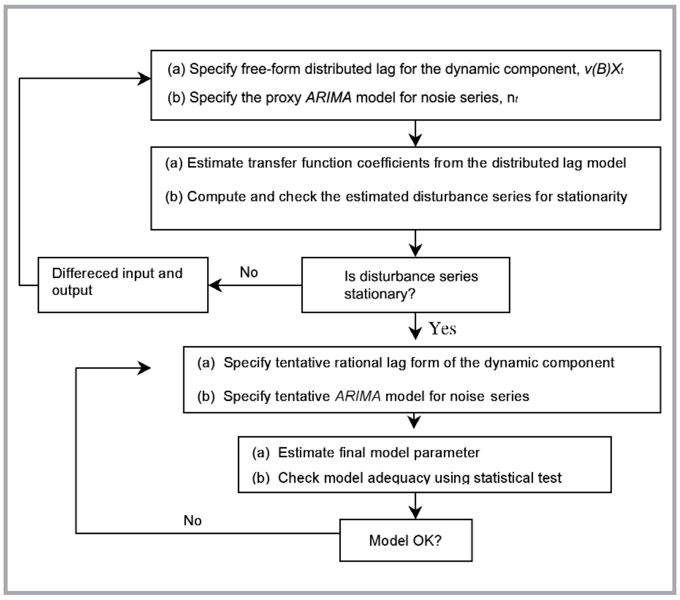

Figure 1: Transfer function modelling algorithm

FREE-FORM DISTRIBUTED LAG MODEL EQUATIONS

As written in the Equation 1 the distributed lag equation form of order k of the regression model is,

$$Y_t = C + v_0X_t + v_1X_{t-1} + v_2X_{t-2} + ... +v_k X_{t-k} + N_t$$

where v_0,v_1, ..., v_k are impulse response weights or transfer function coefficients and N_t is the noise series. The order of $v(B)$ is chosen arbitrarily according to their significant level. The response weight values are estimated using multiple regression approach.

PROXY ARIMA MODEL OF THE NOISE

In the Transfer Function model, a tentative or proxy Auto Regressive Integrated Moving Average *(ARIMA)* model for the noise series is used. The noise series produced by the distributed lag model is compared to that of the proxy model for their stationarity. The possible best-fit *ARIMA* model for the output series of the univariate model can be applied [6]. For instance, the best-fitted ARIMA model for the mean monthly flow series is in the form of *ARIMA (1,0,0)*. Thus *ARIMA (1,0,0)* is choosen as the proxy noise model for the development of a Transfer Function model of mean monthly rainfall-runoff relationship and written as,

$$(1- \phi_1B)\, Y_t = C + a_t \quad [9]$$

Considering only the noise terms,

$$Y_t = \frac{1}{1-\phi_1 B} a_t \quad [10]$$

where ϕ_1 is the $AR(1)$ parameter and a_t is the error series.

PRELIMINARY TRANSFER FUNCTION MODEL

Having known both the dynamic and noise component (proxy *ARIMA*) of the model, a Transfer Function model of order k is thus a combination of its distributed lag model (Equation 6) and *ARIMA* model of the disturbance series (Equation 10) and written as

$$Y_t = c + v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \ldots\ldots \ldots + v_k X_{t-k} + \frac{1}{1-\phi_1 B} a_t \quad [11]$$

EXAMPLE OF THE MODELLING IDENTIFICATION

The following paragraphs provide a step-by-step examples in developing a Transfer Function model for the mean monthly rainfall-runoff series taken from an experimental catchment [9]. Figure 2 shows the monthly rainfall-streamflow series of the study catchment. The rainfall series is notated by P_t , as the input variable while the flow series, Q_t , as the output variable. Rainfall series is the known factor affecting the variable of the runoff series.

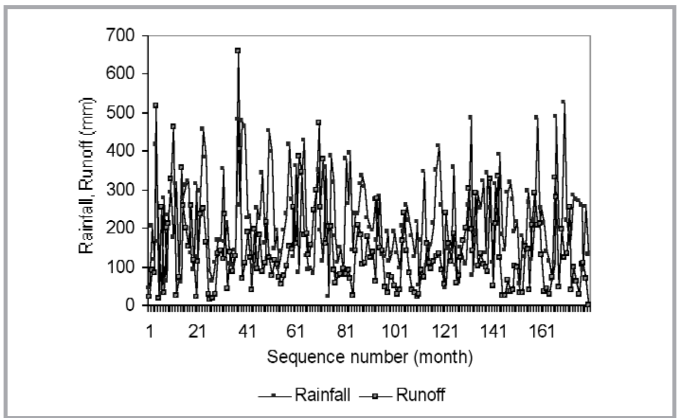

Figure 2: Mean monthly rainfall-runoff records of the study catchment

FEEDBACK CHECKING BETWEEN P_t AND Q_t

Assuming that the serial relationship between P_t and Q_t of order k is written as

$$Q_t = c + v_0 P_t + v_1 P_{t-1} + v_2 P_{t-2} + \ldots .\ v_k P_{t-k} + N_t \quad [12]$$

Thus the serial relationship between the present input series to their past time-lag series and their past output series is written as,

$$P_t = C + b_1 P_{t-1} + b_2 P_{t-2} + \ldots .\ b_k P_{t-k} + c_1 Q_{t-1} + c_1 Q_{t-1} + c_2 Q_{t-2} + \ldots . + c_k Q_{t-k} + N_t \quad [13]$$

Table 1: Statistical output of the feedback effect analysis of the mean monthly rainfall and flow series

Parameter	Standard deviation, σ	Coefficients	t-test	Significance level, P
Constant, C	28.42		5.896	0.000
c_1	0.079	0.386	0.611	0.542
c_2	0.086	-0.034	0.409	0.683
c_3	0.085	0.010	0.123	0.903

Since the purpose of this procedure is to check whether there is feedback effect of the output series on the input series, thus our interest should be focused on the values of c_1, c_2, ..., c_k only. Using multiple regression approach, the estimated of ci values up to order 3 are given in Table 1. Except for the constant, C , it can be seen that all the corresponding t values are small and not significant at 5% level. Thus, at this particular juncture, decision is made that there is no feedback effect from the past of the output (flow) to the input series (rainfall).

FITTED REGRESSION LAG EQUATION

The fitted multiple regression models to the mean-monthly rainfall-runoff data of lagged values up to x_{t-3} is

$$Q_t = 191.68 + 0137P_t - 0.0211P_{t-1} - 0.059P_{t-2} - 0.165P_{t-3} + N_t \quad [14]$$

Table 2: Estimates of the transfer function coefficients and their statistics

Parameter	Coefficients	Std. Error	t-test	Significance level, P
C	191.684	47.682	4.020	0.000
v_0	0.137	0.073	1.888	0.061
v_1	-0.02114	0.073	-0.290	0.772
v_2	-0.05906	0.073	-0.809	0.419
v_3	-0.165	0.072	-2.278	0.024

In fitting this model, it was assumed that the noise series N_t belong to an *ARIMA* (1,0,0) model. The error series of this proxy model appeared stationary. The statistical evidence of Equation 14 is presented in Table 2. The plot of the transfer function coefficients against their lags is shown in Figure 3. It is clearly shown that a non-exponential pattern of the decaying factor exists. Judging from Table 2 and Figure 3 and using identification rules outlined by among others [10,11] the following (*b,r,s*) model order is identified.

i. From the *P* values in column 5 of Table 2, it is clear that there is no delay. The first significant coefficient is at lag 0. So the model constant *b* is set as 0. This has to be expected with the mean monthly rainfall-runoff relationship in hydrology.
ii. From Figure 2 the coefficient began to decay at lag 0. Thus *s* is 0.

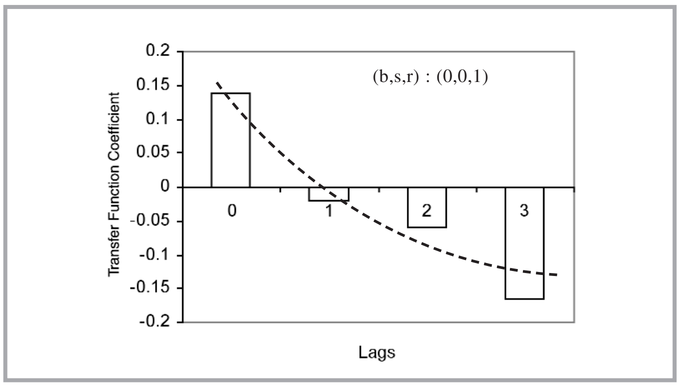

Figure 3: Estimates of transfer function coefficient in equation 14, showing a decaying pattern

iii. The decay pattern of the coefficient (indicated by the dotted line) follows a simple exponential decay. Thus r is 1.

An error series is thus obtained

$$N_t = Q_t - 191.68 - 0.137P_t + 0.0211P_{t-1} + 0.059P_{t-2} + 0.165P_{t-3} \qquad [15]$$

Figures 4 (a), (b) and (c) show the regression errors and their *ACF* and *PACF* plots of model of Equation 15. As noticed by the Auto-Correlation-Function *(ACF)* and the Partial-Auto-Correlation-Function *(PACF)* plots of the N_t series, the significant spikes at lags 1 and 3 suggest that *AR(1)* or *MA(1)* or *AR(3)* or *MA(3)* or combination of *ARIMA* model could be the best fitted model. Nevertheless using the Aikaike Information Criteria *(AIC)* [11], the *AR(1)* model has the smallest *AIC* value. Thus, in the case of mean monthly

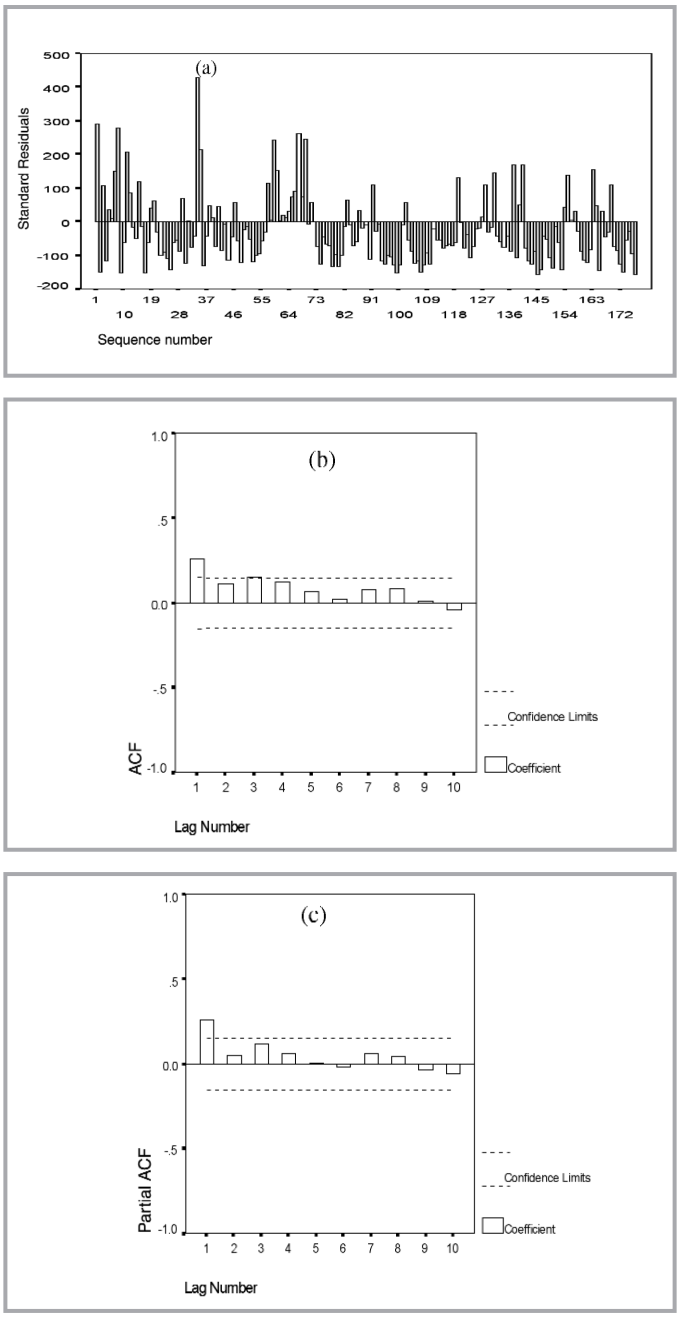

Figure 4: (a) Regression errors from Equation 15
(b) ACF plot and
(c) PACF plot

Rainfall-Runoff relationship, *ARIMA (1,0,0)* model is adopted as the best fitted error series and this conform with that of our first assumed (proxy) model.

Now, with a dead time $b = 0$, the general form of the parsimonious for the full model of Equation 11 is

$$Q_t = C + \frac{\omega(B)}{\delta(B)} P_{t-b} + N_t \qquad [16]$$

where,

$\omega(B) = \omega_0 - \omega_1 B$
$\delta(B) = 1 - \delta_1 B$

$$Q_t = C + \frac{\omega_0 - \omega_1(B)}{1 - \delta_1(B)} P_t + N_t \qquad [17]$$

where,

$N_t = \dfrac{1}{(1- \phi_1 B)}$ at and at is the error series.

Having known the estimated initial values of ω_0 and ϕ_1 , using ordinary Least Square Method (LS) parameter ω_1 , δ_1 and constant C are estimated.

PARAMETER ESTIMATION USING ORDINARY LEAST SQUARE (OLS)

The objective is to find the best value of model parameter ω_0, ω_1, δ_1, ϕ_1 and C, so that a best-fitted model to present the input-output series relationship is obtained. A preliminary estimate is chosen and the computer program refines the estimate iteratively until the Sum of Square Errors *(SSE)* is minimised. For a regression with one independent variable, these estimators are [12],

$$b_1 = \frac{\sum_{i=1}^{n} (X_{1,i} - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_{1,i} - \overline{X})^2}$$

where $\overline{X}$ and $\overline{Y}$ are the sample means of X_i and Y_i .

As practical rules [11] the initial or range value of ω_0 and ϕ_1 can be directly taken from the regression lag model result in Equation 14. Thus, for mean monthly rainfall-flow relationship, the final model is,

$$Q_t = C + \frac{\omega_0 - \omega_1 B}{1 - \delta_1 B} P_t + \frac{1}{(1 - \phi_1 B)} a_t \qquad [18]$$

The final model parameters ω_0, ω_1, δ_1, ϕ_1 and C are estimated iteratively using the ordinary Least Square *(LS)* method algorithm. Equation 18 cannot be solved analytically because it involves non-linear functions. Thus, in the present study the parameters estimation were conducted iteratively using a program written in MATLAB language. For the mean monthly rainfall-runoff relationship of the present example, the following model parameters are obtained:

$C = 159.53$, $\omega_0 = 0.1773$, $\omega_1 = 0.0010$,
$\delta_1 = 0.3030$, $\phi_1 = 0.2348$
The sum of square error of the whole model is 37.45.

As from the previous section, with $\omega(B)$ is of order zero (s = 0), $\delta(B)$ is of order one ($r = 1$) and noise term is *ARIMA (1,0,0)*, the final model can now be written as,

$$Q_t = 159.53 + \frac{0.1773 P_t}{(1 - 0.3030B)} + \frac{1}{(1 - 0.2348B)} a_t \qquad [19]$$

where $B = P_{t-1}$ and a_t is the error series.

The Transfer Function model in Equation 19 has a Mean Square Error of 37.45. The model coefficient also satisfies $|\delta_1| < 1$, a criterion used to check the stability of a first order model [13]. On these counts it is a reasonable model. The final step was to check the error series of the final model. The transfer function is of order (b,s,r) = *(0,0,1)* and noise term is of *(p,d,q) = (1,0,0)*.

MODEL INTERPRETATION

The interpretation for the Transfer Function model in Equation 19 for mean monthly rainfall-streamflow relationships is as follows. When rainfall rises by one unit, runoff responses immediately *(b = 0)*. Then runoff rises (ω_0 is positive) initially by 0.177 units ($\omega_0 = 0.177$). Subsequent time-lagged additions to the runoff series, Q_t get smaller at each succeeding period, according to the first order exponential decay pattern, with the decay coefficient δ_1 = 0.3030. The constant term (C = 159.53) indicates that the flow series rises by 159.53 units at each time period in addition to any other movements dictated by the transfer function or disturbance *ARIMA* pattern.

MODEL PERFORMANCE AND CONCLUSION

The root mean square error (RMSE) and goodness of fit (R^2) method are used to examine the model performance. The RMSE for model in Equation 19 for the mean monthly rainfall-steramflow relationship is 31.88 mm, a value that is reasonably small. Figure 5 compares the simulated flows using the TF model in Equation 19 to be observed values. Figure 6 is the plot of simulated series against observed series. The relationship between the mean monthly rainfall

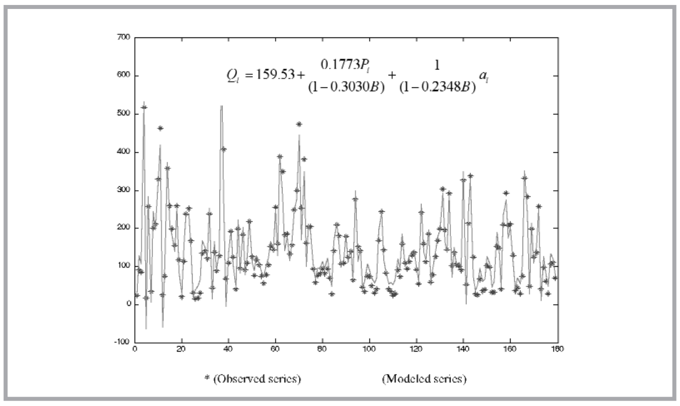

Figure 5: Sequence of observed and modelled flow series

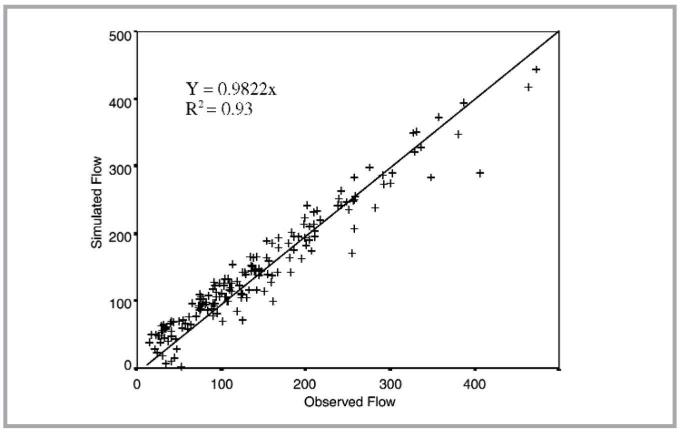

Figure 6: Plot of modelled flow series against observed series

and the mean daily rainfall is fairly represented by a Transfer function model in Equation 19. Nevertheless, based on the scatter diagram in Figure 6 the model was underpredicted by 2% with R^2 = 0.98. The transfer function models of the rainfall-runoff relationship are capable of showing the hydrologic dynamics of the catchment by means of their steady state function.

ACKNOWLEDGEMENT

Rainfall-runoff data provided by Hydrology Section, Department of Irrigation and Drainage is acknowledged. Part of this paper has been presented in the Regional Ecological Modelling Conference (ECOMOD 2004), organised by Mathematical and Biological Department, Universiti Sains Malaysia, 15-16 Sept. 2004. The author would like to thank the two anonymous referees for their constructive comments. ■

REFERENCES

[1] Newson, M. D. and Robinson, M. (1983). *"Effects of Agricultural Drainage on Upland Streamflow: Case Studies in Mid-Wales." J. Env. Manag.* 17:333-348

[2] Robinson, M. (1990). *"Impact of Improved Land Drainage on River Flows."* Report No 113. Wallingford: Institute of Hydrology. 226 pp

[3] Schellekens, J. (2000). *"Hydrological Processes in a Humid Tropical Rain Forest: A Combined Experimental and Modeling Approach."* Vrijie Universiteit: Ph.D. thesis

[4] Beven, K. J. (1989). *"Changing Ideas in Hydrology-the Case of Physically-based Models." J. Hydrol.* 105:157-172

[5] Skaggs R.W. (1991). *"A Computer Simulation Study of Pocosin Hydrology." Wetlands*. 11:399-416 Control." 3rd ed. New Jersey: Prentice Hall.

[6] Hipel, K.W. and McLeod, A. I. (1994). Time Series Modeling of Water Resources and Environmental Systems. Amsterdam: Elsevier

[7] Ayob Katimon, Ahmad Khairi Abd Wahab and Lulie Melling (2002). *"Understanding the Hydrological Behaviour of Tropical Peat Swamps: A Key Factor towards its Sustainable Management"*. Paper presented at the Southeast Asian Natural Resources and Environmental Management Conference, Kota Kinabalu, Sabah, 17-18 Oct.2002

[8] Mulqueen, J. (1986). *"Hydrology and Drainage of Peatland." Env. Geol. Water Sci.* 9(1):15-22

[9] Ayob Katimon, Ahmad Khairi Abd wahab and Kawi Bidin (2003). Hydrologic Mechanism Responsible for Storm Flow Generation Processes of a drained Peat Catchment. *The Journal of The Institution of Engineers, Malaysia.* 64(4): 47-55

[10] Pankratz, A. (1991). Forecasting with Dynamic Regression Models. United States: John Wiley. 386 pp

[11] Makridakis, S., Wheelwright, S. C. and Hyndman, R .J. (1998). "Forecasting Methods and Application." 3rd ed. New Yorks: John Wiley

[12] Gujarati, D. N. (1988). Basic Econometrics. 2nd ed. Singapore: McGraw-Hill. 26-133

[13] Box, G. E. P., Jenkins, G. M. and Reisel, G. C. (1994). "Time Series Analysis Forecasting and Control." 3rd ed. New Jersey: Prentice Hall.

PROFILES

Dr Ayob Katimon
Dr Ayob Katimon is a lecturer in Water Resources, Irrigation and Drainage at the faculty of Civil Engineering, Universiti Teknologi Malaysia. He graduated from various universities, B.Eng (UPM), M.S. (Colorado State University, Fort Collins) and Ph.D. (UTM). He is a corporate member of I.E.M., MANCID and Malaysian Hydrological Society (MHS)

Dr Ahmad Khairi Abd Wahab
Dr Ahmad Khairi Abd Wahab is an Associate Professor in hydraulics and coastal engineering at Faculty of Civil Engineering, Universiti Teknologi Malaysia. Currently, Dr Khairi is the Head, Department of Hydraulics and Hydrology of the faculty. He graduated with B.Eng. (University of Southampton), M.Eng. (UTM) and Ph.D. (University of Swansea, UK).

Dr Ahmad Perwira Tarigan
Dr Ahmad Perwira Tarigan is a senior lecturer at the Department of Civil Engineering, Universitas Sumatera Utara. He is a former postgraduate student at the Department of Hydraulics and Hydrology, Faculty of Civil Engineering, UTM.