Journal of Hydrology, Vol. 268, No. 1-4, 2002, pp. 72-86

Combining a fuzzy optimal model with a genetic algorithm to solve multiobjective rainfall-runoff model calibration

Cheng Chuntian[a,*] , Ou Chunping[b] , K.W.Chau[c]

[a]Department of Civil Engineering, Dalian University of Technology, 116024, Dalian, P.R.China

[b]Intelligent Engineering Lab, Institute of Software, Chinese Academy of Sciences, Box 8718, 100080, Beijing, P.R.China

[c]Department of Civil and Structural Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Abstract

An automatic calibration framework and methodology for the Xinanjiang model that has been successfully and widely applied in China is presented. The automatic calibration of the model consists of two parts: runoff parameter and runoff routing parameter calibration. The former is based on a simple genetic algorithm (GA). The latter is based on a new method which combines a fuzzy optimal model (FOM) with a GA for solving the multiple objective runoff routing parameters calibration problem. Except for the specific fitness where the membership degree of alternative obtained by FOM with limited alternatives and multi-objectives is employed, the multiple objective GA developed in this paper is otherwise the same as the simple GA. The parameter calibration includes optimization of multiple objectives: (1) peak discharge, (2) peak time and (3) total runoff volume. 34 historical floods from 12 years in the Shuangpai Reservoir are applied to calibrate the model parameters whilst 11 floods in recent two years are utilized to verify these parameters. Results of this study and application show that GAs not only can improve forecast accuracy but are also efficient and robust means.

Keywords: rainfall-runoff model; calibration; genetic algorithms; fuzzy optimal model; multiple objectives

1.Introduction

Conceptual rainfall-runoff models(CRRS) have become a basic tool for flood forecasting and for catchment basin management (Franchini and Galeati, 1997). CRRS generally have a large number of parameters, which cannot be directly obtained from measurable quantities of catchment characteristics, and hence model calibration is entailed. In order to calibrate a model, values of the model parameters are selected so that the model stimulates the hydrological behavior of the catchment as closely as possible(Madsen, 2000).

The successful application of CRRS largely depends on how well the model is calibrated (Duan, et al, 1992), i.e., the reliability of operational conceptual rainfall-runoff models used in forecasting is

* Corresponding author. Tel.: +86-411-4708517; Fax: +86-411-4674141
Email address: ctcheng@dlut.edu.cn

highly dependent on the adequacy of the calibration procedure employed (Sorooshian et al., 1983). The parameter calibration has been one of the important topics in CRRS (e.g., Sorooshian and Dracup, 1980;Wang, 1991; Yu and Yang, 2000; Khu et al., 2001). The process of model calibration is normally performed either manually or by using computer-based automatic procedures. Manual calibration is not only difficult to assess explicitly the confidence of the model simulation due to the subjective judgement involved and the need of an experienced hydrologist, but also is a very time consuming task. Thus, automated approaches to calibration have received much attention. There exists a large amount of literature related to the application of traditional optimization methods in solving the automatic calibration of rainfall-runoff model (herein only list part of them, Sorooshian et al., 1983; Gupa and Sorooshian, 1985;Duan et al., 1992; Madsen, 2000). However, none of the traditional optimization methods used for calibrating conceptual rainfall-runoff models, even for those with a moderate number of parameters, are robust and efficient in locating or nearly locating the global optima (Wang, 1991). Local searcher algorithms often encounter the convergence problems. It is not an easy task to get global optima using traditional optimization methods in which skill and experience of the developer are very important. On the other hand, calibration based on a single objective function is often inadequate to simulate all the important characteristics of the observed data. There is a need for effective and efficient multi-objective calibration procedures that are capable of exploiting all the useful information (Yapo et al., 1998; Madsen, 2000; Yu and Yang, 2000).

The objective of the present study is to present a framework and methodology for the automatic calibration of Xinanjiang model (Zhao et al., 1980; Zhao, 1992) with multiple objectives. In section 2, the applied rainfall-runoff model, which is the Xinanjiang model, is briefly described. In section 3, the general steps in the simple GA for the runoff parameter calibration are outlined. In section 4, the multi-objective GA developed in this paper is presented for solving the multi-objective runoff routing parameter calibration problem. In section 5, a case example is presented that illustrates in details the whole process for the calibration of the parameters in the Xinanjiang model. Finally, the conclusions are given in section 6.

2 Xinanjiang rainfall-runoff model: structure and calibrated parameters

The hydrological model used in this study is the Xinanjiang rainfall-runoff model, which is a soil moisture accounting model developed by Zhao et al. (Zhao et al., 1980; Zhao, 1992). The model has been successfully and widely applied in humid and semi-humid regions of China since its initial development in the 1970s. The model structure is shown in Figure 1.It consists of two components, which deal with water balance and routing, respectively. The runoff generating component can be described in seven parameters: U_m, L_m, D_m, B, I_m, K, C whilst the runoff routing component can be described in nine parameters: S_m, E_x ,K_g, K_i, C_g, C_i, C_s, K_e , X_e. The Xinanjiang model uses a single parabolic curve to represent the spatial distribution of the soil moisture storage capacity over the catchment where the exponential parameter B represents the non-uniformity of this distribution and W_m describes the averaged soil moisture storage capacity. The storage depth is divided into three layers where U_m ,L_m and D_m represent the averaged soil moisture storage capacity of the upper layer, the lower layer and the deep layer, respectively. I_m is the percentage of impervious and saturated areas in the catchment. K is the ratio of potential evapotranspiration to pan evaporation whilst C denotes a coefficient of the deep layer. S_m is the areal mean free water capacity of the surface soil layer, which represents the maximum possible deficit of free water storage. E_x is the exponent of the free water capacity curve influencing the development of the saturated area. K_g and K_i are the outflow coefficients of the free water storage to groundwater and interflow relationships, respectively. C_g and C_i are the recession constants of groundwater storage and of the lower interflow storage, respectively. C_s is the recession constant in the lag and route method for routing through the channel system within each sub-basin. K_e and X_e are parameters of the Muskingum method. More details about the Xinanjiang model refer to the reference from Zhao(1992).

From Figure 1, the basin is divided into a set of sub-areas and the runoff is first calculated. Then the outflow hydrograph from each of the sub-area is simulated and finally routed down the channels to the main basin outlet. So, parameter calibration of the model can be divided into two parts: runoff parameter calibration and runoff routing parameter calibration.

3 Parameter calibration of runoff generating component by a simple GA

In recent years, GAs have been shown to have advantages over classical optimization methods (Holland, 1975;Goldberg, 1989) and have become one of the most widely used techniques for solving a number of hydrology and water resources problems (Wang, 1991; Ritzel et al., 1994; Franchini, 1996, 1997;Savic et al., 1999;Vasquez et al., 2000; Sharif and Wardlaw, 2000; Khu et al., 2001). GAs are heuristic iterative search techniques that attempt to find the best solution in a given decision space based on a search algorithm that mimics Darwinian evolution and survival of the fittest in a natural environment. They are distinct from other heuristic iterative search techniques in that they search in parallel, using many individuals in the population instead of a single point. This is a desirable property for many practical applications and is suitable for the runoff parameter calibration. Generally, the error or percentage error of total runoff volume between the simulated and observed runoffs is an important performance measure for the runoff parameter calibration. The runoff parameter calibration is a single objective optimization problem. Thus, a simple GA is sufficient for solving the runoff parameter calibration.

The general steps in the simple GA for the runoff parameter calibration are as follows.

Step 1. The initial parameters for runoff calibration and initial GA parameters are first pre-set. All the parameters of the Xinanjiang model have clear physical meanings. These parameters are to some extent within the same scope for a special basin (Zhao, 1992). W_m varies from 80 mm in South China to 170 mm in North China. Typical values for U_m are from 5 to 20 mm for deforested to forested areas, respectively. L_m is within 60 to 90 mm. Experience indicates that B=0.1 for basins of area less than 10 km^2 and B=0.4 for basins with thousands of square kilometers. For natural basins in humid regions I_m is usually negligible, but in semi-humid or more arid regions the impervious area may be a large proportion of the runoff producing area in the basin. In the absence of the actual percentage of the impervious and saturated areas, I_m is assumed to be the ratio of runoff to rainfall generated by a small event of short duration which occurs after a long dry period. C depends on the proportion of the basin area covered by vegetation with deep roots. It varies from 0.18 in South China to 0.08 in North China. The objective of presetting these initial parameters of the Xinanjiang model is to speed up the automatic calibration process.

Good genetic algorithm performance requires the choice of a high crossover probability, a low mutation probability and a moderate population size. There are four GA parameters: p_c, p_m, P_{size} and T_{max}. p_c means the crossover probability parameter that is typically set so that crossover is performed on most, but not all, of the population. It varies from 0.3 to 0.9. p_m, which is the mutation probability parameter that controls the probability of selecting a gene for mutation, varies usually from 0.01 to 0.1. P_{size}, which is the population size parameter that provides sufficient sampling of the decision space while limiting the computational burden, varies from 10 to 100. T_{max}, which is the maximum number of generation, varies from 10 to 500.

Step 2 An initial population of chromosomes is randomly generated. The total number of chromosomes is controlled by P_{size}. Each chromosome is a finite-length string of numbers that represents the values of the decision variables for that chromosome. The decision variables may be coded using binary or real value. In the problem herein, decision variable is simply a vector of runoff parameters for the Xinanjiang model. The remaining steps described in this section are focused on GAs that use real value. A random function, *f_random(a, b)*, is coded to randomly generate an initial real value for each runoff parameter of the Xinanjiang model, where *a* and *b* are the lower and upper limits respectively, which vary with different runoff parameters.

Step 3. For every chromosome in the population the fitness is computed. Fitness calculation is a problem-oriented process. It has to be overridden by the user according to the system requirement. The problem herein is a maximization problem with a single objective without any constraint. As mentioned above, the objective function in the runoff parameter calibration is taken as the ratio of floods that are qualificatory relative to the total runoff volume. The simulated flood is qualificatory if the absolute error of the total runoff volume between the simulated runoff and the observed one is less than 3mm or if the absolute percentage error of the total runoff volume is less than 20% (NCHI,1985). The objective function value is usually used as a measure of fitness. Equation 1 shows the fitness calculation for maximizing the ratio of qualificatory floods relative to the total runoff volume.

$$f_{runoff} = \frac{M_p}{N} \times 100\% \tag{1}$$

where f_{runoff} is the fitness of this problem, M_p represents the total number of floods that satisfy the qualificatory criteria relative to the total runoff volume, and N is the total number of the calibrated floods.

Step 4. A crossover operation is applied in which two parent chromosomes exchange genetic information by interchanging portions of their strings, thereby generating either one, or more commonly two offspring per pair. The genetic operation of crossover is performed on each mated pair with a certain probability, referred to as crossover probability. The possible crossover operation is uniform, single point, two points and arithmetic crossover. An arithmetic crossover is designed to the crossover operation.

$$v_{i1}^{'} = \alpha_i v_{i1} + (1-\alpha_i) v_{i2} \tag{2}$$

$$v_{i2}^{'} = \alpha_i v_{i2} + (1-\alpha_i) v_{i1} \tag{3}$$

where v_{i1} and v_{i2} are the parent chromosomes, $v_{i1}^{'}$ and $v_{i2}^{'}$ are the children strings, α_i is a random number and $\in (0,1)$, $i=1,2,\ldots,k$. k is the total chromosome pairs for the crossover operation.

Step 5. A mutation operation is applied to avoid being trapped in local optima. Mutation probability controls the rate of mutation in the process of reproduction. Common mutation operation is simple, uniform, boundary, non-uniform and Gaussian mutation. A non-uniform mutation is designed to the mutation operation. If $V = (v_1, v_2, \cdots, v_n)$ is a chromosome and the element v_k was selected for this mutation (the domain of v_k is $[a_k, b_k]$), the result is a vector $V' = (v_1, v_2, \cdots, v_{k-1}, v_k', v_{k+1}, \cdots, v_n)$ and

$$v_k' = \begin{cases} v_k + \Delta(t, b_k - v_k), & if \quad random(0,1) = 0 \\ v_k - \Delta(t, v_k - a_k), & if \quad random(0,1) = 1 \end{cases} \qquad (4)$$

where the function $\Delta(t, y)$ returns a value in the range [0,y] such that the value for the probability of $\Delta(t, y)$ being close to 0 increases as *t* increases:

$$\Delta(t, y) = y \bullet (1 - r^{(1-t/\mathrm{T}_{\max})^{\lambda}}) \qquad (5)$$

where *r* is a random number in the interval[0,1]. T_{max} is the maximum number of generations and λ is a parameter chosen by the user, which determines the degree of dependency with the number of iterations. This property causes this operator to make an uniform search in the initial space when *t* is small , and a very local one in later stages.

Step 6. The most-fit chromosomes are selected to mate and reproduce. Common selection methods include the biased roulette wheel, tournament selection, ranking and ($\mu + \lambda$) selection methods. A ($\mu + \lambda$) selection method is used to produce offspring for the next generation. The method first generates λ children chromosomes from μ parent chromosomes by crossover and mutation operation, then selects μ strong chromosomes as a new population but keeps a mating pool as large as the selected original population.

Step 7. The processes described in Step 2 to 6 are repeated until a specified termination criterion, such as a limit on the maximum number of generation or no obvious change about fitness or pre-set fitness, is satisfied.

4 Parameter calibration of runoff routing component by coupling a fuzzy optimal model with GA

The runoff routing of the Xinanjiang model is divided into two parts. The runoff generated from the water balance component is first transformed into discharge by a linear system. The outflow hydrograph from each sub-area is finally routed down the channels to the main basin outlet by the Muskingum method. The parameter calibration must satisfy the constraints of the Muskingum method for each channel of sub-basin.

$$2K_eX_e \leqslant \Delta t \leqslant 2K_e - 2K_eX_e \qquad (6)$$

where K_e and X_e are the Muskingum coefficients, K_e is a storage constant having the dimension of time, X_e is a dimensionless constant for the reach of the river and Δt is the routing period.

According to the national criteria for flood forecasting in China, the percentage error of peak discharge, peak time and total runoff volume are three important performance measures to evaluate real-time flood forecasting and flood simulation (NCHF, 1985). The result of simulation or forecasting is qualificatory relative to peak value for this flood if the absolute percentage error of peak discharge between the simulated and observed floods is less than 20%. The result is qualificatory relative to peak time if the difference in peak time is within a routing period and relative to total runoff volume if absolute error between the simulated and observed floods is less than 3mm or absolute percentage error less than 20%, respectively. The evaluation of parameter calibration for runoff routing is counting the three ratios of qualificatory criteria relative to the peak discharge, peak time and total runoff volume, respectively. So, the parameter calibration of runoff routing component is a multiple objective optimization problem with constraints.

When solving the multiple objective calibration problem, the problem is usually transformed into a single objective optimization problem by defining a scalar that aggregates the various objective functions. However, it is often difficult to derive all objective functions for practical problems(Gupta and Sorooshian, 1985). Thus, it is worthy of exploring a direct optimization method without transformation of objectives.

The performance of a rainfall-runoff model heavily depends on choosing suitable model parameters, which are normally calibrated by using an objective function. Relevant literature mostly focused on selecting an appropriate objective function in the rainfall-runoff model when using classical optimization techniques (Gupta and Sorooshian, 1985;Yu et al., 2000). However, calibration based on a single performance measure is often inadequate to measure properly the simulation of all the important characteristics of the system that are reflected in the observations (Yapo et al., 1998; Madsen, 2000). There are various skills in selecting the best objective function. The complete and consistent application of these skills is still an art, depending on the problem-oriented process and the adopted optimization method. Most of conventional methods for parameter calibration of a rainfall-runoff model only consider a single objective but hardly deal with multiple objectives at the same time due to the limitation of optimization techniques. Many optimization techniques require large computer resources and derivatives of the objective functions. GAs work with numerical values, and can also establish objective functions without difficulty. They are free from a particular model structure and thereby only require an estimate of the objective function value for each decision set in order to proceed, regardless of whether such information comes from a simple equation or a very complex model. Fitness information, instead of complex and difficult functions, is the only requirement for GAs. The advantages of GAs over conventional parameter optimization techniques are that they are appropriate for the ill-behaved problem, highly non-linear spaces for global optima and adaptive algorithm. As such, GAs are suitable for many practical application and desirable methods for optimization problem with multi-objectives.

The successful application of GAs in solving a given optimization problem greatly depends on the appropriate choice of the fitness. Thus the focus of the following will be on the definition of fitness for the parameter calibration about the runoff routing of the Xinanjiang model.

Except for the specific fitness detailed below, the multiple objective GA developed in this section is basically the same as the simple genetic algorithm described earlier. From the steps of GA mentioned in the previous section, the procedure, with the ($\mu + \lambda$) selection method being used to produce offspring for the next generation, is in fact an evaluation problem with limited alternatives and multiple objectives. Alternatives herein are chromosomes whilst objectives are the flood characteristics such as peak value, peak time and total runoff volume. The above mentioned problem composing of selected alternatives with multi-objectives is a typical multi-objective evaluation with limited alternatives. Here a fuzzy optimal method, in which the optimal rank of alternatives can be obtained by the membership degree of alternative, is employed. (Chen, 1994; Cheng, 1999; Cheng and Chau, 2001).

It is supposed that the total number of objectives for a chromosome evaluation is m, and the total number of chromosomes through crossover and mutation operation is n. The alternative set constituting of n alternatives is denoted by $A = \{A_1, A_2, \ldots, A_n\}$. The decision matrix is represented by $X = (x_{ij})_{m \times n}$, where x_{ij} is the ith objective value of the alternative A_j ($j=1,2,\ldots,n$). In determining the relatively optimal decision among n alternatives, the decision matrix X should be transformed into the matrix of membership degree by the following equations

$$r_{ij} = x_{ij} / x_{i\max} \tag{7}$$

or $$r_{ij} = (1 - x_{ij}) / x_{i\max} \tag{8}$$

where $x_{i\max} = \bigvee_{j=1}^{n} x_{ij}$. If a larger objective represents more optimum membership degree, equation (7) should be adopted. Otherwise, equation (8) should be used. After the transformation, the matrix of membership degree is represented as

$$\mathrm{R} = (r_{ij})_{m \times n} \tag{9}$$

Obviously, the ideal alternative is the vector $\underbrace{(1,1,\ldots,1)^T}_{m}$. It is defined as

$$\boldsymbol{G} = \underbrace{(1,1,\ldots,1)^T}_{m} \tag{10}$$

According to the fuzzy sets theory, the non-ideal alternative, the counterpart of $\boldsymbol{G}$, is defined as

$$\boldsymbol{B} = \underbrace{(0,0,\ldots,0)^T}_{m} \tag{11}$$

Conflicts often exist among objectives and the alternatives $\boldsymbol{G}$ and $\boldsymbol{B}$ are only "fictitious". In order to acquire the optimal solution, it is natural to select an alternative closest to $\boldsymbol{G}$ and farthest away from $\boldsymbol{B}$. The weighted distances are defined as

$$D_j(w) = \sqrt{\sum_{i=1}^{m} [w_i (g_i - r_{ij})]^2} = \sqrt{\sum_{i=1}^{m} [w_i (1 - r_{ij})]^2} \tag{12}$$

$$d_j(w) = \sqrt{\sum_{i=1}^{m} [w_i (r_{ij} - b_i)]^2} = \sqrt{\sum_{i=1}^{m} (w_i r_{ij})^2} \tag{13}$$

In equations (12) and (13), w is the weighting vector, $w = (w_1, w_2, \ldots, w_m)^T$, $\sum_{i=1}^{m} w_i = 1$, $w_i > 0$, $i=1, 2, \ldots, m$.

If the membership degree of alternative A_j relative to $\boldsymbol{G}$ is denoted by u_j, then its counterpart relative to $\boldsymbol{B}$ is $1 - u_j$. The synthetically weighted distance is defined by

$$f_j(u_j) = u_j^2 \sum_{i=1}^{m} [w_i (1 - r_{ij})]^2 + (1 - u_j)^2 \sum_{i=1}^{m} (w_i r_{ij})^2 \tag{14}$$

Then the solution is given by

$$\text{Min}\{f_j(u_j)\}$$

$$s.t. 0 < u_j < 1, j = 1,2,...,n \tag{15}$$

Let $df(u_j)/du_j = 0$ *(j=1,2,...,n),* giving n equations:

$$u_j = \left[1 + \frac{\sum_{i=1}^{m}[w_i(1-r_{ij})]^2}{\sum_{i=1}^{m}(w_i r_{ij})^2} \right]^{-1} \tag{16}$$

where *i=1,2,...,m; j=1,2,...,n*.

From equations (16), when weighting vector ***w*** is known from the experience, u_j can be obtained. According to the maximization rules of membership degree, the optimal order of alternatives can be obtained. Comparing with the procedure reproducing offspring in GAs for the next generation, the membership degree of alternative u_j can be defined as the fitness of *jth* chromosome. The GA combining the fuzzy optimal model can be summarized as follows:

Step 1. The initial parameters for runoff routing calibration and GA parameters are first pre-set.

Step 2. An initial population of chromosome is randomly generated.

Step 3. The multiple objective values for each chromosome are assessed.

Step 4. The chromosome operators, selection, crossover and mutation are processed.

Step 5. The finesses for all chromosomes are evaluated based on the fuzzy optimal model with limited alternatives and multiple objectives.

Step 6. The chromosomes with lower fitness values are eliminated and the new chromosomes with higher fitness values are added.

Step 7. If the termination criterions are satisfied, the process stops; otherwise, it goes back to Step 2.

5 Application example

The model is applied to the Shuangpai Reservoir. The reservoir, with a drainage area of 10,594 km^2 and a water holding capacity of up to 373.8 million cubic meters, is situated in Hunan province of southern China and at the downstream of the Xiaoshui Stream, which is one of tributary rivers in the Xiangjiang River. The reservoir is used for power generation, flood control, as well as irrigation purposes. The temporal distribution of the rainfall during a given year is significantly heterogeneous in this area. The rainfall in this area is mainly due to the thunderstorms. 45.9% of the total rainfall falls between April and June, and 34% of the total rainfall between September and October, which are referred to as the high flow periods. The annual rainfall is 1,500mm; the averaged depth of runoff is 893mm and the averaged discharge is 300m^3/s. Table 1 summarizes the stations used in this study, including the weighting and area for each rain-gauge. The data sets selected for modeling process are rainfall, streamflow and evaporation, as shown in Table 2. 34 historical floods between 1984 and 1995 are used for the parameter calibration whilst 11 floods between 1999 and 2000 are used for the parameter validation.

5.1 Calibration of the parameters

There are no differences in the operation between the runoff parameter calibration and runoff routing parameter calibration except in presetting the initial parameters. Presetting initial parameters is a unique work by manual means. All presetting processes can be performed interactively based on the visual interfaces. After having entered the initial conditions, the model parameters will be automatically calibrated. Table 3 lists the initial parameter values for the Shuangpai reservoir. The four GA parameters are also preset, which are the same for the runoff parameter calibration and runoff routing parameter calibration, i.e., p_c,=0.8, p_m=0.1, P_{size} =20, *and* T_{max} =500. Table 4 shows the results of the runoff parameter calibration. Table 5 lists the performances of the calibrated runoff parameters whilst Table 6 presents the results of runoff routing parameter calibration. Table 7 gives the performances of the calibrated runoff routing parameters. Figure 2 to Figure 13 show the simulated and observed hydrographs from 1984 to 1995 during the calibration.

5.2 Validation of the parameters

A total number of 11 floods are used to validate the model parameters from 1999 to 2000. Table 8 lists performances of the validated runoff parameters whilst Table 9 shows performances of the validated runoff routing parameters. Figure 14 to Figure 15 show the simulated and observed hydrographs from 1999 to 2000 during the validation.

5.3 Analysis of results

All of the calibrated floods and the validated floods are qualificatory relative to the total runoff volume. The results clearly showed that the runoff parameters are able to provide reliable simulation and forecast. 30 floods among the total number 34 of the calibrated floods are qualificatory relative to the peak discharge and peak time, with the qualificatory ratios more than 85%. For the validated floods, all floods are qualificatory relative to peak time and total runoff volume whilst only one flood is not qualificatory relative to peak discharge, with the qualificatory ratios also more than 85%. Thus, the runoff routing parameters are also able to provide reliable simulation and forecasts. Figure 3 to Figure 16 also show that the simulated hydrographs better fitted the observed hydrographs.

6. Conclusions

The paper has addressed the solution in determining the optimal parameters for the Xinanjiang model. The methodology presented herein satisfies at the same time the demand of evaluating three important flood characteristics, such as peak discharge, peak time and total runoff volume. Our results indicate that attempts to calibrate the Xinanjiang model parameters through two procedures consisting of runoff and runoff routing parameter calibration, have been successful. It may be concluded that GA is a robust and efficient tool in solving a complex conceptual rainfall-runoff model parameter calibration problem for a large-scale basin with a drainage area more than of 10,000 km^{2}.

Acknowledgments

This research was supported by “The Teaching and Research Award Fund for Outstanding Young Teachers in Higher Education Institutions of Ministry of Education, P.R.C” (No.200026), “Fok Ying Tung Education Foundation, P.R.C”(No.71072), the Chinese National Natural Science Foundation (No. 60073037).

References

Brian J. Ritzel and J.Wayland Eheart, 1994. Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. *Water Resources Research*, 30(5), 1589~1603

Chen Shouyu, 1994. Theory of fuzzy optimum for multistage and multi-objective decision making system. *Journal of Fuzzy Mathematics*, 2, 163~174

Cheng Chuntian, 1999. Fuzzy optimal model for the flood control system of the upper and middle reaches of the Yangtze River. *Hydrological Sciences Journal* , 44(4), 573~582

Cheng Chuntian, K.W. Chau, 2001. Fuzzy iteration methodology for reservoir flood control operation. *Journal of the American Water Resources Association, (in press)*

David E.Goldberg, 1989. Genetic algorithms in search, optimization & machine learning. Addison-Wesley, Reading, MA

Dragan A.Savic, Godfrey A.Walters and James W.Davidson, 1999. A genetic programming approach to rainfall-runoff modelling. *Water Resources Management*, 12, 219~231

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

Holland, J.H., 1975. Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor, MI.

Jose A. Vasques, Holger R. Mailer, Barbara J.Lence, Bryan A. Tolson and Ricardo O.Foschi, 2000. Achieving water quality system reliability using genetic algorithms. *Journal of Environmental Engineering*, 126(10), 954~962

Marco Franchini, 1996. Using a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal*, 41(1), 21~40

Marco Franchini, Giorgio Galeati, 1997. Comparing several genetic algorithm schemes for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal*, 42(3), 357~379

Moha sharif and Robin Wardlaw, 2000. Multireservoir systems optimization using genetic algorithms: case study. *Journal of Computing in Civil Engineering*, 14(4), 255~263

National Center of Hydrological Information(NCHI), P.R.China, 1985. The national criteria for hydrological forecasting. Hydroelectric Press, Beijing, P.R.China(in Chinese)

Pao-Shan Yu, Tao-Chang Yang, 2000. Fuzzy multi-objective function for rainfall-runoff model calibration. *Journal of Hydrology*, 238, 1~14

Patrice Ogou Yapo, Hoshin Vijai Gupta, Soroosh Sorooshian, 1998. Multi-objective global optimization for hydrologic models. *Journal of Hydrology*, 204, 83~97

Qingyun Duan, Soroosh Sorooshian and Vijai Gupta, 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4), 1015~1031

Qingyun Duan, Soroosh Sorooshian and Vijai Gupta, 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158, 265~284

Q.J.Wang, 1991. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research*, 27(9), 2467~2471

Soon Thiam Khu, Shie-Yui Liong, Vladan Babovic, Henrik Madsen and Nitin Muttil, 2001. Genetic programming and its application in real-time runoff forecasting. *Journal of the American Water Resources Association*, 37(2), 439~451

Soroosh Sorooshian, Qingyun Duan and Vijai Kumar Gupta, 1993. Calibration of rainfall-runoff models: application of global optimization to the sacramento soil moisture accounting model. *Water Resources Research*, 29(4), 1185~1194

Soroosh Sorooshian, John A.Dracup, 1980. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. *Water Resources Research*, 16(2), 430~442

Vijai Kumar Gupta, Soroosh Sorooshian, 1985. The automatic calibration of conceptual catchment models using derivative-based optimization algorithms. *Water Resources Research*, 21(4), 473~485

Zhao, R.J., Zhang, Y.L. and Fang L.R., 1980. The Xinanjiang model. *Hydrological Forecasting Proceedings Oxford Symposium*, IASH 129,351~356

Zhao,R.J., 1992. The Xinanjiang model applied in China. *Journal of Hydrology*, 135, 371~381

Table 1 Details of rain gauge stations in the study area

Station	Kind of station	Station name	Weighting	Area(km^2)
01	Rainfall	Jiangcun	0.0915	745
02	Rainfall	Daoxian	0.0846	691
03	Rainfall	Haofu	0.0689	562
04	Rainfall	Jiangyong	0.1134	926
05	Rainfall	Dalupu	0.1200	979
06	Rainfall	Qingtianhe	0.0326	266
07	Rainfall	Simaqiao	0.0651	531
08	Rainfall	Youxiang	0.0762	622
09	Rainfall	Linyuan	0.0911	744
10	Rainfall	Shuishi	0.0651	532
11	Rainfall	Baijiaping	0.1210	988
12	Rainfall/streamflow	Shuangpai	0.0705	576

Table 2 Data sets in this study

Type	Time period (year)	Purpose	Remarks
Daily rainfall	1984-1995	Calibrate	For 12 rain gauges
	1999-2000	Validate	
Daily evaporation	1984-1995	Calibrate	
	1999-2000	Validate	
Rainfall of 3 hours duration	1984-1995	Calibrate	For 12 rain gauges
	1999-2000	Validate	
Streamflow	1984-1995	Calibrate	For Shuangpai streamflow gauge
	1999-2000	Validate	

Notes：Data between 1996-1998 are lacking because of problem on official right. The data between 1999-2000 are supplied directly by the user.

Table 3 The initially preset parameters

Parameter kind	Parameter name	Lower limit	Upper limit
Runoff	B	0.1	0.6
	I_m	0.01	0.03
	K	0.50	1.10
	C	0.10	0.20
	U_m	10.0	30.0
	L_m	60.0	90.0
	D_m	10.0	50.0
Runoff routing	S_m	10.0	30.0
	E_x	1.0	1.5

K_g	0.25	0.35
K_I	0.35	0.45
C_I	0.10	0.90
C_g	0.80	1.0
C_s	0.01	0.50
K_e	1.0	3.0
X_e	0.1	0.5

Table 4 Results of the calibrated runoff parameters

parameter	B	I_m	K	C	U_m	L_m	D_m
Value	0.55	0.02	0.74	0.14	18.03	62.96	46.05

Table 5 Performances of the calibrated runoff parameters

Floods	Average rainfall (mm)	Observed Runoff (mm)	Simulated runoff (mm)	Absolute error (mm)	Percentage error (%)	Qualificatory
19840416	27.5	24.9	26.2	-1.33	-5.34	Yes
19840530	100.9	98.0	94.1	3.90	3.98	Yes
19850411	29.5	24.1	28.4	-4.25	-17.63	Yes
19850527	84.3	75.7	66.1	9.65	12.75	Yes
19850610	35.8	30.6	26.4	4.18	13.66	Yes
19860706	66.8	51.2	52.5	-1.28	-2.50	Yes
19870404	28.4	24.0	28.4	-4.43	-18.46	Yes
19870515	31.7	29.8	29.8	0.01	0.03	Yes
19870521	30.7	28.8	27.1	1.66	5.76	Yes
19870606	28.3	20.5	18.6	1.91	9.32	Yes
19870614	41.8	22.8	25.4	-2.60	-11.4	Yes
19870722	47.3	21.9	20.1	1.77	8.08	Yes
19870729	40.9	33.7	28.6	5.11	15.16	Yes
19880903	51.7	37.7	40.9	-3.18	-8.44	Yes
19890511	71.0	54.4	56.5	-2.12	-3.90	Yes
19890522	45.3	30.7	29.4	1.31	4.27	Yes
19890529	44.8	38.1	33.2	4.88	12.81	Yes
19900530	67.6	51.9	57.1	-5.21	-10.04	Yes
19900607	49.3	35.5	33.7	1.83	5.15	Yes
19910616	31.1	20.1	17.0	3.10	15.42	Yes
19920423	50.4	36.5	39.0	-2.51	-6.88	Yes
19920516	32.0	27.2	31.9	-4.74	-17.43	Yes
19920705	86.5	72.5	86.3	-13.80	-19.03	Yes
19930513	47.7	35.8	39.0	-3.19	-8.91	Yes
19930607	59.4	45.2	44.3	0.89	1.97	Yes
19930615	30.9	25.4	28.4	-3.04	-11.97	Yes
19940421	90.6	80.7	75.1	5.63	6.98	Yes
19940525	65.9	58.4	58.5	-0.12	-0.21	Yes
19940614	82.5	77.8	77.8	-0.02	-0.03	Yes
19940723	106.3	95.5	93.8	1.68	1.76	Yes
19940805	76.9	62.2	64.6	-2.39	-3.84	Yes
19950425	47.3	38.8	45.6	-6.80	-17.53	Yes
19950526	57.1	24.1	26.8	-2.70	-11.20	Yes
19950614	122.6	115.5	105.5	9.98	8.64	Yes

Notes: 1. The simulated flood is qualificatory if absolute error is less than 3mm or if absolute percentage error is less than 20%

2. The total number of floods is 34. All floods are qualificatory and the ratio of qualifying simulation is 100%

Table 6 Results of the calibrated runoff routing parameters

Parameter	S_m	E_x	K_g	K_i	C_i	C_g	C_s	K_e	X_e
Value	12.90	1.4	0.27	0.38	0.21	0.87	0.15	1.93	0.21

Table 7 Performances of the calibrated runoff routing parameters

Floods	Peak discharge				Peak time			Percentage error in total runoff volume (%)
	Observed (m^3/s)	Simulated (m^3/s)	Error (m^3/s)	Percentage Error (%)	Observed (yyyy-mm-dd hh:mm)	Simulated (yyyy-mm-dd hh:mm)	Error (Number)	
19840416	1130	1422	-292	-25.83	1984-04-17 14:00	1984-04-17 11:00	-1	-5.34
19840530	4310	4890	-580	-13.45	1984-06-01 17:00	1984-06-01 14:00	-1	3.98
19850411	1200	1596	-396	-32.98	1985-04-12 23:00	1985-04-12 20:00	-1	-17.63
19850527	5770	6412	-642	-11.12	1985-05-27 23:00	1985-05-27 23:00	0	12.75
19850610	1370	1195	175	12.74	1985-06-11 23:00	1985-06-11 17:00	**-2**	13.66
19860706	2560	2052	508	19.83	1986-07-06 20:00	1986-07-06 23:00	1	-2.50
19870404	1790	1463	327	18.24	1987-04-05 20:00	1987-04-05 23:00	1	-18.46
19870515	1720	1377	343	19.95	1987-05-16 05:00	1987-05-16 05:00	0	0.03
19870521	1840	1931	-91	-4.94	1987-05-22 02:00	1987-05-22 02:00	0	5.76
19870606	1080	1073	7	0.61	1987-06-07 11:00	1987-06-07 14:00	1	9.32
19870614	1350	1308	42	3.15	1987-06-14 23:00	1987-06-14 20:00	-1	-11.4
19870722	1740	1181	559	32.11	1987-07-23 11:00	1987-07-23 14:00	1	8.08
19870729	2170	1943	227	10.46	1987-07-29 20:00	1987-07-29 20:00	0	15.16
19880903	1960	1891	69	3.53	1988-09-05 08:00	1988-09-04 23:00	**-3**	-8.44
19890511	2870	2672	198	6.90	1989-05-13 14:00	1989-05-13 11:00	-1	-3.90
19890522	1890	1885	5	0.28	1989-05-22 23:00	1989-05-22 23:00	0	4.27
19890529	1880	2181	-301	-16.02	1989-05-31 05:00	1989-05-30 23:00	**-2**	12.81
19900530	2770	2610	160	5.76	1990-06-01 02:00	1990-06-01 02:00	0	-10.04
19900607	2110	2305	-195	-9.25	1990-06-08 11:00	1990-06-08 11:00	0	5.15
19910616	1010	945	65	6.44	1991-06-16 23:00	1991-06-16 20:00	-1	15.42
19920423	2320	2389	-69	-2.98	1992-04-24 11:00	1992-04-24 14:00	1	-6.88
19920516	3020	2710	310	10.27	1992-05-17 17:00	1992-05-17 17:00	0	-17.43
19920705	3760	4842	-1082	-28.78	1992-07-06 20:00	1992-07-06 17:00	-1	-19.03
19930513	2560	2339	221	8.62	1993-05-14 14:00	1993-05-14 14:00	0	-8.91
19930607	2010	1737	273	13.56	1993-06-09 11:00	1993-06-09 14:00	1	1.97
19930615	1940	1940	0	-0.01	1993-06-16 02:00	1993-06-15 23:00	-1	-11.97
19940421	5070	5443	-373	-7.35	1994-04-23 20:00	1994-04-23 17:00	-1	6.98
19940525	2700	3055	-355	-13.17	1994-05-26 14:00	1994-05-26 17:00	1	-0.21
19940614	3330	2870	460	13.81	1994-06-18 08:00	1994-06-17 17:00	**-5**	-0.03
19940723	5810	5914	-104	-1.79	1994-07-24 05:00	1994-07-24 02:00	-1	1.76
19940805	2660	2884	-224	-8.44	1994-08-06 11:00	1994-08-06 14:00	1	-3.84
19950425	2040	1814	226	11.06	1995-04-26 11:00	1995-04-26 14:00	1	-17.53
19950526	1400	1636	-236	-16.84	1995-05-26 17:00	1995-05-26 20:00	1	-11.20
19950614	3880	4650	-770	-19.85	1995-06-17 02:00	1995-06-17 02:00	0	8.64

Notes: 1. The total number of floods, which are qualificatory relative to the percentage error of peak discharge, is 30 and the ratio of qualifying simulation is 88.24%

2. The total number of floods, which are qualificatory relative to the error of peak time, is 30 and the ratio of qualifying simulation is 88.24%

3. The total number of floods, which are qualificatory relative to the percentage error of total runoff volume, is 34 and the ratio of qualifying simulation is 100%

Table 8 Performances of the validated runoff parameters

Floods	Average rainfall (mm)	Observed Runoff (mm)	Simulated runoff (mm)	Absolute error (mm)	Percentage error (%)	qualificatory
19990426	52.3	47.8	39.6	8.2	17.15	Yes
19990526	99.9	82.9	74.8	8.11	9.78	Yes
19990618	51.4	30	25.9	4.13	13.77	Yes
19990627	53	32.9	34.6	1.66	5.05	Yes
19990831	49.9	38.4	34.4	4.05	10.55	Yes
20000403	37.3	25.1	28.5	3.4	13.55	Yes
20000410	39.2	35.4	38.9	3.46	9.77	Yes
20000426	26.4	14.5	14.7	0.22	1.52	Yes
20000510	34.9	19.7	18.3	1.39	7.06	Yes
20000528	88.5	62.6	72.5	9.94	15.88	Yes
20000611	44.8	24.3	21.6	2.74	11.28	Yes

Note: The total number of floods is 11. All floods are qualificatory and the ratio of qualifying simulation is 100%

Table 9 Performances of the validated runoff routing parameters

Floods	Peak discharge				Peak time			Percentage error in total runoff volume (%)
	Observed (m^3/s)	Simulated (m^3/s)	Error (m^3/s)	Percentage Error (%)	Observed (yyyy-mm-dd hh:mm)	Simulated (yyyy-mm-dd hh:mm)	Error (Number)	
19990426	2550	2512	38	1.47	1999-04-25 20:00	1999-04-25 23:00	1	3.55
19990526	4893	5083	-190	-3.88	1999-05-26 17:00	1999-05-26 17:00	0	0.33
19990618	1690	1786	-96	-5.66	1999-06-18 20:00	1999-06-18 23:00	1	5.16
19990627	1111	1144	-33	-3.01	1999-06-25 17:00	1999-06-25 17:00	0	4.83
19990831	1965	1946	19	0.96	1999-08-31 23:00	1999-08-31 23:00	0	2.75
20000403	1628	1928	-300	-18.43	2000-04-02 23:00	2000-04-03 02:00	1	-7.68
20000410	1798	2044	-246	-13.70	2000-04-09 20:00	2000-04-09 17:00	-1	-12.23
20000426	909	800	109	11.95	2000-04-26 20:00	2000-04-26 17:00	-1	-0.27
20000510	1039	1080	-41	-3.90	2000-05-10 02:00	2000-05-10 02:00	0	7.13
20000528	3163	3298	-135	-4.27	2000-05-28 14:00	2000-05-28 17:00	1	-6.03
20000611	1096	1498	-402	**-36.65**	2000-06-12 05:00	2000-06-12 02:00	-1	5.33

Notes: 1. The total number of floods, which are qualificatory relative to the percentage error of peak discharge, is 10 and the ratio of qualifying simulation is 90.91%

2. The total number of floods, which are qualificatory relative to the error of peak time, is 11 and the ratio of qualifying simulation is 100%

3. The total number of floods, which are qualificatory relative to the percentage error of total runoff volume, is 11 and the ratio of qualifying simulation is 100%

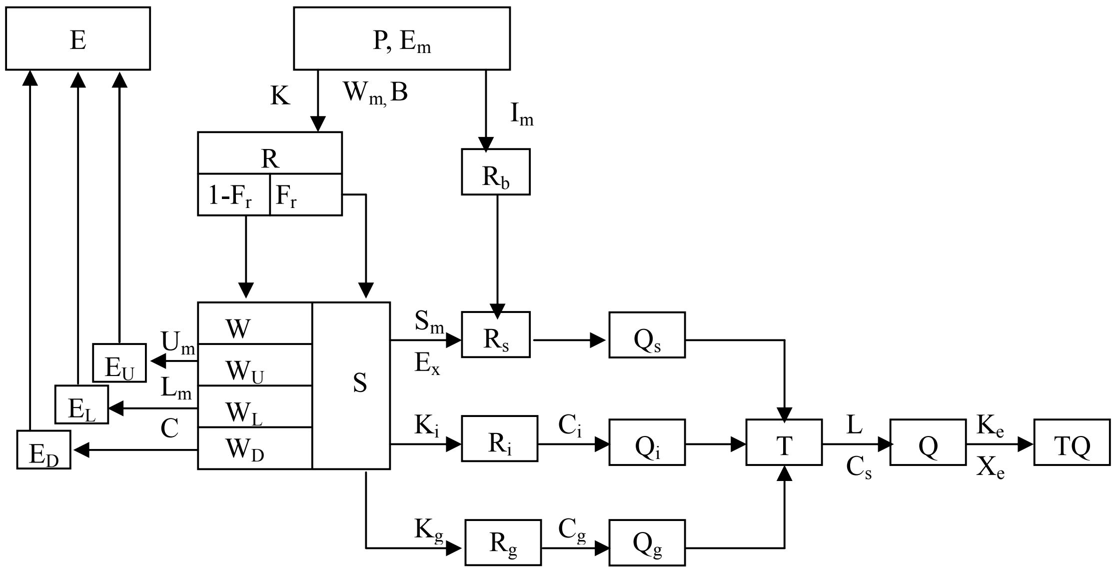

Figure 1 Flow chart for the Xinanjiang model

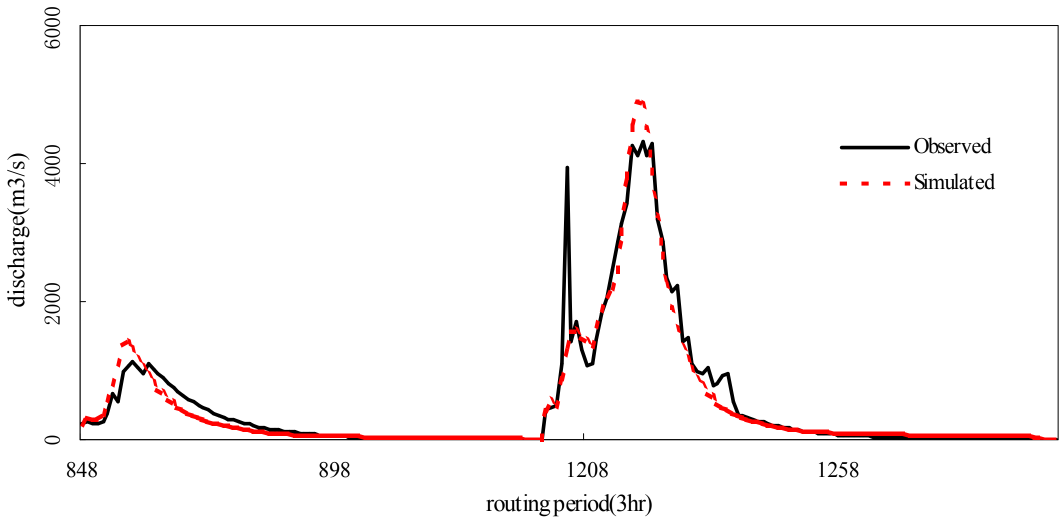

Figure 2 The simulated and observed hydrographs for 1984 during calibration

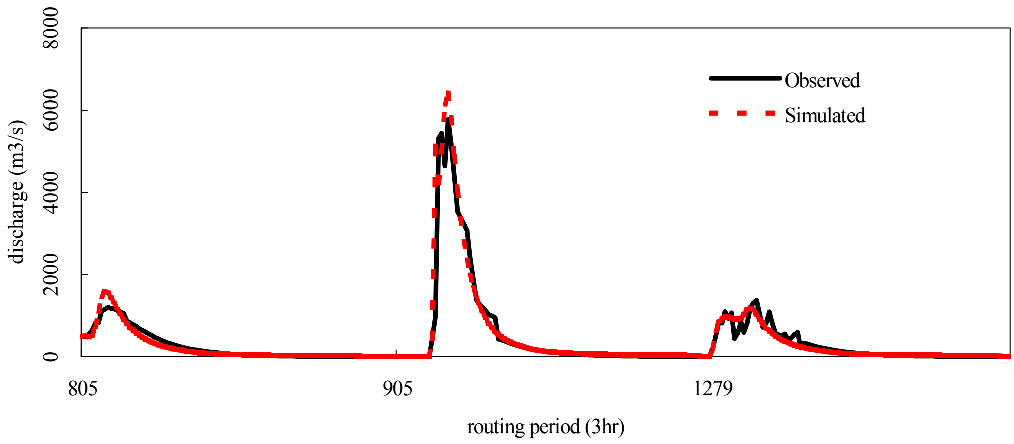

Figure 3 The simulated and observed hydrographs for 1985 during calibration

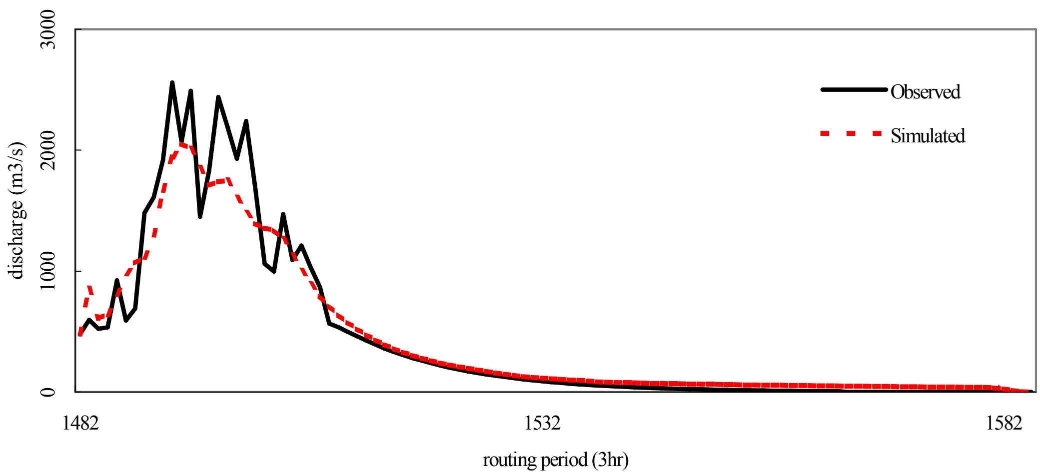

Figure 4 The simulated and observed hydrographs for 1986 during calibration

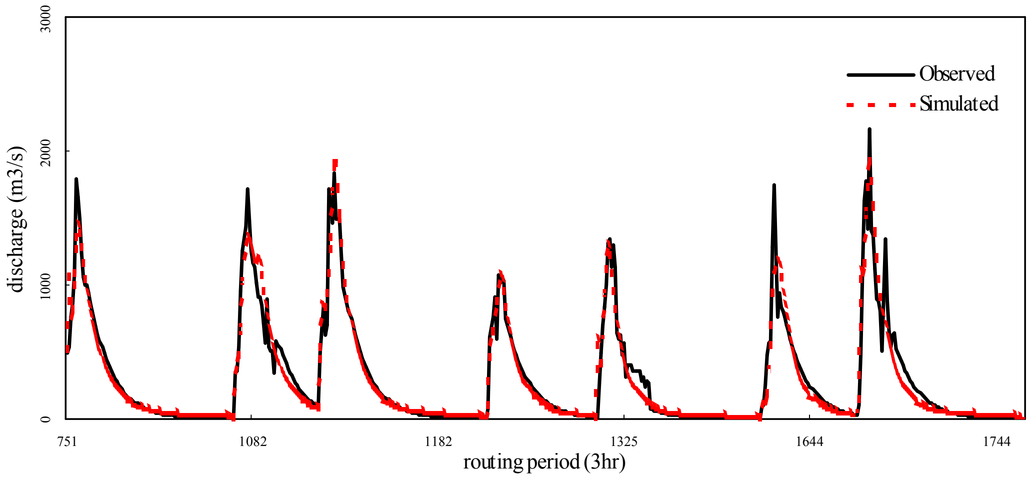

Figure 5 The simulated and observed hydrographs for 1987 during calibration

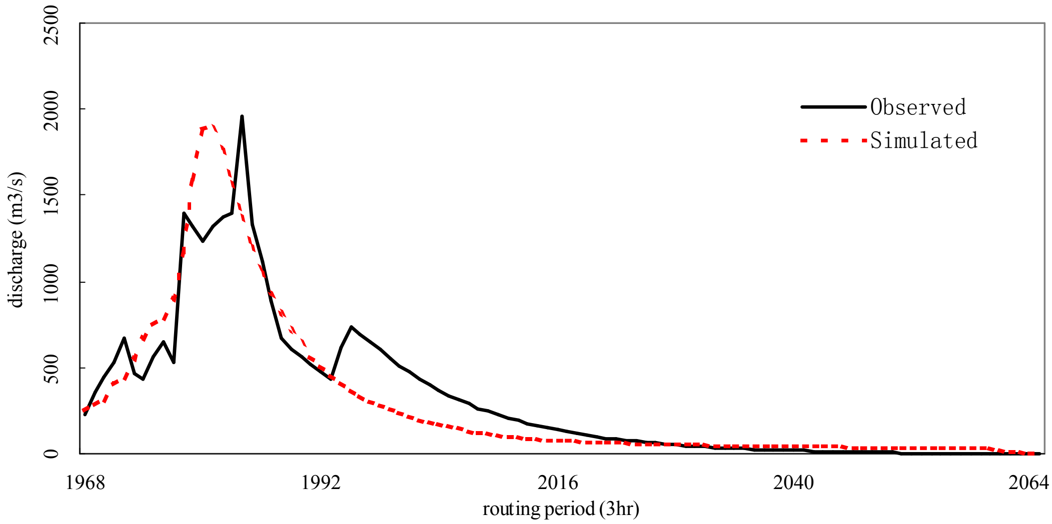

Figure 6 The simulated and observed hydrographs for 1988 during calibration

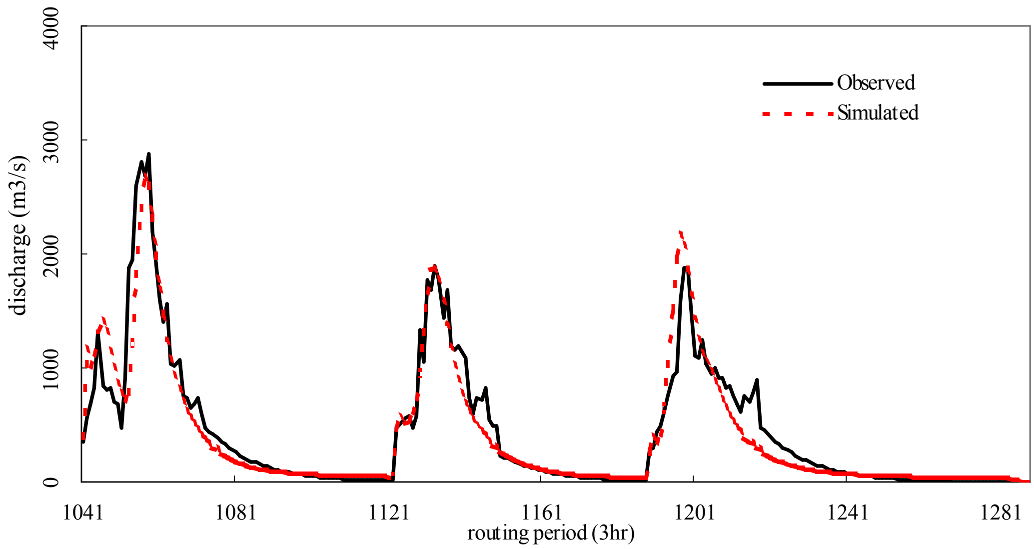

Figure 7 The simulated and observed hydrographs for 1989 during calibration

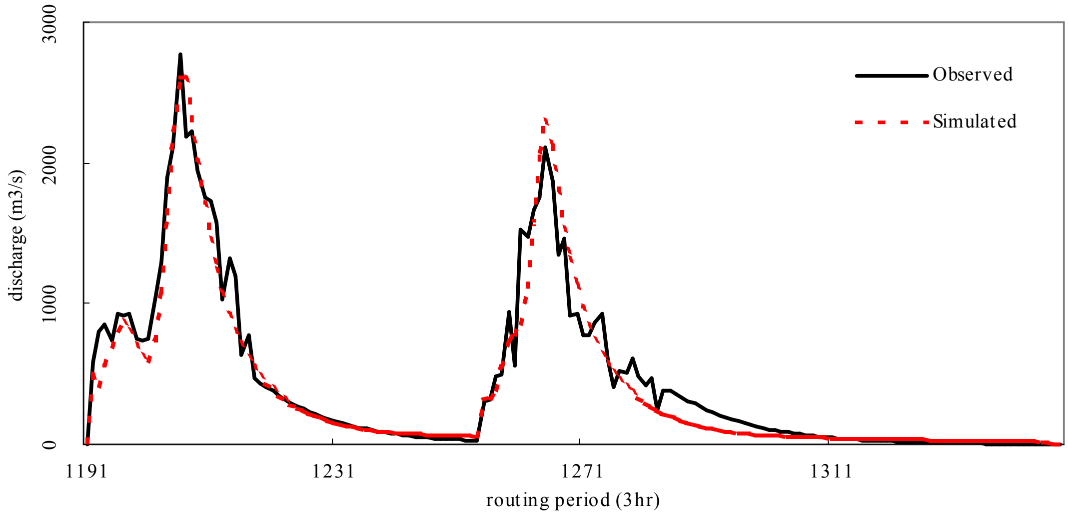

Figure 8 The simulated and observed hydrographs for 1990 during calibration

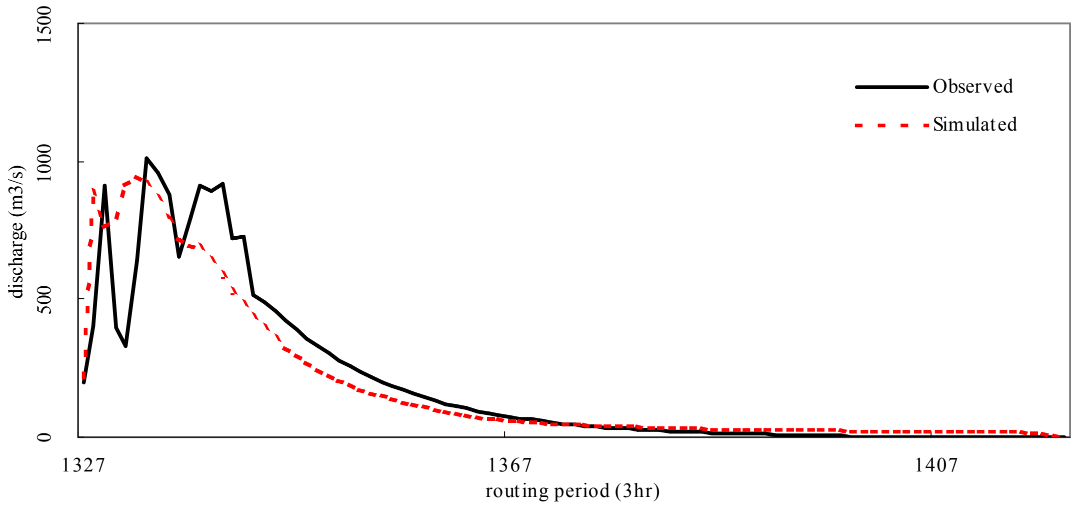

Figure 9 The simulated and observed hydrographs for 1991 during calibration

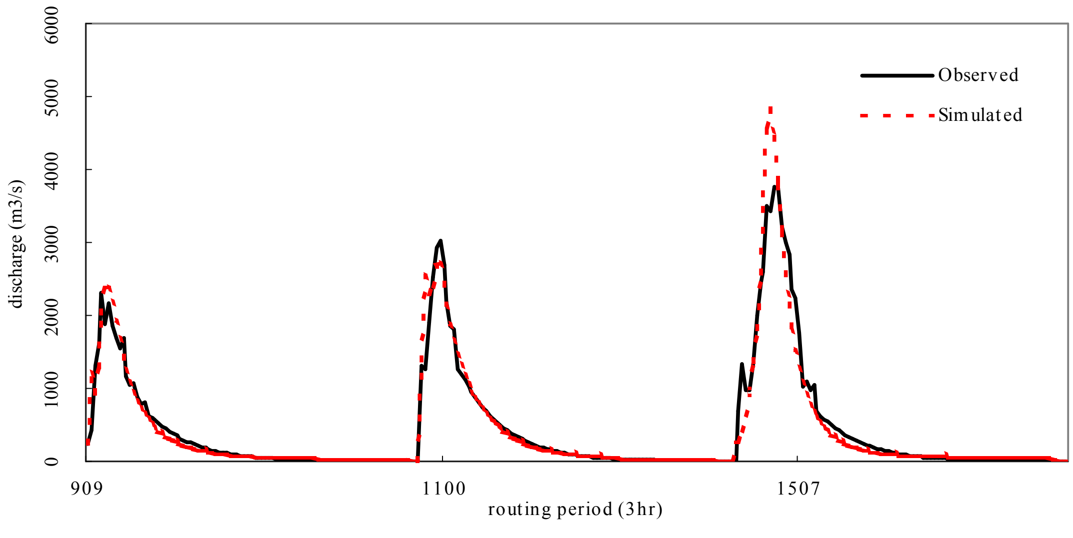

Figure 10 The simulated and observed hydrographs for 1992 during calibration

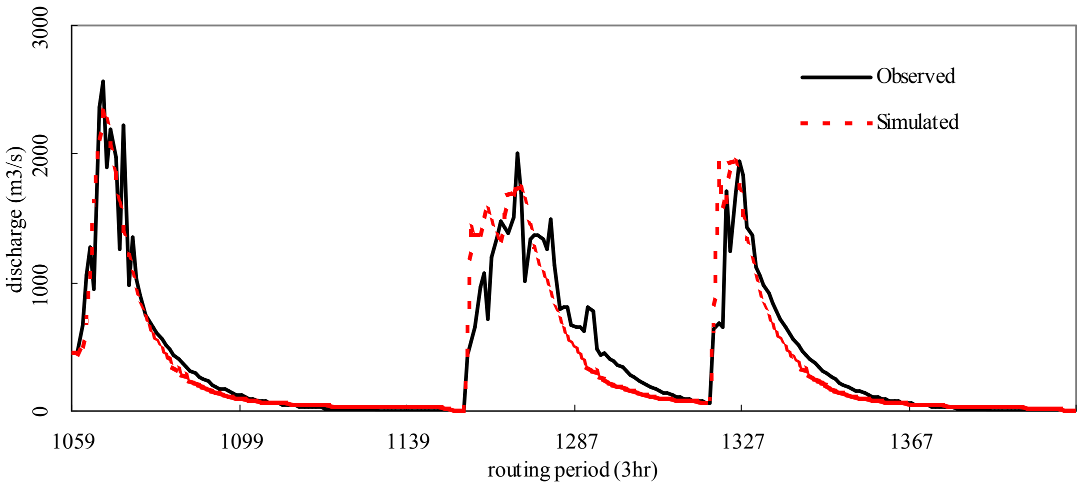

Figure 11 The simulated and observed hydrographs for 1993 during calibration

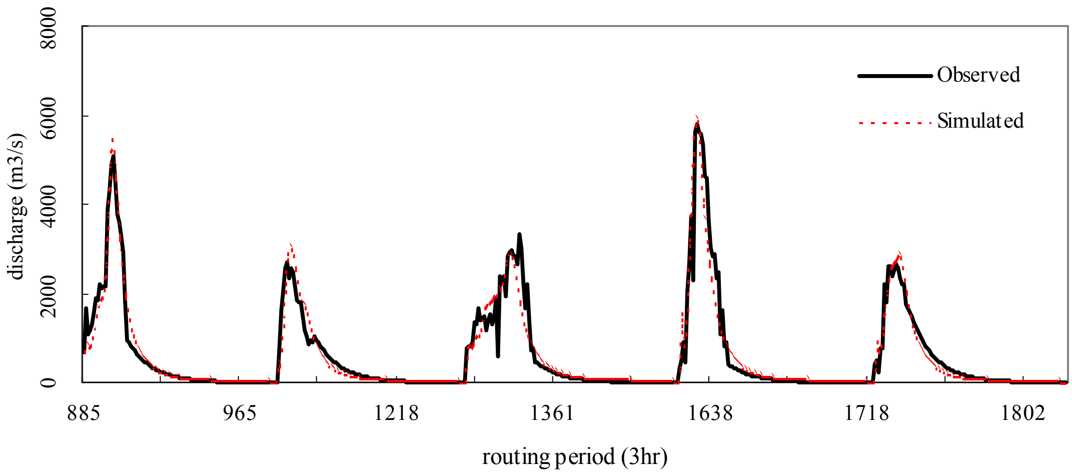

Figure 12 The simulated and observed hydrographs for 1994 during calibration

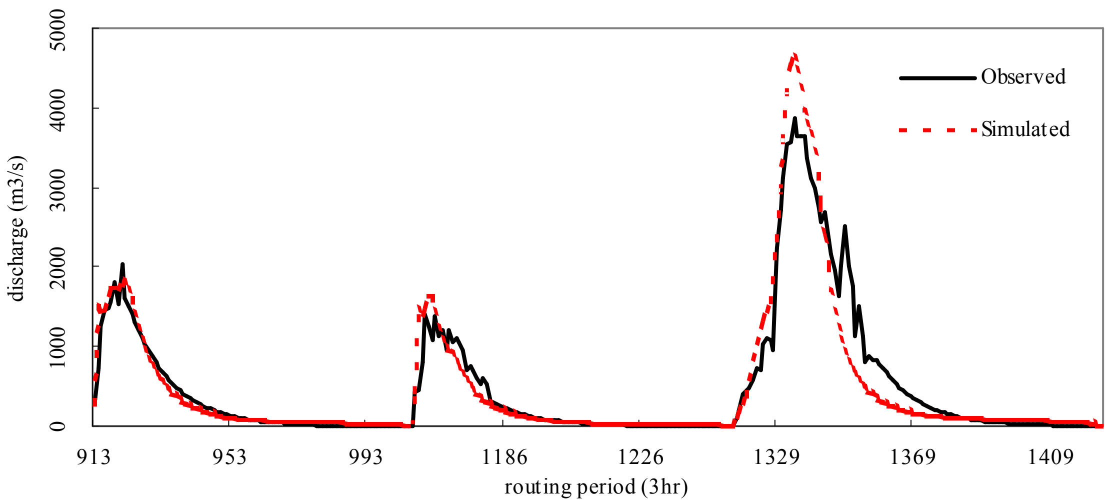

Figure 13 The simulated and observed hydrographs for 1995 during calibration

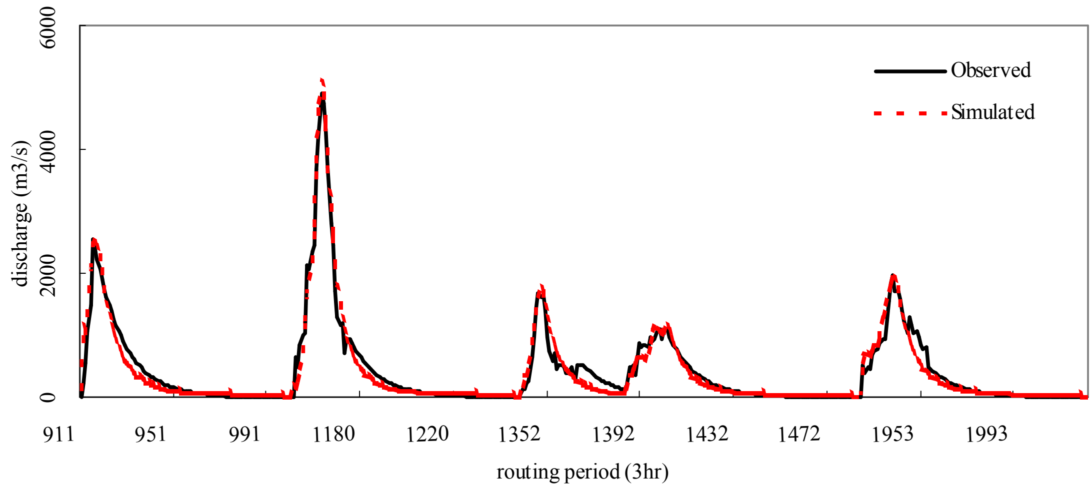

Figure 14 The simulated and observed hydrographs for 1999 during validation

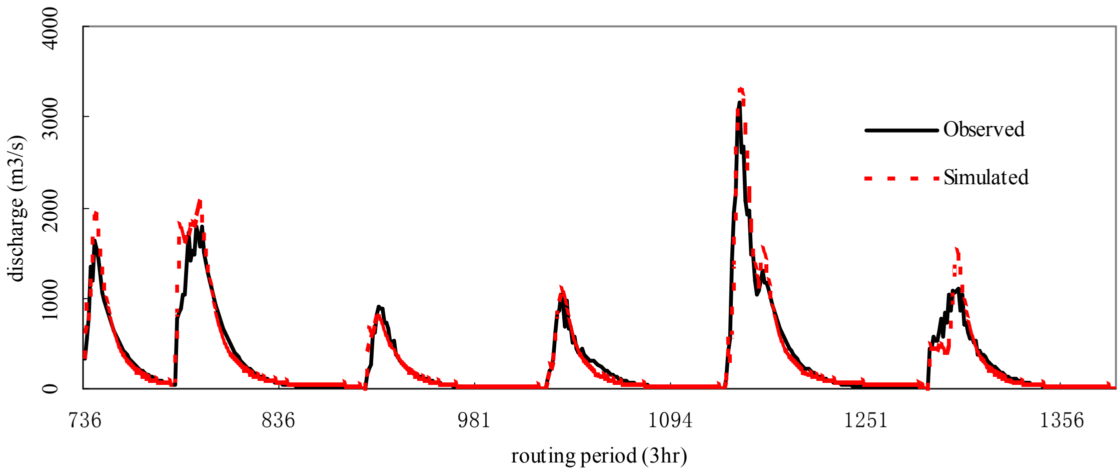

Figure 15 The simulated and observed hydrographs for 2000 during validation