

Development of a Multi-site Flow Generation Model in Northern Taiwan

Ray-Shyan Wu and Chih-Cheng Chang

Department of Civil Engineering, National Central University, Chung Li, Taiwan

Abstract Most streamflow generation models are developed by taking only a series of recorded stream flow data into consideration. However, with the trend of integrated water resource systems management, it becomes unacceptable to treat each stream in the system independently. Furthermore, the climate pattern, neglected in most models, should be considered in the model. This is because the variation of hydrological factors, such as terrain, climate, and soil types are limited in a nearby area. That is, there must be a certain correlation between the streamflows. With this in mind, a multi-site streamflow generation model is developed in this study. The distinctive feature of this model is to focus on three flow components, the base flow, the monsoon-induced flow, and the Typhoon flood, in the northern Taiwan area. The mean and variation of each flow are discussed based on these three components. Moreover, the correlation, both autocorrelation and cross correlation, in the flow series among rivers is calculated to maintain lag one time series of flow in the overall watershed. Tanshuei River Watershed, consisting Peishih River, Nanshih River, and Tahan River, is chosen as a case study. The result shows a much better fit of characteristics and cross correlation between streamflows. This model provides a practical multi-site flow generation model for water resource management.

1. INTRODUCTION

In water resource studies, flow in stream is the determining factor in assessing potential water supply capacity for reservoir systems. Traditionally, in-situ flow record is used in deciding reservoir capacity. However, recorded flow is only one realization among all-possible cases. Therefore, an adequate flow generation model is required in order to study all possible system performance.

Most streamflow generation models are developed by taking only an individual streamflow record into consideration. Then, a flow generation model, such as an autoregressive Markov model, is applied. Other researches proposed the disaggregation model for the generation of synthetic flows at the annual level and seasonal level (Lin and Lee, 1988). However, with the trend of applying integrated water resource systems management, it becomes unacceptable to treat each stream in the system independently. This is because the variation of hydrological factors, such as terrain, climate, and soil types are limited in a nearby area. That is, there must be a certain correlation between the streamflows. To deal with the multi-site streamflows, researchers, e.g., Wilson [1973], proposed multivariate time series models.

Furthermore, local characteristics of climate should be considered in flow generation. With this in mind, a carefully study on the flow components for each stream should be required. The distinctive feature in this study is to focus on three major flow

components, the base flow, the monsoon-induced flow, and the Typhoon flood, in the northern Taiwan area. The mean and variation of each flow are discussed based on these three components.

2. MODEL FORMULATION

2.1 Normalization of Streamflow

It is important that the generated streamflow should carry statistic characteristics of a certain distribution. However, in the real world the minimum of a streamflow is zero and the maximum is almost unlimited (Matalas, 1967). To deal with this fact, a common used procedure in streamflow transformation is applied (Box and Cox, 1964). Firstly, the streamflows are assumed to be a 2-parameter log-GAMMA distribution, and a transformation of data into normal distribution is performed, as follows.

(1). Logarithmic transformation

$$Y_t = \log(X_t) \quad (1)$$

(2). Normalization

$$Y'_{v,\tau} = \frac{Y_{v,\tau} - \hat{\mu}_\tau}{\hat{S}_\tau} \quad (2)$$

where, $\hat{\mu}_\tau$ and \hat{S}_τ , the mean and standard deviation of the streamflow in time τ , are

commonly represented with a Harmonic synthetic model. However, applying the Harmonic synthetic model is simply a data fitting procedure, with no physical meaning what so ever. In other words, it is possible to perfectly reproduce the recorded data set with carefully assigned parameters. However, there is no guarantee that the future streamflow would be the same. To tackle this, a decomposition approach based on the precipitation patterns is proposed, which will be discussed in next section.

(3). Wilson-Hilferty transformation

$$Y'_{u,r} = \begin{cases} \max(Y'_{u,r}, -2/g_r(Y')) & g_r(Y') > 0 \\ \min(Y'_{u,r}, -2/g_r(Y')) & g_r(Y') < 0 \end{cases} \quad (3)$$

$$Z_{u,r} = \begin{cases} \frac{6}{g_r(Y')} \cdot \left[\left(\frac{g_r(Y') \cdot Y'_{u,r}}{2} + 1 \right)^{1/3} - 1 \right] + \frac{g_r(Y')}{6} & g_r(Y') \neq 0 \\ Y'_{u,r} & g_r(Y') = 0 \end{cases} \quad (4)$$

where, $g_r(Y')$ denotes the skewness of $Y'_{u,r}$.

The transformed flow data is then tested for its normality with the Kolmogorov-Smirnov test. This test provides bounds within which every observation should lie if the sample is actually drawn from a specific distribution, in this case the Normal distribution.. In this study, the test specifies that

$$\Pr\left[G^{-1}\left(\frac{i}{n} - C_\alpha\right) \leq X_{(i)} \leq G^{-1}\left(\frac{i}{n} + C_\alpha\right)\right] = 1 - \alpha \quad (5)$$

where C_α is the critical value of the test at significant level α .

2.2 Formulation of a Single Site Streamflow Model

The Harmonic synthetic model is widely applied in generating streamflows. A standard model is

$$\hat{\mu}_\tau = A_0 + \sum_j \left[A_j \cdot \cos \frac{2\pi j \tau}{36} + B_j \cdot \sin \frac{2\pi j \tau}{36} \right] \quad (6)$$

$$\hat{S}_\tau = A_0 + \sum_j \left[A_j \cdot \cos \frac{2\pi j \tau}{36} + B_j \cdot \sin \frac{2\pi j \tau}{36} \right] \quad (7)$$

where, τ denotes the time of the year, in 10-days.

A common approach in deciding the number of parameter in the model is through some statistical analysis such as investigating AIC and BIC. However, due to the complexity of the data set, it may require more than 5 parameters to provide a good fit. Instead, temporal characteristics of the

streamflow are carefully studied and a decomposition procedure is proposed.

Figure 1 indicates two peaks in streamflow data annually. The occurrence of these two peaks is due to the monsoon season and the Typhoon season. Therefore, streamflow can be break into three components, the base flow, the monsoon-induced flow, and the Typhoon flood. Furthermore, each of the three components can be modeled separately and then added together, as in Figure 2. That is,

$$\hat{\mu}_\tau = f_c + f_{b1} \cos \frac{2\pi(\tau - t_b)}{36} + f_{b2} \sin \frac{2\pi(\tau - t_b)}{36} + F_m + F_T \quad (8)$$

where, f_c, f_{b1}, f_{b2} are three parameters that are used to formulate a simple Harmonic synthetic model for the base flow. F_m and F_T , the monsoon-induced flow and the Typhoon flood, are expressed as

$$F_m = f_m \sin \frac{2\pi(\tau - t_m)}{36} \quad (\tau \in \text{monsoonseason}) \quad (9)$$

$$F_T = f_T \sin \frac{2\pi(\tau - t_m)}{36} \quad (\tau \in \text{Typhoonseason}) \quad (10)$$

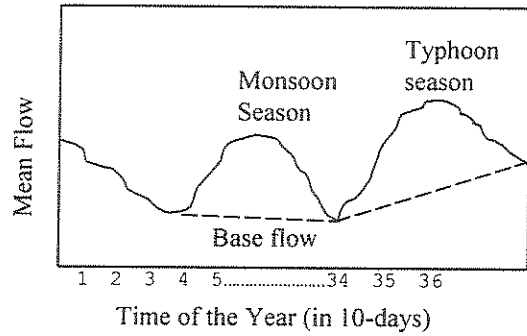


Figure 1: Decomposition of Flow

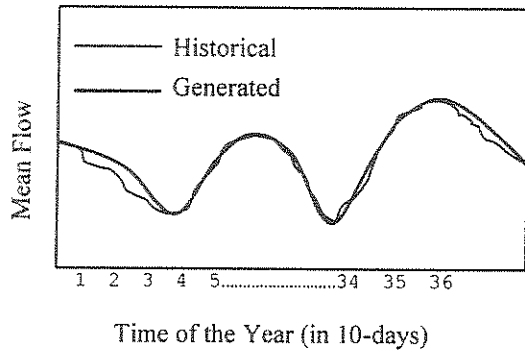


Figure 2: Mean Annual Flow from Generation Model

2.3 Formulation of a Multiple Site Streamflow Model

Following Loucks et al. (1981), a sequence of synthetic flows can be generated by the model

$$Z_{y+1} = AZ_y + BV_y \quad (11)$$

where, A and B are $(n \times n)$ matrices whose elements are chosen to reproduce the lag 0 and 1 cross covariances of the transformed flows at each site; V_y is a column vector of standard normal random variables.

3. APPLICATION OF MODEL

3.1 Case Study Area

Tanshuei River Watershed, consisting Peishih River, Nanshih River, and Tahan River, is chosen as a case study, as shown in Figure 3. The river is in the northern of Taiwan and flows though the capital city Taipei.

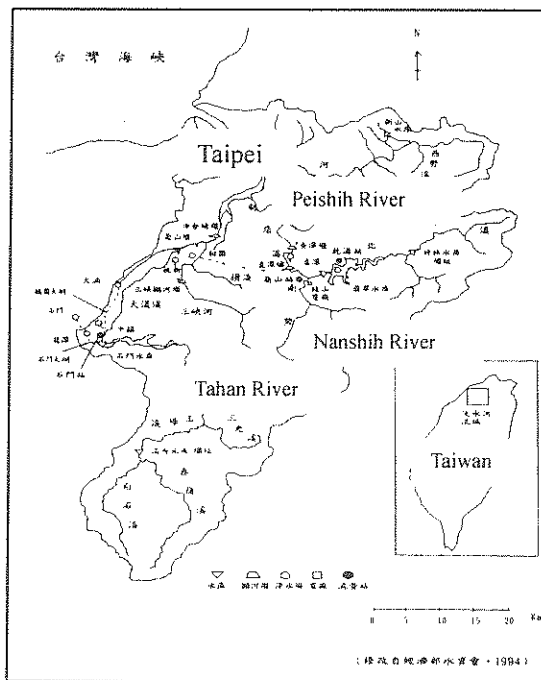


Figure 3: Tanshuei River Watershed

The model is applied on flow data for 43 years. That is, three sets of 1548 10-day flow record for each river is analyzed. The cross correlation coefficients are reported in Table 1, with the highest being 0.8. The data indicates 2 peaks, one is between time period 14 to 21 and the other between 22 to 33, corresponding to the monsoon season and Typhoon season, respectively. The data is complied with 2-parameter log-GAMMA transformation and the model previously described is applied to estimate all parameters necessary.

Table 1: Cross Correlation among Three Rivers

Cross Correlation Coefficient	Peishih River	Nanshih River	Tahan River
Peishih River	1.00	0.89	0.69
Nanshih River	0.89	1.00	0.80
Tahan River	0.69	0.80	1.00

3.2 Streamflow Generation Model

The resulting streamflow generation model is as follows,

(1) Peishih River

$$\begin{aligned} \hat{\mu}_\tau = & 4.8 + 0.6 \times \cos \frac{2\pi(\tau-1)}{36} \quad (FOR \tau=1-36) \\ & + 1.35 \times \sin \frac{2\pi(\tau-13)}{14} \quad (FOR \tau=13-20) \\ & + 1.0 \times \sin \frac{2\pi(\tau-18)}{34} \quad (FOR \tau=18-35) \end{aligned} \quad (12)$$

$$\hat{S}_\tau = 0.8 + 0.5 \times \sin \frac{2\pi(\tau-18)}{34} \quad (FOR \tau=18-35) \quad (13)$$

(2) Nanshih River,

$$\begin{aligned} \hat{\mu}_\tau = & 5.2 + 0.3 \times \cos \frac{2\pi(\tau-1)}{36} \quad (FOR \tau=1-36) \\ & + 1.35 \times \sin \frac{2\pi(\tau-14)}{14} \quad (FOR \tau=14-20) \\ & + 1.0 \times \sin \frac{2\pi(\tau-18)}{34} \quad (FOR \tau=18-35) \end{aligned} \quad (14)$$

$$\hat{S}_\tau = 0.55 + 0.4 \times \sin \frac{2\pi(\tau-18)}{34} \quad (FOR \tau=18-35) \quad (15)$$

(3) Tahan River

$$\begin{aligned} \hat{\mu}_\tau = & 5.3 + 0.2 \times \cos \frac{2\pi(\tau-15)}{36} + 0.2 \times \sin \frac{2\pi(\tau-15)}{36} \quad (FOR \tau=15-50) \\ & + 0.7 \times \sin \frac{2\pi(\tau-14)}{12} \quad (FOR \tau=14-20) \\ & + 1.0 \times \sin \frac{2\pi(\tau-19)}{28} \quad (FOR \tau=19-33) \end{aligned} \quad (16)$$

$$\begin{aligned} \hat{S}_\tau = & 0.6 + 0.1 \times \cos \frac{2\pi(\tau-15)}{36} + 0.1 \times \sin \frac{2\pi(\tau-15)}{36} \quad (FOR \tau=15-50) \\ & + 0.3 \times \sin \frac{2\pi(\tau-19)}{28} \quad (FOR \tau=19-35) \end{aligned} \quad (17)$$

The coefficient of correlation between predicted and record flow in Peishih River, Nanshih River, and Tahan River are 0.91, 0.97, and 0.98, respectively.

Furthermore, let P_t , N_t and T_t denote the flow in Peishih River, Nanshih River, and Tahan River in period t , respectively. Then,

$$\begin{bmatrix} P_{t+1} \\ N_{t+1} \\ T_{t+1} \end{bmatrix} = \begin{bmatrix} 0.51 & -0.14 & -0.02 \\ 0.22 & 0.33 & 0.04 \\ 0.07 & -0.07 & 0.65 \end{bmatrix} \begin{bmatrix} P_t \\ N_t \\ T_t \end{bmatrix} + \begin{bmatrix} 0.88 & 0 & 0 \\ 0.72 & 0.40 & 0 \\ 0.58 & 0.26 & 0.54 \end{bmatrix} \begin{bmatrix} V_{1t} \\ V_{2t} \\ V_{3t} \end{bmatrix} \quad (18)$$

4. RESULTS AND DISCUSSION

The flow model is applied to generate 43 year of 10-days flow for three rivers. Figure 4 and 5 report mean and standard derivation of the synthetic flows and recorded flows for Peishih River in logarithmic scale. Clearly, the model is able to provide a reasonable good fit both in the mean and the variance of the flow. However, a slight distortion in reproducing the same variance during the Typhoon season is noted. The reason is that in the field a high variance happens in the beginning of the Typhoon season (during period 22 to 33), while in this model a symmetric sinusoid function is assumed. Similar results are obtained for Nanshih River and Tahan River.

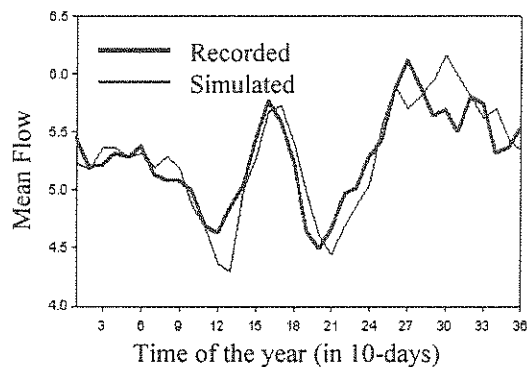


Figure 4: Mean of Simulated and Recorded Flow in Peishih River in Log-scale

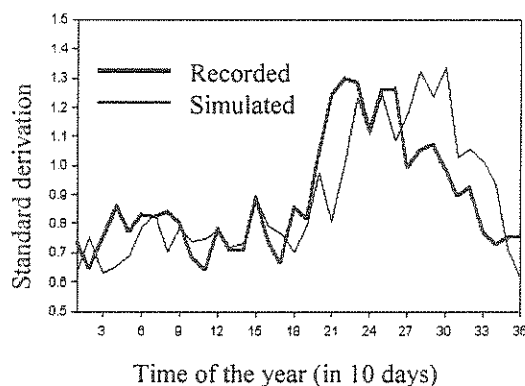


Figure 5: Standard Derivation of Simulated and Recorded Flow for Peishih River in Log-Scale

To check the performance of the model in practice, the mean and standard derivation of the synthetic flows and recorded flows are compared, as shown in Figure 6 and 7. The fit is slightly worse than that in the logarithm base. In fact, an unavoidable distortion in the high flow condition is observed. This is due to the logarithmic transformation and the assumption of Normal distribution in the analyses. As a result, the standard derivation is over estimated in the high flow conditions. The procedure in taking the logarithm of the flow and then exponential the simulated flow back would magnify the flow more or less. The logarithmic transformation should be applied in flow data with little or no skewness (Stedinger, 1980). The skewness of the flow data in Peishih, Nanshih and Tahan River are 0.01, 0.01 and 0.4, respectively. The skewness of the simulated flows are all less than 0.1. This result is discussed in Lettenmaier and Burges (1977).

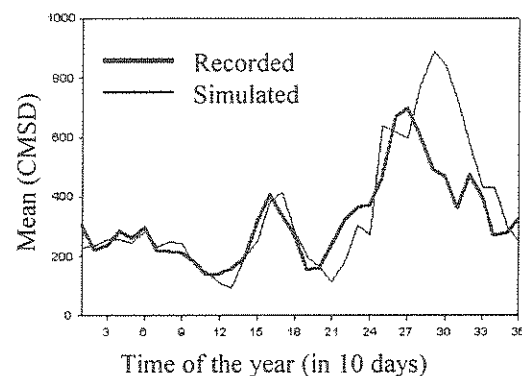


Figure 6: Mean of Simulated and Recorded Flow for Peishih River

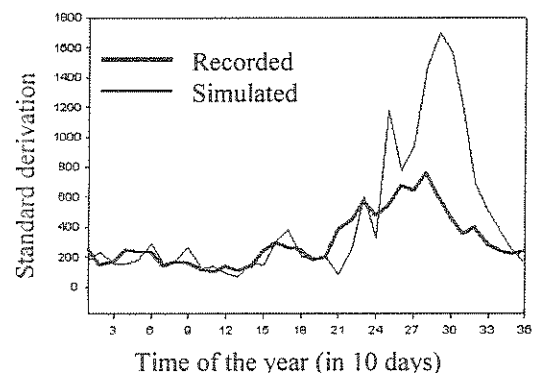


Figure 7: Standard Derivation of Simulated and Recorded Flow for Peishih River

Apparently, this model can be classified as an ARMA(1,0) model. The cross-correlation between rivers is maintained. For example, Figure 8 and 9 shows the cross-correlation between Peishih River and Nanshih River and that between .

The lag one cross-correlation in simulated flow is 0.8, while in flow record that value is 0.89. The periodical pattern in the cross-correlation diagram indicates the annual cycle of dry and wet seasons.

The reproduction of the cross-correlation structure is crucial when a system of reservoirs is operated as a whole. In the field, it is likely that when drought occurs, rivers are all in low flow condition. With the help of the proposed model, such flow condition can be simulated. This is the idea that would be implemented in the case study.

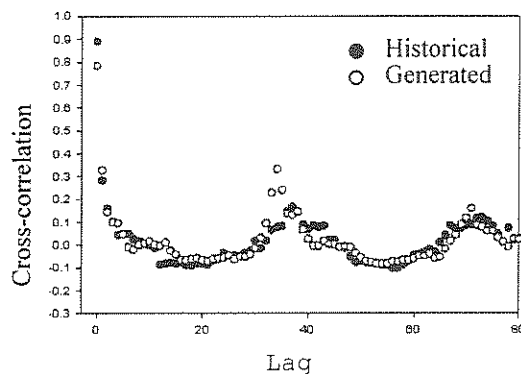


Figure 8: Cross-correlation between Peishih River and Nanshih River

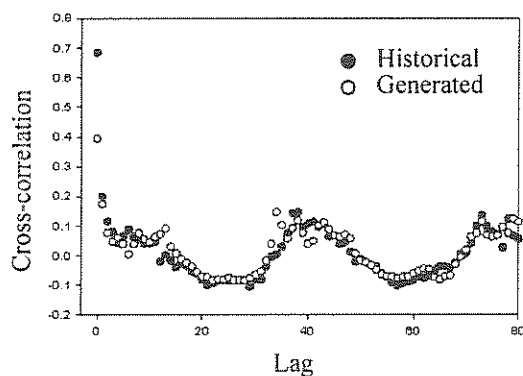


Figure 9: Cross-correlation between Peishih River and Tahan River

5. CONCLUSIONS

1. A multi-site streamflow generation model is developed in this study. The distinctive feature of this model is to focus on three flow components, the base flow, the monsoon-induced flow, and the Typhoon flood. The synthetic flow models are constructed for the mean and variation of each flow component. Moreover, the correlation, both autocorrelation and cross

correlation, in the flow series among rivers is simulated with an ARMA(1,0) model.

2. The break down of the flow into three major flow components, the base flow, the monsoon-induced flow, and the Typhoon flood, are based on local characteristics of climate in Taiwan. It is possible to apply this model to other regions by carefully selecting local flow components. Furthermore, the procedures in choosing the time periods corresponding to each flow components can be refined. In this paper, a somewhat subjective approach is applied.
3. The reproduction of the cross-correlation structure is crucial when a system of reservoirs is operated as a whole. The challenging task in reservoir operation is to maintain water supply during drought periods when rivers are all in low flow condition. With the help of the proposed model, such flow condition can be simulated.
4. The log-GAMMA transformation used in this study unavoidable overestimates the mean and variation in high flows. The impact may have to be investigated if flood control is of concern.
5. The proposed model is tested in the case study, Tanshuei River Watershed. The result shows a much better fit of characteristics and cross correlation between streamflows. This study proposes a practical multi-site flow generation model for water resource management.

6. ACKNOWLEDGMENT

Support for this work was provided, in part, by a grant from the National Science Council of the Republic of China in Taiwan through contact NSC 86-2621-E008-011.

REFERENCE

- Box, G.E.P., and D.R. Cox, An analysis of transformation, *Journal Royal Statistical Society.*, B26, 211, 1964.
- Lettenmaier, D.P., and S.J. Burges, An operational approach to preserving skew in hydrologic models of long-term persistence," *Water Resource Research*, 13(1), 113-124, 1977.
- Lin, G.F. and F.C. Lee, Multi-seasonal key series for disaggregation scheme, Fifth International Symposium on Stochastic Hydraulics, University of Birmingham, UK, 1988.
- Loucks D.P., J.R. Stedinger. and. D.A. Haith, *Water Resource System Planning and Analysis*, Prentice-Hall, USA, 1981.

Matalas, N.C., Mathematical assessment of synthetic hydrology, *Water Resource Research*, 3(4), 937-945, 1967.

Stedinger, J.R., Fitting Log Normal distribution to hydrologic data, *Water Resource Research*, 16(3), 481-490, 1980.

Wilson, G.T., The estimation of parameters in multivariate time series models, *Journal of the Royal Statistical Society B*, 35(1), 76-85, 1973.