Assessing Hydrological Model Behaviors by Intercomparison of the Simulated Stream Flow Compositions: Case Study in a Steep Forest Watershed in Taiwan

Jr-Chuan Huang[1*],Tsung-Yu Lee[1],Jun-Yi Lee[1],Shih-Chieh Hsu[2], Shuh-Ji Kao[2,3],Fi-John Chang[4]

1. Department of Geography, National Taiwan University, Taipei, Taiwan
2. Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan
3. State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, China
4. Department of Engineering for Sustainable Environment, National Taiwan University, Taiwan

*Corresponding Author: Jr-Chuan Huang, Assistant Professor, Department of Geography, National Taiwan University, Taipei, TAIWAN

Email: riverhuang@ntu.edu.tw

Submitted to Hydrology and Earth System Sciences

Abstract

The accuracy of streamflow composition simulated by different models has been rarely discussed. In this study, total 23 flood events covering full rainfall spectrum were simulated by using HBV and TOPMODEL. Simulated streamflow compositions were compared with hydrograph decomposed by independent geochemical data via end-member mixing analysis (EMMA). Results showed that both models gave satisfactory streamflow simulation in terms of the Nash efficiency coefficient, correlation coefficient, and discharge volume. However, the modeled interflow and base flow behaved differently with the changing storm intensity and duration. The HBV simulated base flow considerably increased as the storm duration prolonged; by contrast, the TOP-derived base flow remained stable. On the other hand, HBV prefers generating less interflow to percolate more to the base flow for fitting the stream flow. Accordingly, HBV is more suitable for thin soil layer. We suggested that a proper model selection should take the implicit environmental background into account for simulating reliable streamflow composition. Compared with the EMMA-derived flows, both models showed a significant time lag (2-4 hr). If EMMA-derived hydrograph is real, the modeled base flow responses are required to speed up. Our model's intercomparison against independent validation by geochemical data is a good means of studying the model behaviors. The selection of a more appropriate hydrological model should consider the characterization of the model structure and the watershed characteristics.

Keywords: stream flow composition, HBV, TOPMODEL, end-member mixing analysis (EMMA), Taiwan

註解 **[R1]:** For detailed comment #1

註解 **[R2]:** For technical comment #1

1. Introduction

Simulating the stream flow accurately is one of the main concerns of scientists and managers, particularly in hydrology science and water resource assessment. For this goal, hydrological models are implemented through different conceptualizations of simplified representations of the real world (Beven, 2001; Refsgaard and Henriksen, 2004). Therefore, a number of hydrological models with different model structures were proposed and applied around the world. Undoubtedly, this significant progress in hydrological modeling works has been facilitating many discharge-relevant applications. Currently, the attentions have shifted to understanding more on the model structure and the corresponding behaviors for advanced interpretations (Reed *et al.*, 2004; Clark *et al.*, 2008). For example, some previous studies applied different model structures (e.g., runoff generations or routings) on the same catchment to determine the model suitability and applicability (Winchell *et al.*, 1998; Valeo *et al.*, 2001; Johnson *et al.*, 2003). These comparative studies revealed that models with different structures could satisfactorily simulate the stream discharge for the same catchment. However, the selection of the hydrological models and the model structure uncertainties are still not fully understood. Clark *et al.* (2008) have applied 79 unique model structures by combining the components of four existing hydrological models into catchments. They concluded that the model structure uncertainty is as important as the parameter uncertainty, indicating that intercomparison among models can give insight into the understanding of hydrological models. In addition, Weiler *et al.* (2003) integrated the instantaneous unit hydrograph and the temporal variability of rainfall isotopic composition to interpret the runoff processes and pathways. The series of studies conducted by the McDonnell's laboratory demonstrated that the transit time plays a crucial role in testing the hydrological models and indicated the importance of geochemistry on hydrological modeling (Fenicia *et al.*, 2008; Sayama and McDonnell, 2009).

Although the abovementioned studies made a significant step forward on the choice and suitability of hydrological models, the accuracy of the simulated stream flow composition controlled by different model structures still needed further studies. Such result raises two interesting issues. 1. Why does the model prefer to provide such streamflow composition? 2. What kind of streamflow composition from the models is more realistic or reliable? Obviously, the model preference or model behavior is dominated by the model structure, governing equations, and calibration. Therefore, this study incorporated the same base-flow equation (linear reservoir concept) into different model structures (Hydrologiska Byrans Vattenbalans-avdelning (HBV) and TOPMODEL) to investigate the influence of the model structure on the simulations (Wagener *et al*., 2010). Understanding the two questions not only provide insight into hydrology but also water resource planning.

To investigate the two model behaviors, this study applied the hydrological models with the same base-flow component to a steep mountainous watershed in Taiwan. Altogether, 23 events covering a wide rainfall spectrum in hourly basis were used. The Nash efficiency coefficient (*Nash_EC*) and its logarithmic form ($Nash\_EC_{log}$) were used to calibrate parameters. The *Nash_EC*, volume bias ratio, and correlation coefficient were used to evaluate the model applicability. The simulated streamflow compositions were intercompared to investigate the model behaviors among the different events. Finally, the two rainstorms, supplemented by an intensive geochemical dataset, were independently introduced to assess and validate the simulated streamflow composition. This study improved our understanding on model selection and the role of the parameters in streamflow composition,

註解 **[R3]:** For technical comment #1

2. Materials and Methods

2-1 Study Area

This study chose the Chi-Chia-Wan watershed in central Taiwan as study area which is a typical forested, steep, and mountainous watershed with a drainage area of 105 km^2 (Fig. 1). The elevation varies from 1,131 m to 3,882 m above sea level, and the steep slope (average of approximately 33.3°) represents a high runoff velocity and sediment transport (Kao *et al.*, 2011; Lee *et al.*, 2013). The majority of soils are colluvial soils (including greyish yellow and dark greyish) and lithosols with high permeability. The soil depth is various due to the frequent mass movements, but most soil depths vary from 40 -120 cm (Soil and Water Conservation Bureau, 1985). The annual average air temperature is 15.8 °C, and the monthly average air temperatures in January and July are 4 and 23 °C in 2000–2009 (Huang *et al.*, 2006). The original and secondary forests covering nearly 87% of the area are the dominant land cover in this watershed. Most agricultural lands (e.g., orchard and vegetable farms) locate along the road or the riparian zone. The annual precipitation is as high as 2,551 mm (based on 2000–2011 data) with distinct seasonality. Approximately 75% of the annual precipitation rains during the wet season (May to October), and tropical cyclone (typhoon) is the main contributor. The annual evapotranspiration here is estimated between 600-1200 mm and the daily evapotranspiration in summer may be as high as 6-8 mm (Water Resources Agency, 2011). The annual discharge is approximately 2,129 mm (from 2000–2011 data) with a mean daily discharge of 7.09 m^3/s (equivalent

註解 **[R4]:** For detailed comment #1

註解 **[R5]:** For technical comment #1

註解 **[R6]:** For technical comment #5

註解 **[R7]:** For detailed comment #3

to 5.83mm/day).

For the rainstorm selection, a total of 23 rainstorms with significant water level rise (over 2 m) were selected to evaluate the model applicability (Table 1). In general, the average cumulative rainfall was approximately 430 mm within 102 h. The total rainfalls varied from 184.5 mm to 836.4 mm and the maximum rainfall intensity ranged from 10.7 mm/h to 39.5 mm/h. The total runoff depth ranged from 37.8 mm to 672.6 mm and was significantly positively correlated with the total rainfall and rainfall duration. The peak discharges ranged from 66.5 m^3/s to 510.4 m^3/s (equivalent to 2.28 mm/h to 17.5 mm/h) and were positively correlated with the total rainfall, average rainfall intensity, and maximum rainfall intensity (Table 1). The streamflow responding to the rainfall within a short time lag (generally less than 2 h) indicates that the watershed has a steep slope and short traveling time. These events, which crossed a wide spectrum in terms of the total rainfall and duration, are the critical factors to detect the limit of model applicability.

2-2 Hydrological Modeling: TOPMODEL and HBV

The two hydrological models [HBV (Hydrologiska Byråns Vattenbalans-avdelning) and TOPMODEL (hereafter, TOP)] were used in this study. Both models are regarded as conceptual distributed model and have been widely applied in many studies. Here, the two models were briefly introduced.

2-2.1 HBV Model

The HBV model was originally developed by the water balance section of the Swedish Meteorological and Hydrological Institute and has been modified into several versions (e.g., Bergström and Forsman, 1973; Bergström, 1992; Lindström *et al.*, 1997; Krysanova *et al.*, 1999; Haberlandt *et al.*, 2001; Blöschl *et al.*, 2008). Aghakouchak and Habib (2010) modified this model into a distributed-based model, and we used this version in the current study. The HBV model consists of four modules: (1) snowmelt and snow accumulation, (2) soil moisture and effective precipitation, (3) evapotranspiration, and (4) runoff response. For rainstorm (short-term) simulation, the hourly time step was used. In this study, the snow accumulation, snowmelt, and evapotranspiration modules were turned off. Although the elevation in the study site is high (> 3,000m), very little or no snow appears in subtropical summer. The amount of evapotranspiration during rainstorms (~3 days) should be less than

註解 **[R8]:** For detailed comment #2

註解 **[R9]:** For technical comment #1

註解 **[R10]:** For technical comment #6

註解 **[R11]:** For technical comment #1

24mm which is much smaller compared to the average of rainstorm precipitation (~ 430mm). Therefore, the two modules can be neglected to reduce the parameters involved in calibration.

註解 [R12]: For detailed comment #3

In the HBV model, precipitation is usually divided into two components: the first part contributes to the soil root zone, and the second one contributes to the interflow storage. The second component is usually known as effective precipitation. This component is estimated by an exponential coefficient and the saturation level in the soil root zone. In the soil root zone, saturation level is defined as the soil moisture over the field capacity (*FC*), the parameter that describes the maximum water storage. Adopting this concept, the higher the saturation is, the larger is the precipitation proportion recharged into the inter flow storage. Equation (1) describes the calculation of the effective precipitation, which is a function of the current soil moisture content.

$$P_{eff} = P \cdot \left(\frac{SM}{FC}\right)^{\beta} \qquad (1)$$

where P_{eff} is the effective precipitation [L], *SM* is the current soil moisture [L], *FC* is the maximum soil storage capacity [L], *P* is the hourly precipitation [L], and β is a model parameter (shape coefficient) [–]. The soil moisture status temporally evolves by receiving the rest of rainfall in each time step until reaching FC. An initial value of the soil moisture is required to start the calculations.

註解 [R13]: For detailed comment #4

The interflow and baseflow estimations at the watershed outlet are based on the linear reservoir concept (Fig. 2a). The reservoirs are directly connected to each other by a constant percolation rate (P_r). Two outlets (Q_s and Q_i) are in the upper reservoir and one outlet (Q_b) is in the lower reservoir. When the water level in the upper reservoir exceeds the threshold value (*L*), surface runoff (Q_s) occurs. For the surface flow routing, the unit response function implemented by the diffusive transport approach is used (Liu *et al.*, 2003; Huang *et al.*, 2009). The water level in the upper reservoir is used to generate the interflow (Q_i). The base-flow response in the lower reservoir is relatively slower and controlled by the water level in that reservoir. The recession coefficients K_s, K_i, and K_b, control the response functions of the three flows. The three recession coefficients and the percolation rates are all model parameters, which are estimated via calibration. Equations (2) and (3) illustrate the calculation of the three flows in the outlet:

註解 [R14]: For technical comment #1

$$Q_s = \begin{cases} K_s(S_i - L)A_c & if\, S_i > L \\ 0 & if\, S_i < L \end{cases} \qquad (2)$$

$$Q_i = A \cdot K_i \cdot S_i$$

$$Q_b = A \cdot K_b \cdot S_b \qquad (3)$$

where Q_s, Q_i, and Q_b represent the surface flow, interflow, and base flow [L^3/T], respectively. The parameters K_s, K_i, and K_b are the recession coefficients of the surface flow, interflow, and base flow [T^{-1}], respectively. S_i is the upper reservoir water level [L], S_b is the lower reservoir water level [L], and L is the threshold of water level [L]. A and A_c are the watershed and cell area [L^2], respectively (Fig. 2a).

2-2.2 TOPMODEL

TOPMODEL proposed by Beven and Kirkby (1979) has been applied widely around the world (Beven, 1996). The kernel feature of this model is to use the topographic index (defined as the contributing area over the gradient, see Equation 4) to estimate the variable source area and then simulate the discharge.

$$\gamma_i = \ln\left(\frac{a_i}{T_i \cdot tan\beta_i}\right) \qquad (4)$$

註解 **[R15]:** For technical comment #11

where γ_i is the local topographic index. Parameter T_i is the lateral transmissivity as the soil is saturated [L^2/T]. a is the specific contributing area defined as the drainage area per unit contour length [L], and tan β is the local gradient [-]. Because of its concise structure, numerous modifications have been introduced in the past three decades. We used the three-layer TOPMODEL (Huang *et al.*, 2009) in this study. This modification has been widely used in Taiwan for relevant hydrological applications either for hourly or for daily time step input (Huang *et al.*, 2011; Huang *et al.*, 2012). The conceptual scheme is shown in Fig 2(b). This model divides the soil column into three layers: upper, middle, and bottom layers, to simulate the surface flow, interflow and base flow, respectively (composing the stream discharge). In this model, the following 9 parameters need calibration: maximum root zone storage ,S_{rmax} [L], initial root zone storage, S_{r0} [L], Mannings' surface roughness, n [-], maximum draining capacity in the middle layer, T_d [L/T], lateral transmissivity, T_i [L^2/T], interflow recession coefficient, m_i [L], base-flow recession coefficient, K_b [T^{-1}], groundwater recharge or percolation, K_{per} [L/T] and bypass flow rate, Q_{by} [L/T].

For the upper layer, there are two ways to reduce the storage. One way is the evapotranspiration, which is turned off as mentioned before. The other way is the quick bypass flow (Q_{by}) from the upper layer to the bottom layer when the saturation exceeds 0.6. When the storage is fully filled by precipitation, the surplus rainfall infiltrates into the middle layer. However, the infiltrating water depends on the remaining space in the middle layer or on the maximum draining capacity T_d. Therefore, the saturation excess runoff can be described as the following:

$$\delta_i = \begin{cases} P_i - D_i & if\ T_d > D_i \\ P_i - T_d & if\ T_d \leq D_i \end{cases} \quad (5)$$

where P_i and D_i are the rainfall on i[th] cell [L] and the local soil moisture deficit [L], respectively. δ_i is the surplus rainfall [L], which transforms to the surface runoff. In this equation, D_i is only the condition to determine the rainfall converted to surface runoff. For the second equation in equation (5), the maximum drainage capacity, T_d, is the upper limit to avoid too much rainfall infiltrating into dryer or near-ridge cell so rapidly. The observational study in New Zealand revealed the larger contribution from new water in ridge top sites, indicating the possible generation of infiltration excess runoff (Sklash et al., 1986). Although the second equation is somehow similar to infiltration capacity based runoff, this equation mainly followed the concept of saturation excess runoff in most cases.

註解 **[R16]:** For detailed comment #5

For the middle layer, the local soil moisture deficit D_i can be estimated by

$$D_i = \bar{D} + m_i(\gamma - \ln(\frac{a_i}{T_i \cdot tan\beta_i})) \quad (6)$$

where $\bar{D}$ is the mean value of the soil moisture deficit over the catchment area. This equation uses the difference between the local topographic index and the average topographic index to estimate the possible local soil moisture deficit everywhere. Meanwhile, the subsurface flow for each time step can be estimated by the following recession curve function:

$$Q_i = Q_0 \cdot \exp(-\frac{\bar{D}}{m_i}) \quad (7)$$

where Q_i is the interflow [L^3/T] and $Q_0 = A \cdot \exp(-\gamma)$ is the discharge when the average soil moisture deficit is zero.

For the base flow, the same linear reservoir concept is applied to simulate the base flow as follows:

$$Q_b = S_b K_b \quad (8)$$

where Q_b is the base flows [L^3/T]. S_b [L] and K_b [L^2/T] are water level and the recession coefficient, respectively. The initial S_b can be derived from the initial observed discharge at time t = 0. The above three flows compose the stream discharge at the catchment outlet.

The two models show different model structures, particularly in the surface flow generation. The TOPMODEL generates the surface flow depending on the variable source area. The saturation in the middle layer is the key factor in generating surface flow. However, the HBV model separates the rainfall into root zone and inter-flow storage through the effective precipitation calculation, which is proportional to the soil moisture content and shape factor. In other words, the effective precipitation is the valve that controls the recharge into interflow storage before reaching full saturation in the root zone. In addition, the surface flow occurs only when the water level in the interflow storage is higher than the threshold *L*, which means that the surface flow in HBV is controlled by threshold *L*. Hence, the maximum interflow is somewhat limited. To set the common ground for the two models, we introduced the same base-flow governing equation to investigate the different model behaviors. The experiment design can aid us in understanding more about the model behaviors. Further, the sensitivities of the parameters were also evaluated to clarify the role of the allocation of the three flows. Hydrograph shapes, runoff volumes, and correlation coefficient were three measures used to discuss the model performance. Finally, the two rainstorms, supplemented by the intensive geochemical dataset for the stream flow compositions, were used to validate the simulated compositions.

2-3 Calibration and Performance Evaluation

In hydrological modeling, calibration is intensively used to determine the unknown and/or non-measurable parameters by ranking the performance measure between simulations and observations. However, many previous studies showed that no unique performance measure is better suited than another for the calibration of a model (Gupta *et al.*, 1998; Yapo *et al.*, 1998; Madsen, 2000; Vrugt *et al.*, 2003); therefore, multi-objective calibration has been proposed and applied widely. For the multi-objective calibration, the simulations laid on the Pareto front can be regarded as the best simulations, and the corresponding parameters are good candidates for further applications (e.g., parameter uncertainty estimation). Here, we use two performance measures for calibration. One measure is Nash efficiency coefficient, *Nash_EC* (proposed by Nash and Sutcliffe, 1970). This coefficient (Equation 9) varies from negative infinity to unity, where unity represents a perfect match and zero indicates that the simulation performance is identical to the expected value (mean) of the observations. However, this coefficient using the squared difference between simulation and observation leads to high sensitivity in the high flow. To consider the low-flow properly, a variant $Nash\_EC_{log}$, which transfers the simulated and observed discharges into a logarithmic scale, is applied as the other performance measure. In this study, over 80,000 parameter sets were generated by the uniform or log-uniform distribution for the two models. The detailed description of parameter range and used distribution was illustrated in Table 2. The best simulations and the corresponding parameter sets, defined by the highest values of the *Nash_EC* and $Nash\_EC_{log}$, are selected for further discussion.

$$Nash\_EC = 1 - \frac{\sum_{i=1}^{T}(Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^{T}(Q_{obs,i} - \overline{Q_{obs}})^2} \qquad (9)$$

where Q_{sim} and Q_{obs} are the simulated and observed discharges, respectively, and T is the total of time step during the evaluation period.

In addition to the two performance measures for calibration, we also used the following three indexes, namely, *Nash_EC*, *EQV*, and *CC*, to show the extent of the agreement between simulations and observations. *EQV* defines as the ratio of the total simulated volume over the total observed volume. This index is useful in investigating the volume bias which is important for irrigation, reservoir operation, and flood control. *CC* is the correlation coefficient between simulations and observations. Notably, a high *CC* with poor *EQV* indicates that the simulation has a highly similar shape with the observations but biases in the runoff volume. Based on the three indexes, the simulations in terms of hydrograph shape, volume, and correlation can be assessed comprehensively.

3. Results

After intensive simulations and calibration, the performances of *Nash_EC* and $Nash\_EC_{log}$ for the HBV- and TOP-derived simulations are shown in Fig. 3. The overall *Nash_EC* and $Nash\_EC_{log}$ values were scattered in an awl shape, and the maximum values met at approximately 0.65 in both axes for the two models. In general, both models could simulate the rainstorm fairly well showing a good agreement between the two model-derived simulations and observations. Notably, the pareto front may not exist when the performance measures are inherently similar or the simulation has the similar tradeoff weight between the performance measures. In this circumstance, all the simulations approach to a specific point. Therefore, we selected the best 15 simulations (the highest values of the sum of the two measures) as the representative simulations, and their corresponding parameter sets were regarded as the well-performed sets for each model (discussed later). The detailed simulation results are tabulated in Table 3.

註解 **[R17]:** For detailed comment #7

In the HBV-derived simulations, the *Nash_EC* values varied from 0.16 to 0.91 with a mean of 0.70. For the TOPMODEL, the *Nash_EC* values ranged from 0.10 to 0.89 with a mean of 0.64. The HBV-derived simulations were slightly better than the TOP-derived simulations. The standard deviations for the HBV- and TOP-derived models were 0.22 and 0.19, respectively, showing a similar level of variations among events. For *EQV*, the average performance of the two models was similar (0.92 and 0.93 for HBV and TOP, respectively). However, the range varied from 0.56 to 1.52 for HBV and from 0.68 to 1.20 for TOP, respectively, showing that TOPMODEL could make the simulated volume more consistent (less variation) with observations. For the correlation coefficient, the average

CC for HBV and TOP was 0.94 and 0.88, respectively, which indicated that the HBV simulations might give a higher correlation than the TOP simulations. The standard deviation of the TOP-derived *CC* was larger than that of the HBV-derived *CC*.

In summary, both model could simulated the streamflow in the similar performance level in terms of the hydrograph shape. TOP-derived simulation has the more consistent discharge volume than the HBV-derived simulation. However, HBV gave higher correlation coefficient than TOPMODEL. However, the outperformed parameter sets do not guarantee applicability for the all events (Huang *et al.*, 2009). Although the pursuit of higher performance measures (e.g., average *Nash_EC* and $Nash\_EC_{log}$ in this study) is the main consideration of calibration, pursuing the smaller variation in order to increase the applicability for all events (e.g., different hydrological conditions) should be emphasized as well.

4. Discussions

4-1 Well-performed simulation and corresponding parameter sets

The well-performed simulations and the performance in terms of *Nash_EC*, *EQV*, and *CC* are illustrated to reveal the variation in the simulations among the rainstorms (Fig. 4). In Fig. (4a.1) and 4(b.1), we found that the *Nash_EC* values of the 15 well-performed simulations for each event were quite diverse for the models, particularly in the small events. For the correlation coefficient, the HBV model presented good and consistent simulations for all events (Fig. 4a.2). The higher correlation coefficient values indicated that the simulations and observations all agreed well in terms of hydrograph shape. In contrast, the TOP-derived simulations for small events were highly divergent (Fig. 4b.2). For the runoff volume estimations, the HBV-derived simulations for small events were distinctly overestimated but were underestimated for large events (Fig. 4a.3). However, the TOP-derived simulations estimated the runoff volume better and remained more consistent compared to the HBV model (Fig. 4b.3).

This comparison showed the both models could not simulate the small events very well. It may be due to the fact that the spatial rainfall distributions of the small rainstorm are usually more heterogeneous than that of large events (Huang *et al.*, 2012). Nevertheless, HBV could outperform in correlation coefficient, but it significantly overestimated the discharge. In contrast, TOPMODEL showed a promising and consistent result in runoff volume, but failed in correlation coefficient. The diverse values for some small events in TOPMODEL may result from the surface runoff mechanism.

註解 **[R18]:** For detailed comment #9

註解 **[R19]:** For detailed comment #10

The surface runoff mostly generates in source area. Given a biased precipitation pattern on source area it would lead to a significant over- or underestimation in surface runoff rather than infiltration excess runoff (Huang *et al*., 2011). We cannot expect that the hydrological models can simulate such events well only based on the limited data. As for runoff volume estimation, the TOP-derived simulations maintained the water balance better than that derived by HBV. Taking a closer look at the model structures we found that the runoff estimation by HBV strongly depends on the storage status and the yield parameters (e.g., K_s, K_i, and K_b); therefore, it may not keep the water mass balance. In other words, HBV is more flexible in adjusting the simulated streamflow. In reality, many watersheds may not follow the mass balance, but it has been the basic assumption in many hydrological models. Therefore, the water mass balance assumption may need other environmental backgrounds to support.

註解 **[R20]:** For detailed comment #11

For the corresponding parameter sets, the retrieved parameter values were normalized to the upper and lower limits and linked to one another for showing the connectivity (Fig. 5). This figure shows that different parameter combinations could produce virtually equal model simulations. However, some parameters are constrained within a limited range indicating the parameters are more sensitive and dominant in simulation (Madsen, 2000; Madsen *et al.*, 2003). The pattern of parameter combination also represents the model behavior. For HBV, once the parameters, S_{ramx}, K_s, L, and K_b are fixed or determined; the similar simulations can be expected. Meanwhile, the similar parameter combinations show a similar model behavior in simulation. In contrast, only the parameter DE_i and K_{per} are sensitive and other parameters are diverse in TOPMODEL simulations. It seems that more than one type of parameter combinations can achieve similar performance which indicates that more than one type of streamflow compositions can be obtained. In this regard, a model which gives the more types of parameter combinations with similar performance has high flexibility.

註解 **[R21]:** For detailed comment #12

註解 **[R22]:** For technical comment #1

4-2 Comparison of HBV- and TOP-derived Stream Composition

The simulated three flows for all rainstorms are listed in Table 4. The proportions of the HBV-derived flows are 0.22, 0.29, and 0.49 for the surface flow, interflow, and base flow, respectively. By contrast, the TOP-derived flows for the surface flow, interflow, and base flow are 0.27, 0.50, and 0.23, respectively. Obviously, the base flow in HBV model plays a dominant role in simulating the streamflow; however, in TOPMODEL interflow is the major component for streamflow. Furthermore, the three flow proportion against average rainfall intensity and storm duration are shown in Figs. 6–8. Figure 6 shows that both simulated surface flow proportions increase from 0.1 to 0.5 with the increase in the average rainfall intensity from 2.0 mm/h to 11.0 mm/h. Meanwhile, both simulated surface flow proportions decrease from 0.5 to 0.1 with the increase in the storm duration from 40 h to 160 h. However, the HBV-derived surface flows among events show larger variation than those derived from TOPMODEL. Nevertheless, the consistent results in surface flow derived from the two model structures, even for extreme events, reveal that the surface flow proportions which, as expected, increase and decrease with the average rainfall intensities and storm durations, are fairly reliable and realistic.

For the interflow, the models showed discrepant relationships with the average rainfall intensity and storm duration (Fig. 7). The HBV model showed lesser interflow (~0.3) than TOPMODEL (~0.5). Notably, the two model behaviors had opposite responses to the storm duration. As the storm duration increased, the TOP-derived interflow increased from ~0.3 to 0.6. However, the HBV-derived interflow decreased from ~0.4 to 0.15. The opposite behaviors were due to the model structure. Theoretically, TOPMODEL simulates a larger interflow as using the decrease in the average soil deficit, which is also used to determine the variable source area. In our case, the maximum variable source area was approximately 30% to 65%. Therefore, TOPMODEL can be expected to give even larger interflow for more torrential rainstorms. By contrast, the HBV model simulates the interflow using the limited depth of L. When heavy rainfall exceeds L, surface runoff occurs and the inter-flow storage is reduced rapidly. Therefore, the inter flow proportion is relatively limited even when the storm duration increases.

For the base flow, the two models showed similar patterns with average rainfall intensity (Fig. 8). However, the HBV ranges from 0.7 to 0.1, which is much wider than that derived from TOP (from 0.4 to 0.1). It is an indication that for small events, the base flow is dominant in HBV, but the inter flow is important in TOPMODEL. Meanwhile, the HBV-derived base flow increases with the increase in the storm duration; however. TOP-derived base flow (avg = 0.23; std. = 0.098) is stable for all events. In the TOPMODEL structure, the middle layer yields interflow efficiently and thus the base flow remains unchanged. By contrast, the HBV model recharges more to the lower reservoir in order to fit the streamflow. Therefore, the base flow is compelled to increase, particularly during extreme rainstorms, which indicates that HBV may be more suitable for watersheds with thin soil layer. Likewise, TOPMODEL is expected to be preferable for watersheds with thick soil layer. In this regard, we could expect that the proper model choice should be based on the extensive spectrum of rainstorms and the extra environmental background, instead of intensive calibration. Meanwhile, such intercomparison between models also increased our understanding of the model structure and behavior.

4-3 Comparison with Chem-hydrograph

Furthermore, geochemical dataset was introduced to derive the streamflow composition through the end-member mixing analysis (EMMA). Lee *et al.* (2010; 2011; 2013) collected water samples in wells and soil columns for the end member of the base flow and interflow. Besides, we sampled the stream and rainwater at high frequency (~3 h interval) during event no. 15 and 17. The chem-hydrographs of the three components after the EMMA are shown in Figs. 9(a) and (b) for event nos. 15 and 17, respectively. It shows that the base flow was quite stable and only changed during the flood peak time. In general, the base flow occupied approximately 25% of the total runoff. The response of interflow surged and diminished rapidly. The interflow proportion was similar to the base flow. The remaining discharge was attributed to the surface flow. From the geochemical perspective, the surface flow is the most important component during the rainstorm period, which occupies approximately 40% to 50% of the total runoff volume.

EMMA is recognized as a useful analysis tool for hydrograph separation, although the number and selection of geochemical tracers are sometimes questionable (Barthold *et al.*, 2011; Carrera *et al.*, 2004). Despite the uncertain proportion of discharge components and the objective identification of the end members, the result of the stream composition, in terms of relative proportion, is relatively reliable. It substantially provides another perspective for stream flow composition. More importantly, the time-series changes of the flows should be realistic. In this regard, the EMMA-derived stream composition could be a good reference for comparison with the model-derived ones.

The EMMA- and the model-derived results are listed in Table 4. The HBV model simulation shows that the interflow and base flow are dominant components for event nos. 15 and 17, respectively. By contrast, TOPMODEL considers the interflow as the superior component; the surface flow is only secondary. No model yields the same EMMA-derived composition with regard to the proportion. From the quantity perspective, the TOP-derived surface flow shows a good agreement with that derived by EMMA. By contrast, the HBV-derived interflow is close to the EMMA-derived results, although the surface flow is underestimated.

The quantity and the response-time results are shown in Figs. 10 and 11 for event nos. 15 and 17, respectively. These two figures show that the HBV-simulated discharge is slightly underestimated, and TOPMODEL overestimated in the streamflow. However, TOPMODEL exhibits a good agreement in the recession segment for the two events. For the interflow, the two models produce fair results. HBV model simulates the base flow as a gentle dome. By contrast, the TOP-derived base flow shows a quick-response steep-bell shape. In the shape comparison, the TOPMODEL outperforms the HBV

Model. However, a significant time lag of approximately 2 h to 4 h is observed. In our case, the base flow responds with the streamflow simultaneously. The base flow could thus be considered a type of piston flow. In this regard, incorporating the piston flow theory into the hydrological models can improve the time lag, which aids in the interpretation of the base flow.

5. Summary

Many hydrological models can simulate the stream flow satisfactorily and plausibly. However, different runoff compositions can result in the similar streamflow. Therefore, recent attention has shifted to model structures to ensure the accuracy of inferences derived from modeling. In our study, HBV presented consistent parameter combinations; however, TOPMODEL achieved more parameter combinations, which implied that HBV preferred to give only one composition for simulated streamflow, but TOPMODEL could yield more. Rethinking is thus necessary to identify which model structure is better.

In the comparison of the simulated components, both simulated surface flows realistically reflect the nature. The simulated surface flows increased with the increasing the rainfall intensity and decreased with the increasing storm duration. Both base flows also showed the same patterns, although HBV-derived base flow was the dominant. However, the two modeled interflows exhibited a contrasting relationship with the storm duration. The HBV interflow decreased with the increase of duration. Because of the limited interflow storage, this model compelled to percolate much water to the base flow storage in order to fit the observed streamflow, which indicated that HBV could be more suitable for the thin-soil environment. On the other hand, TOPMODEL could be a better choice for catchments with thick soil. Compared with the EMMA-derived flows, a significant 2 h to 4 h time lag was observed, which indicated that the real base flow responses are faster than the models have presented. Possibly, an explicit consideration of the piston-flow characteristics in the base flow should be incorporated to improve the time lag and aid in the interpretation of the base flow.

Obviously, intercomparison between models under a wide spectrum of rainstorms is a good way to better understand the model behaviors. Besides, the independent geochemical data (e.g. EMMA-derived components) provides another perspective in examining the model behaviors. Undoubtedly, rejecting a model completely is difficult. Alternatively, it is very likely that more than one model structure is essential to capture the streamflow and tracer dynamics simultaneously when the rainstorm cases and environment background are insufficient. In this regard, we need to revisit the model behavior and the model structure again independently validation for testing hydrological models.

Acknowledgement

This study is supported by Taiwan National Science Council and National Taiwan University under grants NSC 101-2116-M-002-017, NSC 100-2621-M-002-027-MY3, NSC 101-2923-B-001-MY3, and NTU 10R70604-2. The authors also thank Taiwan Power Company and Shei-Pa National Park for providing the hydrological data and samples. Ms. Lee, Li-Chin’s work in drawing figures is appreciated.

Reference

Aghakouchak, A., Habib, E. (2010) Application of a Conceptual Hydrologic Model in Teaching Hydrologic Processes, *International Journal of Engineering Education*, 26(4), 963-973.

Barthold, F.K., Tyralla, C., Schneider, K., Vachè, K.B., Frede, H.-G., Breuer, L. (2011) How many tracers do we need for end member mixing analysis (EMMA)? A sensitivity analysis, *Water Resources Research*, 47, W08519, doi:10.1029/2011WR010604.

Bergström, S. (1992) *The HBV model — its structure and applications*. SMHI RH no. 4, Norrköping, Sweden.

Bergström, S., Forsman, A. (1973) Development of a conceptual deterministic rainfall-runoff model *Nordic Hydrology*, 4, 147–170.

Beven, K. J., and M. J. Kirkby (1979) A physically based variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24(1), 43–69.

Beven, K. (1996) The limits of splitting: Hydrology, *Science of The Total Environment*, 183, 89-97.

Beven, K. (2001) Revisiting the Problem of Model Choice. In Beven K. (ed.) *Rainfall-Runoff Mofelling The Primer*. John Wiley and Sons, New York, pp.297-304.

Beven, K. J. and Freer, J. (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, *Journal of Hydrology,* 249, 11–29.

Blöschl, G., Reszler, C., Komma, J. (2008) A spatially distributed flash flood forecasting model, *Environmental Modelling and Software*, 23(4), 464–478

Burns, D.A., McDonnell, J.J., Hooper, R.P., Peters, N.E., Freer, J.E., Kendall, C., Beven, K. (2001) Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research watershed (Georgia, USA), *Hydrological Processes*, 15, 1903-1924.

Carrera, J., Vázquez-Suñé, E., Castillo, O., Sánchez-Vila, X. (2004) A methodology to compute mixing ratios with uncertain end-members, *Water Resources Research*, 40, W12101, doi:10.1029/2003WR002263.

Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T., Hay, L.E. (2008) Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resources Research*, 44, W00B02, doi:10.1029/2007WR006735.

Fenicia, F., McDonnell, J.J., Savenije, H.H.G. (2008) Learning form model improvement: On the contribution of complementary data to process understanding, *Water Resources Research*, 44, W0619, doi:10.1029/2007WR006386.

Gupta, H. V., Sorooshian, S. and Yapo, P. O. (1998) Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information, *Water Resources Research,* 34(4), 751-763.

Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowita, G., Ye, M. (2012) Towards a comprehensive assessment of model structural adequacy, *Water Resources Research*, 48, W08301, doi:10.1029/2011WR011044.

Haberlandt, U., Klocking, B., Krysanova, V., Becker, A. (2001) Regionalisation of the base flow index from dynamically simulated flow components - a case study in the Elbe River Basin, *Journal of Hydrology* , 248, 35-53.

Huang, J.C., Kao, S.J., Hsu, M.L., Lin, J.C. (2006) Stochastic procedure to extract and to integrate landslide susceptibility maps: an example of mountainous watershed in Taiwan, *Natural Hazards and Earth System Sciences,* 6(5), 803-815

Huang, J.C., Lee, T.Y., Kao, S.J. (2009) Simulating typhoon-induced storm hydrographs in subtropical mountainous watershed: An integrated 3-layer TOPMODEL, *Hydrology and Earth System Sciences*, 13, 27-40.

Huang, J.C., Kao, S.J., Lin, C.Y., Chang, P.L., Lee, T.Y., Li, M.H. (2011) Effect of Subsampling Tropical Cyclone Rainfall on Flood Hydrograph Response in a Subtropical Mountainous Catchment, *Journal of Hydrology*, doi: 10.1016/j.jhydrol.2011.08.037.

Huang, J.C., C.K. Yu, J.Y. Lee, L.W. Cheng, T.Y. Lee, S.J. Kao (2012), Linking typhoon tracks and spatialrainfall patterns for improvi ng flood lead time predictions over a mesoscale mountainous watershed, *Water Resources Research*, 48, W09540, doi:10.1029/2011WR011508.

Johnson, M.S., Coon, W.F., Mehta, V.K., Steenhuis, T.S., Brooks, E.S., Boll, J. (2003) Application of two hydrologic models with different runoff mechanisms to a hillslope dominated watershed in the northeastern US: a comparison of HSPF and SMR, *Journal of Hydrology*, 284, 57-76.

Katsuyama, M., Ohte, N., Kobashi, S. (2001) A three-component end-member analysis of streamwater hydrochemistry in a small Japan forested headwater catchment, *Hydrological Processes*, 15, 249-260.

Krysanova, V., Bronstert, A., Müller-Wohlfeil, D.-I. (1999) Modelling river discharge for large drainage basins: from lumped to distributed approach, *Hydrological Sciences Journal*, 44(2), 313–331

Lin, Y.-F., Lin, C.-Y., Chou, W.-C., Lin, W.-T., Tsai, J.-S., Wu, C.-F. (2004) Modeling of riparian vegetated buffer strip width and placement A case study in Shei Pa National Park, Taiwan, *Ecological Engineering*, 23, 327-339.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S. (1997) Development and test of the distributed HBV-96 hydrological model, *Journal of Hydrology*, 201, 272–288.

Lee, T.Y. (2010) Study on Integration of Monitoring and Modeling for Analyzing the Effects of Land Use Changes on Water Quality and Stream Temperature in Chichiwan Creek. Doctoral Dissertation, Graduate Institute of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan.

Lee, T.Y., Huang, J. C., Kao, S. J. (2011) Typhoon hydrograph separation in the mountainous watershed of Taiwan. Asia Oceania Geosciences Society 2011 8th Annual Meeting, August, Taipei, Taiwan.

Lee, T.Y., Huang, J.C., Kao, S.J., Tung, C.P. (2013) Temporal variation of nitrate and phosphate transport in headwater catchments: the hydrological controls and land use alteration, Biogeosciences, 10, 2617-2632, doi:10.5194/bg-10-2617-2013.

Liu, Y.B., Gebremeskel, S., De Smedt, F., Hoffman, L., Pfister, L. (2003) A diffusive approach for flow routing in GIS based flood modeling, *Journal of Hydrology,*. 283, 91–106.

Madsen, H. (2000) Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *Journal of Hydrology*, 235, 276-288.

Madsen, H. (2003) Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, 26(2), 205-216.

McMillan, H., Tetzlaff, D., Clark, M., Soulsby, C. (2012) Do time-variable tracers aid the evaluation of hydrological model structure? A multimodel approach, *Water Resources Research*, 48, W05501, doi:10.1029/2011WR011688.

Michaud, J., Sorooshian, S. (1994) Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed, *Water Resources Research*, 30(3), 593-605.

Nash, J. E., Sutcliffe, J. V. (1970) River flow forecasting through conceptual models 1. A discussion of principles, *Journal of Hydrology,* 10, 282–290.

Perrin, C., Michel, C., Andreassian, V. (2001) Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242, 275–301.

Pokhrel, P., Gupta, H.V. (2011) On the ability to infer spatial catchment variability using streamflow hydrographs. *Water Resources Research*, 47, W08534, doi:10.1029/2010WR009873.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., DMIP participants, 2004. Overall distributed model intercomparison project results. *Journal of Hydrology*, 298(1-4), 27-60.

Refsgaard, J.C., Henriksen, H.J. (2004) Modelling guidelines – terminology and guiding principles, *Advances in Water Resources*, 27, 71-82.

Sayama, T., McDonnell, J.J. (2009) A new time-space accounting scheme to predict stream water residence time and hydrograph source components at the watershed scale, *Water Resources Research*, 45, W07401, doi:10.1029/2008WR007549.

Sklash, M.G., Stewrt, M.K., Pearce, A.J. (1986) Storm Runoff Generation in Humid Headwater Catchments 2. A Case Study of Hillslope and Low-Order Stream Response, *Water Resources Research*, 22(8): 1273-1282.

Soil and Water Conservation Bureau (1985) Hillslope Soil Survey in Taichung County, pp. 154, Republic of China Min. Econ. Affairs, Taipei, Taiwan (in Chinese)

Sorooshian, S., Dracup, J.A. (1980) Stochastic Parameter Estimation Procedures for Hydrologic Rainfall-Runoff Models: Correlated and Heteroscedastic Error Cases, *Water Resources Research*, 16(2), 430-442.

Valeo, C., Moin, S.M.A. (2001) Hortaonian and variable source area modeling in urbanizing basins, *Journal of Hydrologic Engineering*, **6**(4), 328-335.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., Sorooshian, S. (2003) Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resources Research,39(8),1214.*

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., Sorooshian, S. (2003) A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research,39(8),1201.*

Wagener, T., Sivapalan, M., Troch, P.A., McGlynn, B.L., Harman, C.J., Gupta, H.V., Kumar, P., Rao, P.S.C., Basu, N.B., Wilson, J.S. (2010) The future of hydrology: An evolving science for a changing world, *Water Resources Research*, 46, W05301, doi:10/1029/2009WR008906.

Water Resources Agency (2011) Hydrological Year Book of Taiwan, 2011. Republic of China Min. Econ. Affairs, Taipei, Taiwan (in Chinese).

Weiler, M., McGlynn, B.L., McGuire, K.J., McDonnell, J.J, (2003) How does rainfall become runoff? A combined tracer and runoff transfer function approach. *Water Resources Research*, 39(11), 1315, doi:10.1029/2003WR002331.

Winchell, M., Gupta, H.V., Sorooshian, S. (1998) On the simulation of infiltration- and saturation-excess runoff using radar-based rainfall estimates: Effects of algorithm uncertainty and pixel aggregation, *Water Resources Research*, 34(10): 2655-2670.

Yapo, P.O., Gupta, H.V., Sorooshian, S. (1998) Multi-objective global optimization for hydrologic models. *Journal of Hydrology*, 204, 83-97.

Table 1 the rainstorm characteristics in Chi-Chia-Wan catchment since 1986

Event no.	Date	Duration (hr)	Rainfall (mm)	Max. RI*[2] (mm/hr)	Runoff volume (mm)	Peak flow (mm/hr)	RC*[3]
1	1986/09/18	108	316.0	26.5	140.3	5.4	0.44
2	1986/08/22	95	247.2	15.6	69.9	2.9	0.28
3	1989/09/10	120	595.3	38.4	330.7	13.4	0.56
4	1990/09/07	96	454.9	28.6	237.6	7.9	0.52
5	1990/08/18	121	425.6	24.9	245.5	6.9	0.58
6	1990/06/22	105	359.7	28.8	132.2	3.7	0.37
7	1996/07/30	110	451.1	27.3	363.8	13.5	0.81
8	1997/08/28	90	228.5	17.5	120.8	5.2	0.53
9	1998/10/15	120	273.6	23.3	128.6	2.8	0.47
10	2000/08/22	94	398.8	25.7	107.3	3.5	0.27
11	2004/08/23	80	452.9	25.9	351.1	17.5	0.78
12	2004/07/02	96	431.3	35.3	112.3	4.5	0.26
13	2005/08/31	40	426.9	39.5	198.3	17.4	0.46
14	2006/06/08	144	409.4	20.3	247.9	5.5	0.61
15*[1]	2007/08/17	87	490.2	38.5	334.3	13.9	0.68
16	2007/09/17	97	184.5	15.0	75.8	2.3	0.41
17*[1]	2007/10/05	120	629.7	35.4	403.2	15.4	0.64
18	2008/07/17	90	200.0	21.7	93.0	2.6	0.47
19	2008/09/12	144	836.4	10.7	672.6	11.6	0.80
20	2008/09/27	91	672.9	33.3	483.4	16.5	0.72
21	2009/08/06	154	829.4	22.0	622.6	11.6	0.75
22	2009/10/05	72	220.5	14.9	37.8	2.0	0.17
23	2010/9/19	72	253.1	28.9	103.7	4.9	0.41
	Average	102	430.0	26.1	245.1	8.4	0.52

*[1] meant the events had the chem-hydrographs for validation

*[2] Max RI was the maximum rainfall intensity during the event

*[3] RC, runoff coefficient indicated the total runoff over the total rainfall

Table 2 the descriptions, ranges, and distributions of parameters for HBV and TOPMODEL

註解 [R23]: For detailed comment #6

HBV				TOPMODEL			
Parameter	Unit	Range	Distribution	Parameter	Unit	Range	Distribution
n	[-]	0.0 – 1.0	Uniform	n	[-]	0.0 – 1.0	Uniform
S_{rmax}	[L]	0.0 – 35.0	Uniform	S_{rmax}	[L]	0.0 – 35.0	Uniform
S_{r0}	[L]	0.0 – 35.0	Uniform	S_{r0}	[L]	0.0 – 35.0	Uniform
Beta	[-]	1.0 – 10.0	Uniform	T_d	[L/T]	10.0 –	Uniform
L	[L]	-3.0 – -1.0	Log uniform	m_i	[L]	0.0 – 2.0	Log uniform
K_s	[T^{-1}]	0.0 – 1.0	Uniform	T_i	[L^2/T]	1.0 – 10.0	Uniform
K_i	[T^{-1}]	-3.0 – -1.0	Log uniform	K_b	[T^{-1}]	-3.0 – -1.0	Log uniform
K_b	[T^{-1}]	-3.0 – -1.0	Log uniform	K_{per}	[L/T]	-3.0 – -1.0	Log uniform
K_{per}	[L/T]	-2.0 – 1.0	Log uniform	Q_{by}	[L/T]	-3.0 – -1.0	Log uniform
Q_{by}	[L/T]	0.0 - 13.0	Uniform				

Table 3 the HBV- and TOP-derived simulations evaluated by *Nash_EC*, *EQV*, and *CC*

Event	HBV Model			TOP Model		
no.	*Nash EC*	*EQV*	*CC*	*Nash EC*	*EQV*	*CC*
1	0.87	0.91	0.96	0.86	0.91	0.94
2	0.72	1.37	0.97	0.68	0.81	0.86
3	0.80	0.72	0.97	0.89	0.84	0.95
4	0.16	0.56	0.96	0.79	0.85	0.93
5	0.25	0.56	0.97	0.78	0.76	0.97
6	0.84	1.07	0.97	0.67	1.11	0.93
7	0.81	0.73	0.95	0.75	0.78	0.92
8	0.77	0.86	0.92	0.47	0.87	0.74
9	0.86	0.92	0.96	0.43	0.83	0.73
10	0.91	0.97	0.97	0.61	0.84	0.82
11	0.84	0.89	0.95	0.88	1.05	0.97
12	0.66	1.11	0.87	0.54	1.06	0.89
13	0.69	0.98	0.84	0.70	1.17	0.89
14	0.56	0.79	0.92	0.78	1.00	0.92
15*	0.82	0.78	0.96	0.61	1.00	0.94
16	0.75	1.14	0.93	0.60	0.94	0.83
17*	0.83	0.82	0.97	0.81	1.15	0.98
18	0.85	1.06	0.95	0.10	0.68	0.63
19	0.43	0.64	0.92	0.59	0.95	0.90
20	0.89	0.86	0.96	0.40	1.20	0.94
21	0.60	0.70	0.92	0.69	0.88	0.86
22	0.30	1.52	0.95	0.62	0.82	0.85
23	0.78	1.14	0.97	0.38	0.80	0.75
Average	0.70	0.92	0.94	0.64	0.93	0.88
Std.	0.22	0.24	0.03	0.19	0.14	0.09

*Note that the value for each event is the average of the representative simulations.

Table 4 The proportion of the simulated surface-, inter-, and base-flows derived from the two models

Event no.	HBV Model derived			TOP Model derived		
	Surface flow	Inter flow	Base flow	Surface flow	Inter flow	Base flow
1	0.26	0.31	0.43	0.30	0.51	0.19
2	0.17	0.32	0.51	0.26	0.46	0.28
3	0.27	0.31	0.41	0.33	0.50	0.16
4	0.17	0.23	0.60	0.21	0.56	0.23
5	0.15	0.22	0.63	0.16	0.57	0.27
6	0.18	0.27	0.55	0.18	0.56	0.26
7	0.34	0.35	0.31	0.44	0.44	0.13
8	0.14	0.26	0.60	0.10	0.58	0.32
9	0.11	0.23	0.66	0.10	0.55	0.35
10	0.26	0.32	0.42	0.32	0.49	0.18
11	0.41	0.39	0.20	0.51	0.41	0.07
12	0.18	0.30	0.52	0.24	0.56	0.20
13	0.43	0.40	0.17	0.53	0.39	0.08
14	0.07	0.16	0.77	0.09	0.54	0.37
15*	0.31	0.42	0.27	0.39	0.46	0.16
16	0.09	0.24	0.67	0.09	0.52	0.40
17*	0.29	0.31	0.40	0.30	0.55	0.15
18	0.11	0.25	0.64	0.16	0.46	0.38
19	0.25	0.27	0.48	0.27	0.58	0.15
20	0.36	0.34	0.31	0.40	0.44	0.16
21	0.20	0.25	0.55	0.22	0.61	0.17
22	0.13	0.29	0.58	0.27	0.38	0.36
23	0.26	0.35	0.39	0.45	0.30	0.24
Average	0.22	0.29	0.49	0.27	0.50	0.23
Std.	0.10	0.06	0.16	0.13	0.08	0.10

Table 5 Stream discharge composition derived from two models and EMMA

Event	Flow type	Chem-hydrograph		HBV Model		TOP Model	
		Amount (m^3/sec)	Proportion (%)	Amount (m^3/sec)	Proportion (%)	Amount (m^3/sec)	Proportion (%)
No. 15	Surface-flow	3949	40.5	2428	31.1	3711	38.5
	Inter-flow	2905	29.8	3305	42.3	1689	45.5
	Base-flow	2896	29.7	2076	26.6	1537	16.0
No. 17	Surface-flow	5562	47.3	2690	29.2	4116	30.4
	Inter-flow	2928	24.9	2837	30.8	7460	55.1
	Base-flow	3269	27.8	3685	40.0	1963	14.5

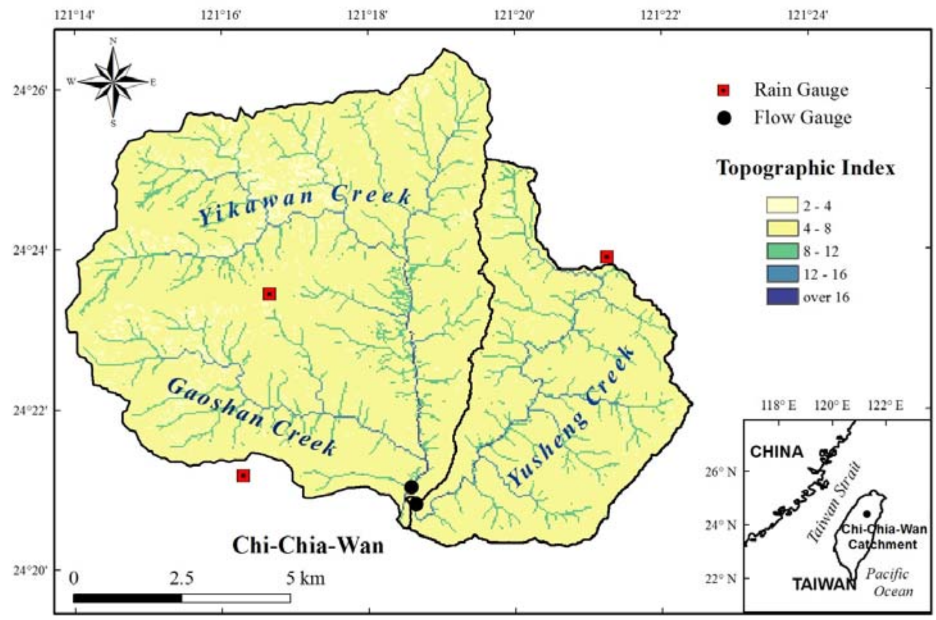

Figure 1. The landscape, stream network and topographic index pattern within the Chi-Chia-Wan catchment. The raingages and flow stations are labeled by red square and black dot.

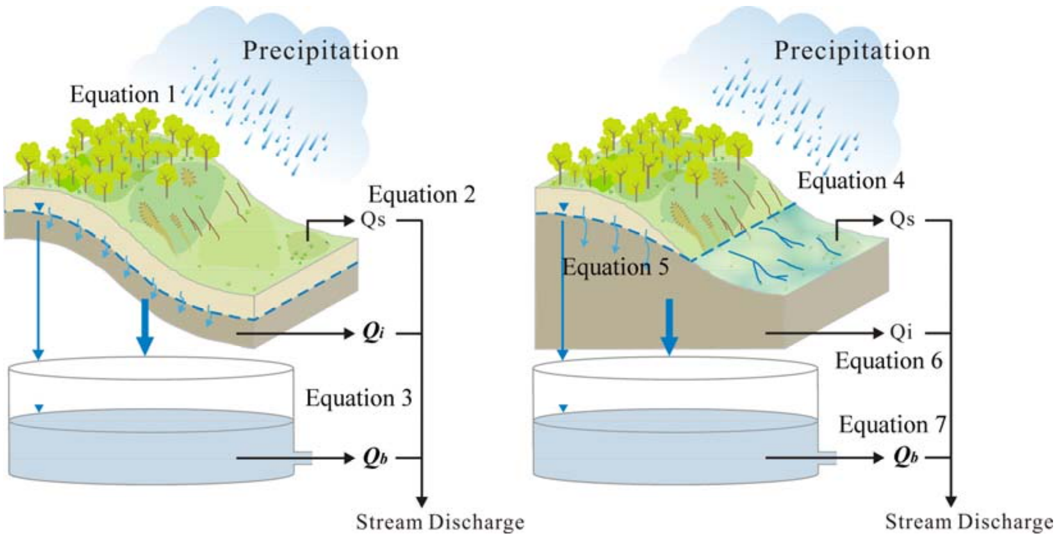

Figure. 2 The conceptual diagrams of HBV-Model (left) and TOP-Model (right). The bold and italic symbols mean the flow is modeled by linear reservoir

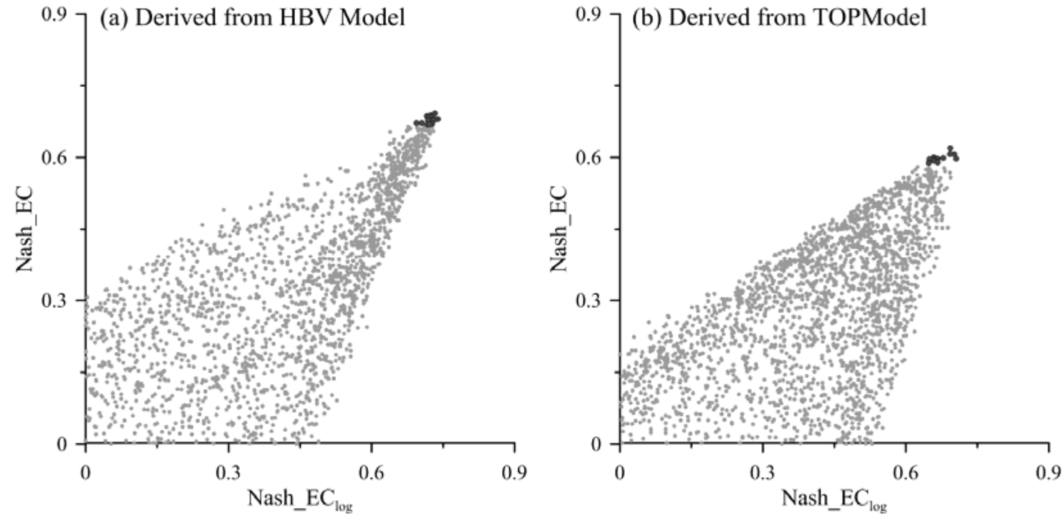

Figure. 3 The Nash_EC and Nash_EC_{log} values corresponding to the generated parameter sets in calibration process. The HBV-derived and TOP-derived results are shown in left and right panel, respectively. The *x*-axis and *y*-axis represent the Nash_EC_{log}, and Nash_EC, respectively. The gray circles represent the simulations partly and the black ones are the best 15 simulations sorted out by the equally weighted ranking.

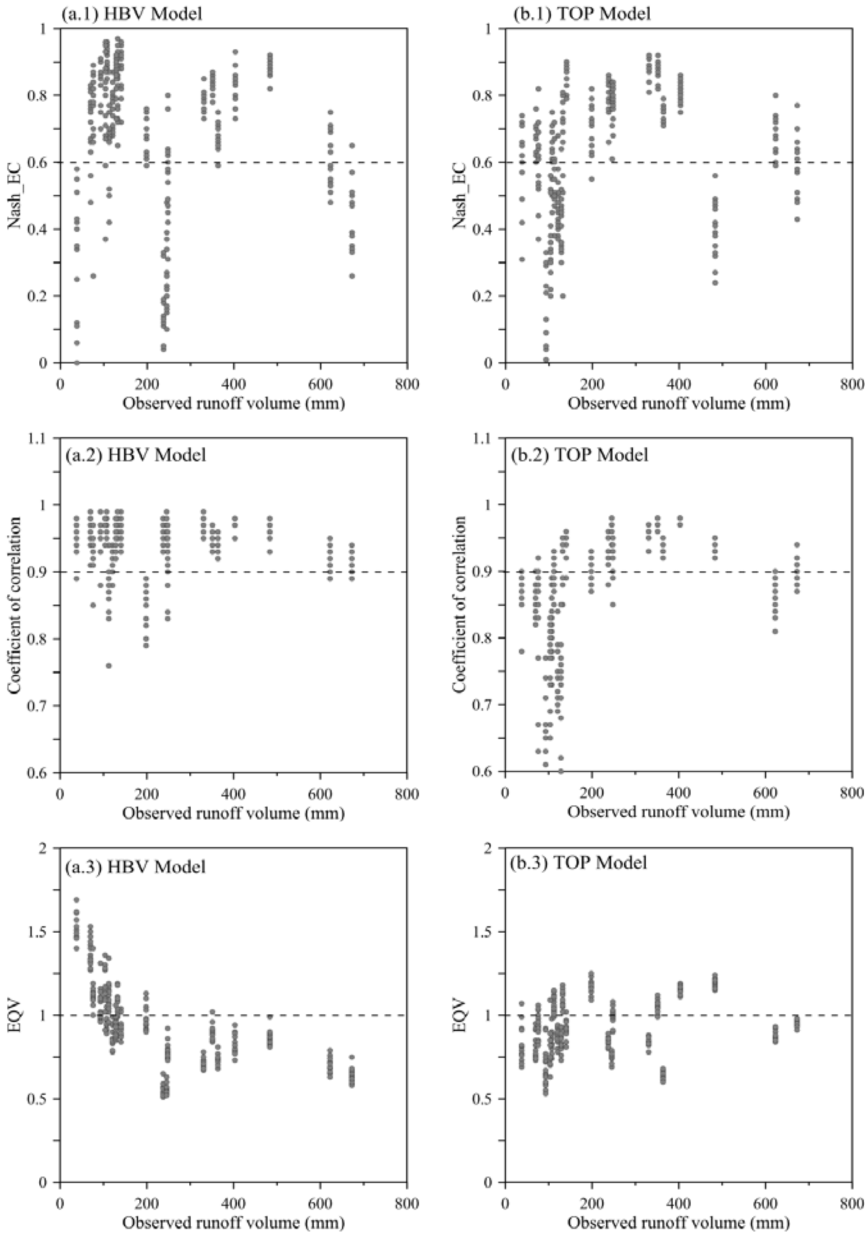

Figure 4. The performances of the best HBV- (a) and TOP-derived simulations (b), respectively, against the rainstorm magnitude in terms of observed runoff volume. The performance measure of Nash_EC, coefficient of correlation, and volume bias are shown in the upper, middle, and bottom panels, respectively. The dash lines are shown for reference.

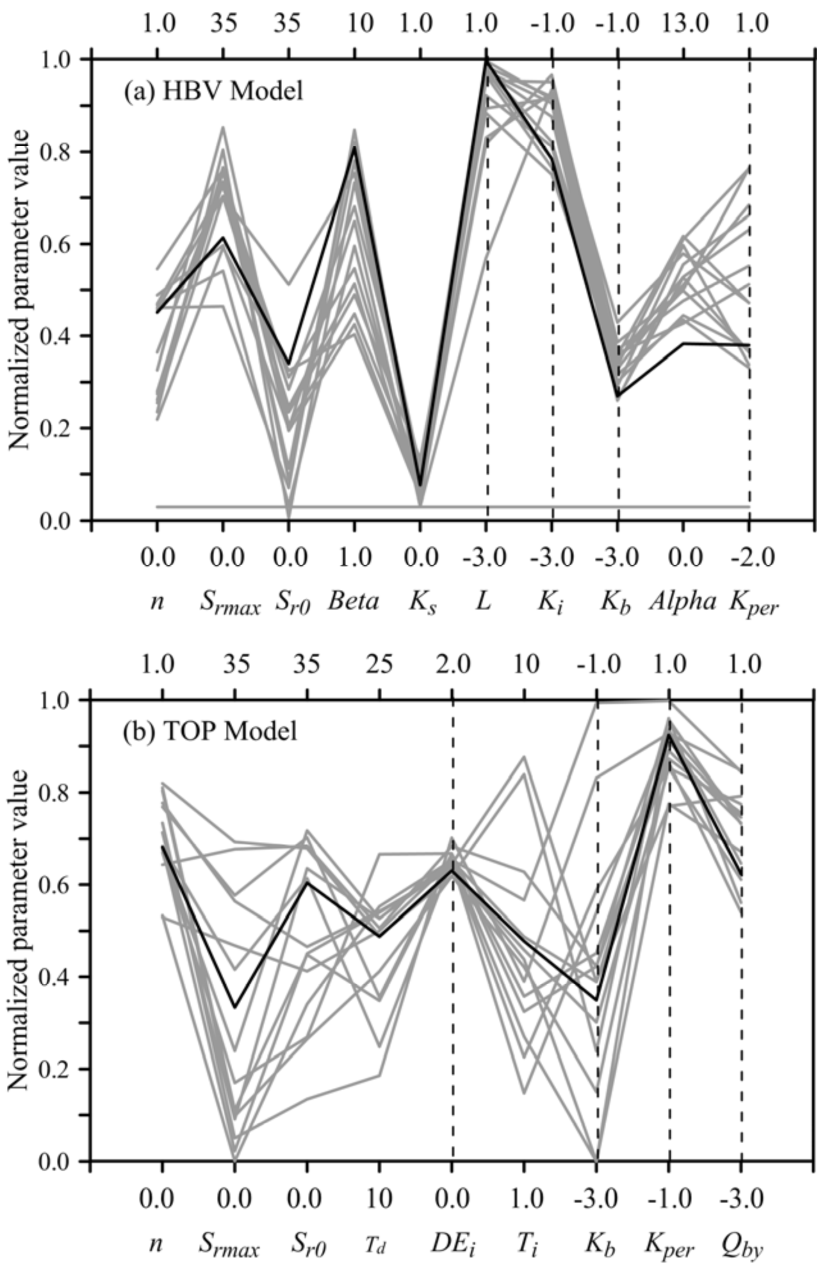

Figure 5. Normalized range of parameter values of the 15 best simulations (gray lines) for HBV Model (a) and TOP Model (b). The vertical dashed lines represent the parameter in logarithmic scale. The black lines indicate the best one for the two models.

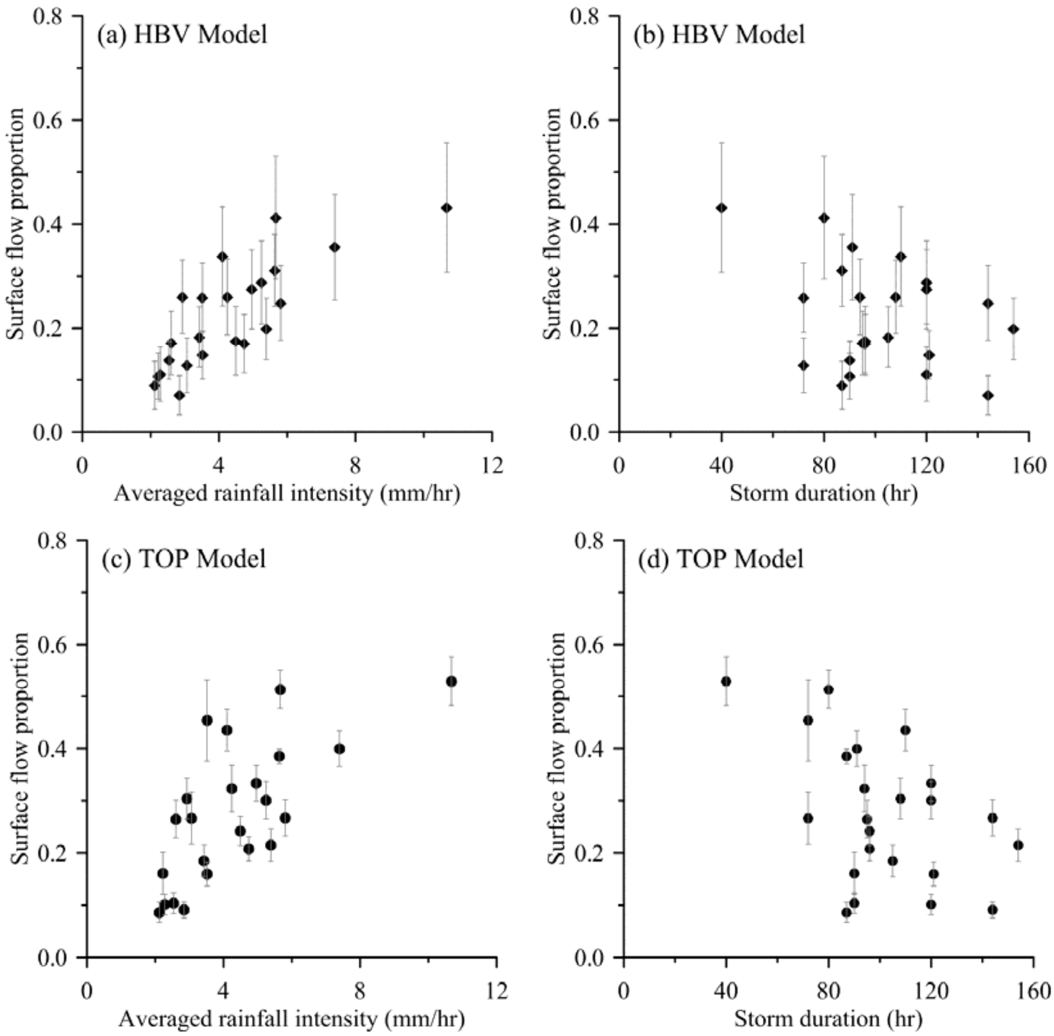

Figure 6. The variation of HBV-derived surface flow against averaged rainfall intensity (a) and storm duration (b). The variation of TOP-derived surface flow against averaged rainfall intensity (c) and storm duration (d). The black dot and gray line represent the mean and the standard deviation among the best simulations.

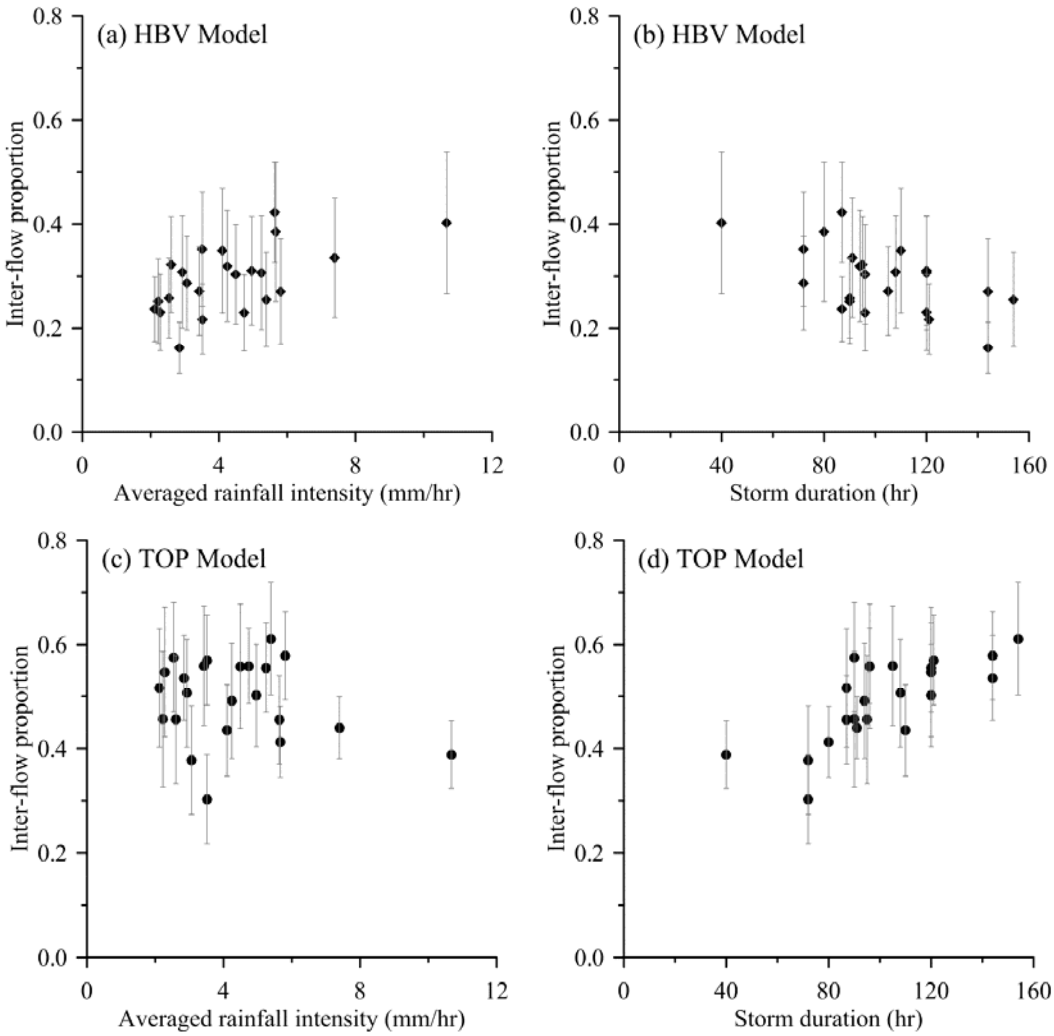

Figure 7. The variation of HBV-derived inter-flow against averaged rainfall intensity (a) and storm duration (b). The variation of TOP-derived inter-flow against averaged rainfall intensity (c) and storm duration (d). The black dot and gray line represent the mean and the standard deviation among the best simulations.

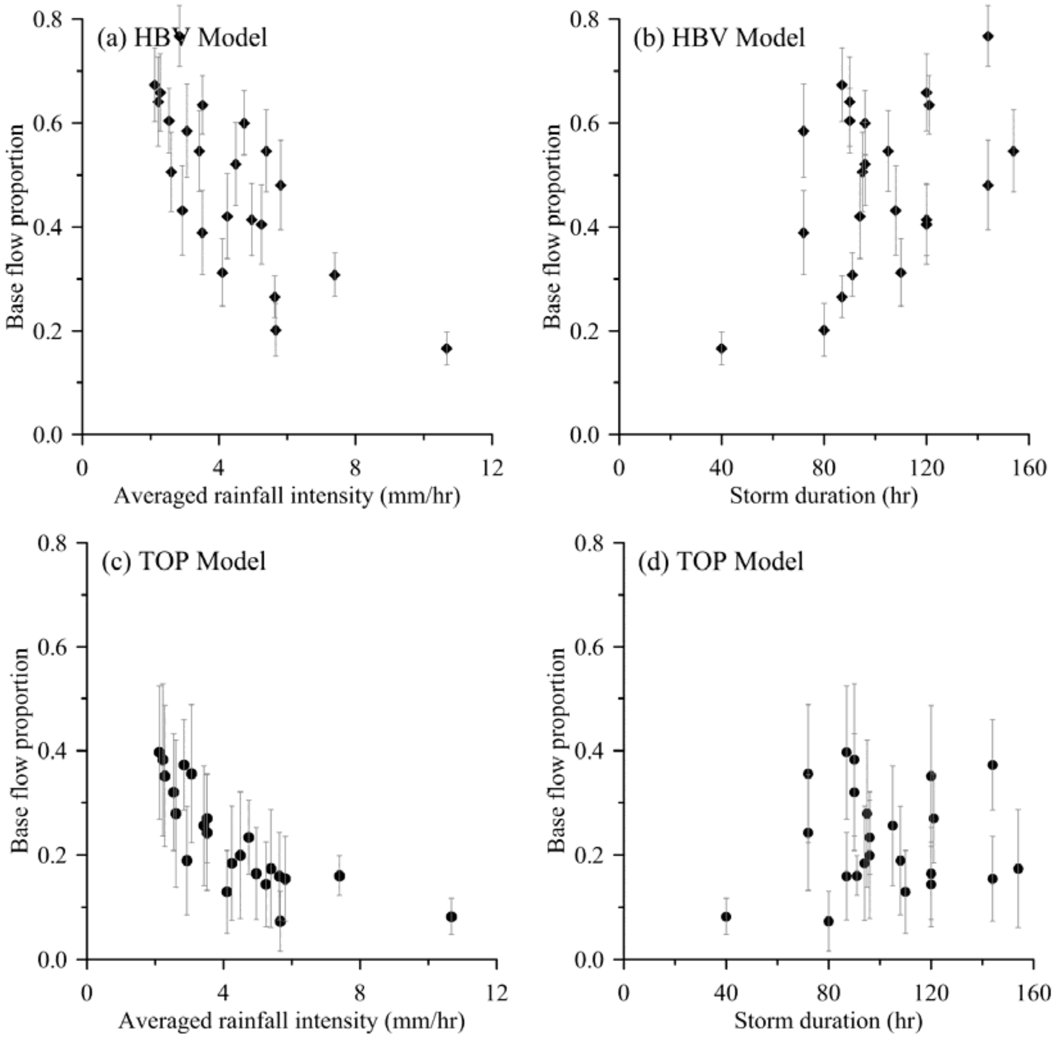

Figure 8. The variation of HBV-derived base-flow against averaged rainfall intensity (a) and storm duration (b). The variation of TOP-derived base-flow against averaged rainfall intensity (c) and storm duration (d). The black dot and gray line represent the mean and the standard deviation among the best simulations.

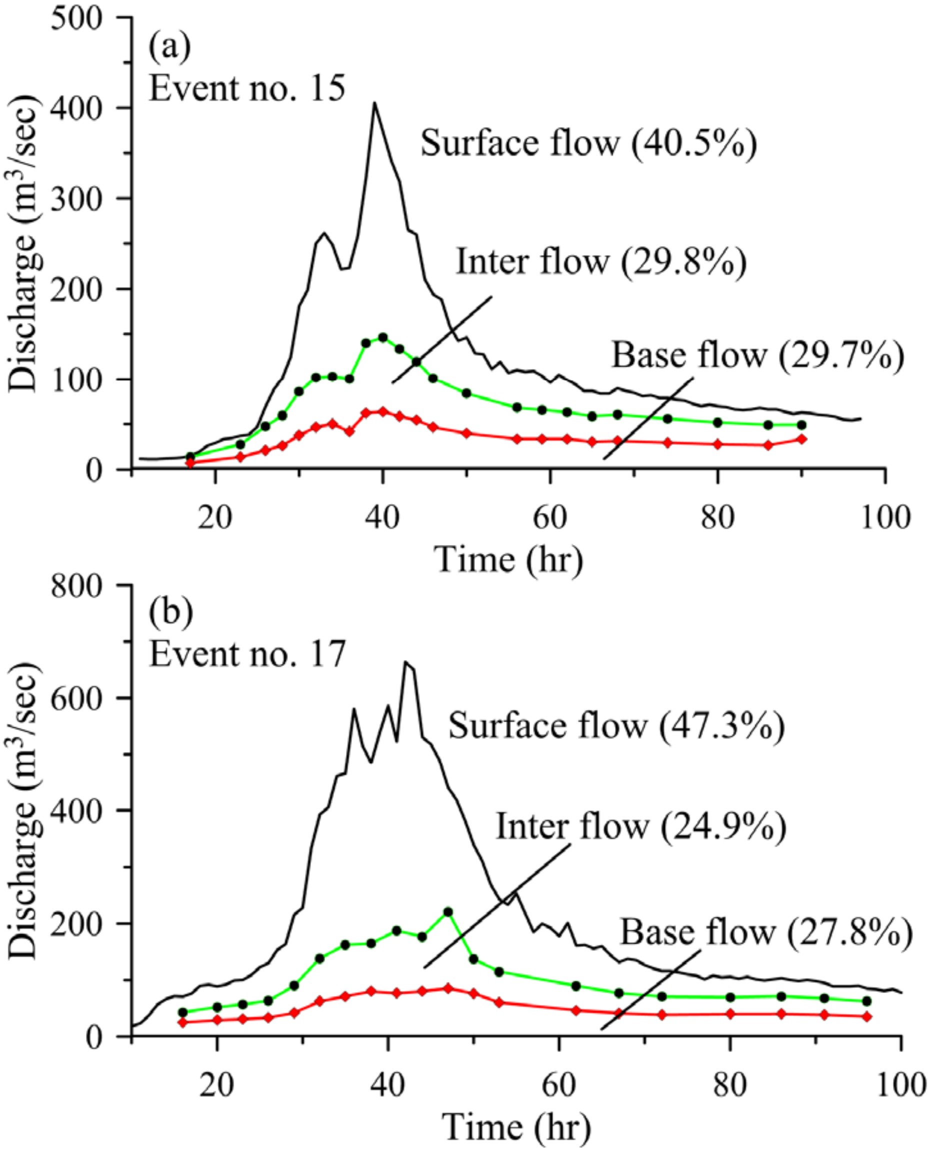

Figure 9. The EMMA-estimated three discharge components of event no. 15 and no. 17 are shown in (a) and (b), respectively. The black lines represented the observed stream discharge. The green and red lines indicate the estimated inter- and base-flow derived from EMMA. (seeing text in section 4-3 for details)

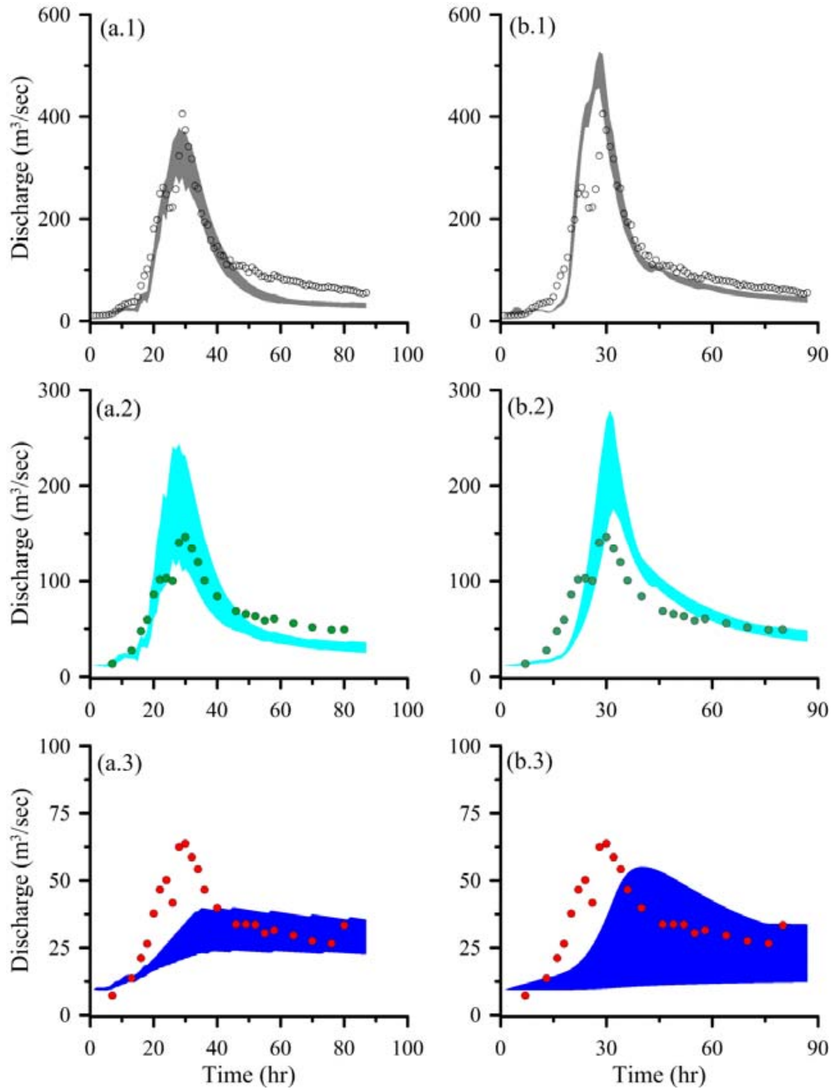

Figure 10. Comparison between the measured stream discharges (event no. 15) and the best 15 simulations derived from HBV model (a.1) and TOP model (b.1). The comparison of interflow derived from mixing analysis (green dots) with the simulated inter-flows (sky blue zone) derived from HBV model (a.2) and TOP model (b.2), respectively. The comparison of base flow derived from mixing analysis (red dots) with the simulated inter-flows (blue zone) derived from HBV model (a.3) and TOP model (b.3), respectively.

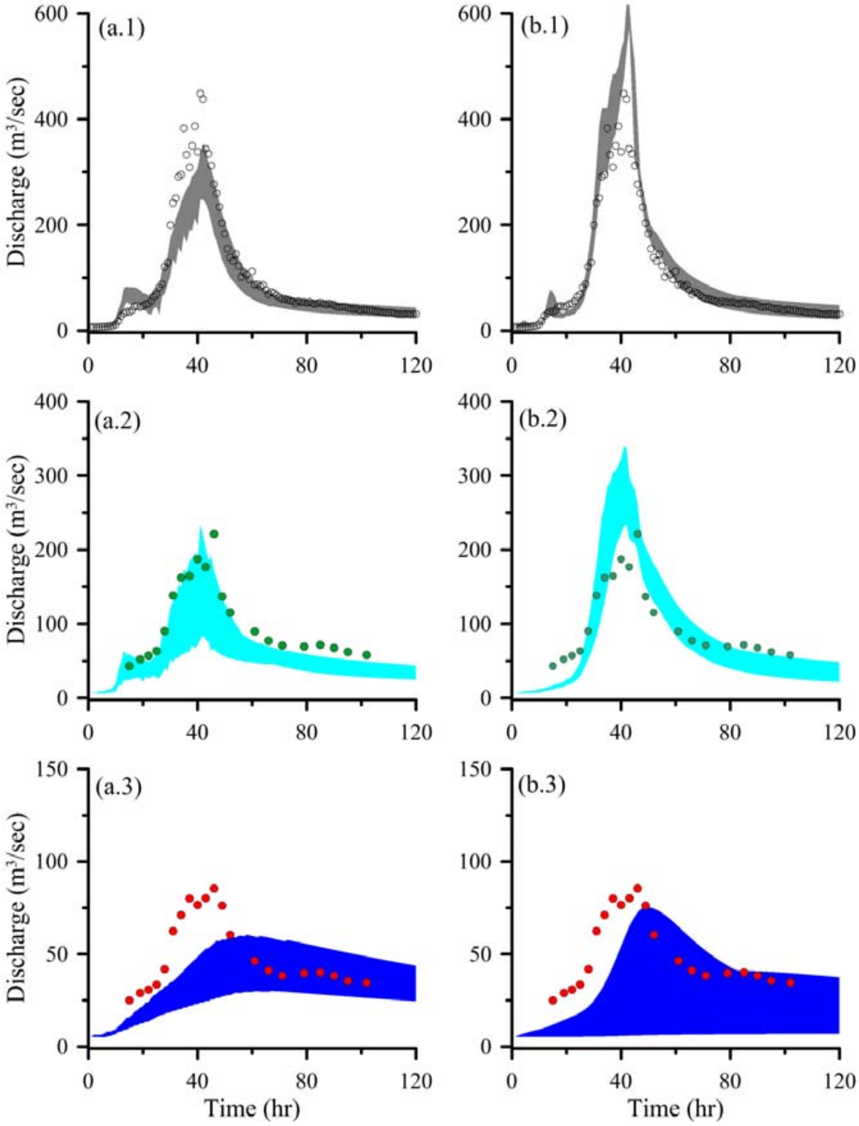

Figure 11. Comparison between the measured stream discharges (event no. 17) and the best 15 simulations derived from HBV model (a.1) and TOP model (b.1). The comparison of interflow derived from mixing analysis (green dots) with the simulated inter-flows (sky blue zone) derived from HBV model (a.2) and TOP model (b.2), respectively. The comparison of baseflow derived from mixing analysis (red dots) with the simulated inter-flows (blue zone) derived from HBV model (a.3) and TOP model (b.3), respectively.