

文章编号:1671-8844(2004)03-018-05

概念性流域水文模型参数优选技术研究

张洪刚, 郭生练, 王才君, 刘攀, 彭定志

(武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉 430072)

摘要:简要介绍了概念性流域水文模型参数优选的目标函数与参数优选算法,应用江西省七一水库的汛期遥测雨洪资料,分析了新安江模型参数的灵敏度,深入研究了多目标参数自动优选问题的参数解空间,着重探讨了多目标函数的组合技术及其解法.实例研究结果表明,多目标参数自动优选方法能综合考虑水文过程的各种要素,优于传统的单目标优选结果.

关键词:水文模型;参数优选;目标函数;灵敏度分析

中图分类号:P 338.8 **文献标识码:**A

Review and research of calibration techniques for conceptual hydrological model

ZHANG Hong-gang, GUO Sheng-lian, WANG Cai-jun, LIU Pan, PENG Ding-zhi

(State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China)

Abstract: Objective functions and parameter optimization algorithms for conceptual hydrological model calibration were introduced firstly. By using observed hourly rainfall and discharge data of Qiyi Reservoir basin in Jiangxi Province, the parameter space and the parameter sensitivity of the Xinanjiang model were analyzed deeply. The multi-objective functions and their combination as well as the solutions were also discussed in detail. It was found that automatic calibration using multiple objectives could consider of different aspects of the discharge hydrograph and has a better goodness-of-fit than that of the traditional single objective function.

Key words: hydrological model; parameter calibration; objective function; sensitivity analysis

1 概述

概念性流域水文模型广泛应用于洪水预报和水资源管理等众多领域,它可以帮助我们分析各种不同的信息,解决一些复杂的水资源和水环境问题.模型的模拟结果与模型结构以及模型的参数值密切相关,为此我们面临两个问题:如何选取一个适用于所选流域的水文模型;如何选择一组模型参数使得模拟结果与实测资料尽可能接近.从理论上讲,模型参数可以从流域直接或间接获得,但由于概念性水文模型参数既有其物理意义,又有其

推理概化的成分,因此大部分模型参数只能在对实测资料进行分析的基础上通过参数优选得到.在过去的 20 年中,水文工作者对水文模型参数优选方法做了大量的研究,可归结为以下四点:实测数据中各种误差的处理方法研究;寻求一种解决模型参数优选问题的最佳方法;选择多少数据以及选择什么样的数据来进行参数优选;如何有效统计分析模型结构与模型参数的不确定性,以及这种不确定性对模型输出的影响^[1].

参数优选常用手工优选(如试错法)和自动优选两种.手工参数优选算法最终优选的参数值因人

收稿日期:2002-09-22

作者简介:张洪刚(1978-),男,山东淄博人,博士研究生,从事现代水文预报方向的研究.

基金项目:国家自然科学基金(50179026)和国家基础研究“973”资助项目(G19990436).

而异,与经验和所受的训练以及对模型结构的理解程度有关,该法主要的不足之处是难以判断优选过程什么时候达到最优以及所得到的解是否为最优解.随着计算机技术的迅猛发展,参数自动优选方法得到了普遍的应用,大大加快了参数优选的速度,弥补了工作人员确乏经验的不足,增加了模拟结果的客观性与可信度,但它还不能完全替代手工优选方法,需要与手工优选方法结合使用^[2].

通过分析发现水文资料的质量对模型参数优选的影响远大于所选取水文资料的数量对模型参数优选的影响.水文资料的质量依赖于数据中所包含的有关水文过程信息的多少与数据本身存在的误差,Gupta与Sorooshian认为数据包含的信息多少取决于水文过程的变幅,如果数据涵盖了丰水、中水、枯水年,则认为数据中包含的水文信息较多.我们总是希望信息足够的多而误差尽可能的小,但实际上,无论选取怎样的数据,由于测量仪器的系统偏差、数据转化算法的误差、反推入库流量的误差等因素,误差不可避免,尤其是降雨资料往往会低估了流域上的实际降雨量^[3],因此对模型率定数据要慎重选择,需要对数据进行三性(数据的代表性、可靠性与一致性)审查.

2 目标函数

目标函数用来评价实测流量与模拟流量过程的吻合程度,不同的目标函数用来评价水文过程的不同特征,目标函数的选择对优选结果至关重要,下面简要介绍最常用的几个目标函数^[4].

(1) 总体水量误差,评价总体水量是否平衡.

$$F_1(\quad) = \frac{1}{N} \left| \sum_{i=1}^N [Q_{\text{obs},i} - Q_{\text{sim},i}(\quad)] \right| \quad (1)$$

(2) 均方误,评价实测与模拟流量过程线的吻合程度.

$$F_2(\quad) = \left[\frac{1}{N} \sum_{i=1}^N [Q_{\text{obs},i} - Q_{\text{sim},i}(\quad)]^2 \right]^{\frac{1}{2}} \quad (2)$$

(3) 实测与模拟洪水流量过程的吻合程度,其吻合程度用均方误表示:

$$F_3(\quad) = \frac{1}{M_p} \sum_{j=1}^{M_p} \left[\frac{1}{n_j} \sum_{i=1}^{n_j} [Q_{\text{obs},i} - Q_{\text{sim},i}(\quad)]^2 \right]^{\frac{1}{2}} \quad (3)$$

(4) 实测与模拟小流量过程的吻合程度,其吻合程度也用均方误表示:

$$F_4(\quad) = \frac{1}{M_1} \sum_{k=1}^{M_1} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} [Q_{\text{obs},i} - Q_{\text{sim},i}(\quad)]^2 \right]^{\frac{1}{2}} \quad (4)$$

式中: $Q_{\text{obs},i}$ 为实测流量序列, $Q_{\text{sim},i}$ 为模拟的流量过程; N 为流量序列数; M_p 为洪峰个数; M_1 为小水过程数; n_j 为第 j 个洪峰过程序列数; n_k 为第 k 个小水过程序列数; \quad 为待优选参数.

Nash 与 Sutcliffe 在 1970 年提出了模型效率系数(也称确定性系数)来评价模型模拟结果的精度,确定性系数是式(2)的另一种表现形式,它更直观的体现了实测与模拟流量过程拟合程度的好坏^[4],确定性系数公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^N [Q_{\text{obs},i} - Q_{\text{sim},i}]^2}{\sum_{i=1}^N [Q_{\text{obs},i} - \overline{Q_{\text{obs}}}]^2} \quad (5)$$

式中, $\overline{Q_{\text{obs}}}$ 表示实测流量过程的均值,其余符号同前. R^2 越大表示实测与模拟流量过程拟合的越好,模拟精度越高,另外还有洪峰合格率、峰现时差等评价指标.

3 参数优选算法

对于某一特定的流域,当水文模型和目标函数选定以后,优选算法的选择对模型参数的最终取值起决定性作用,参数优选方法一般分为局部寻优与全局寻优.

局部寻优法通过给定参数初值,给定搜索步长,通过计算新的目标函数值,逐步用较优的点代替次优的点,在给定终止条件的前提下,经过反复试算逐步确定参数最优点,比如:罗森布瑞克法,单纯形法,模式搜索法等.由于水文模型大多数是非线性的,模型的响应面是多峰的,也就是说在参数空间里有若干个局部极低点,因此局部寻优法对参数初值的要求较高,给定不同的参数初值,往往会得到不同的优选结果,因此采用局部寻优法很难确定优选结果是否为全局最优^[5].

全局最优法能有效的对参数空间内的多个极值点进行综合考虑,从整个参数空间中寻求全局最优解,全局最优法分为确定优选、随机优选及二者综合的方法,由于非线性水文模型的目标函数值的连续性不明确,因此参数优选中最常用的是随机优选法与综合优选法.如基因算法是一种基于自然基因和自然选择机制的寻优方法.该法按照“择优汰

劣'的法则,将适者生存与自然界基因变异、繁衍等相结合,从各参数的若干可能取值中,逐步求得最优值^[5]。

目前在水文模型参数优选中应用最为广泛的方法是基因法,罗森布瑞克法和单纯形法。三种优化方法中以罗森布瑞克法的运算速度最快,单纯形法次之,基因法略差;参数初值的选定对基因法的影响较小,而对罗森布瑞克法和单纯形法的影响较大;各方法以单纯形法的精度最高,罗森布瑞克法次之,基因法略差。综合上述三种方法的优点,建议以基因法的优选结果作为参数初值,然后采用罗森布瑞克法,最后再采用单纯形法进一步优化,一般可得到模型参数的最佳值^[2,6]。

4 模型参数灵敏度分析

随着水文模型的结构越来越复杂,模型参数也显著增加,典型的概念性水文模型一般有 10~15 个参数,大多数参数只能通过参数的物理意义以及流域水文地质特性粗略的给定一个取值范围,这给参数优选带来了极大的不便。为此需要对参数进行灵敏度分析,通常的方法是在保证其他参数值不变的前提下,改变一个或几个参数值,通过目标函数值的变化,来分析参数对模拟结果的灵敏度。本文选取江西省七一水库 1995~2000 年汛期遥测雨洪资料,以新安江三水源模型为例(计算时段长取 1 h)进行分析。

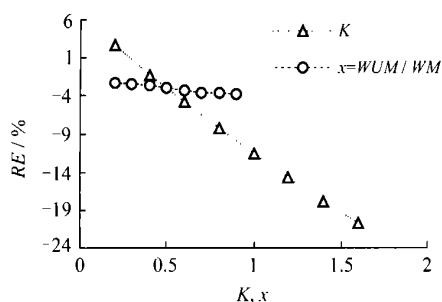


图1 不同的 K 与 x 对水量相对误差的影响

5 多目标参数优选技术

传统的单目标参数优选往往不能恰当的描述由观测资料所反映出来的各种水文特征。因此需要研究探讨多目标参数自动优选方法,多目标参数优选方法在模型率定过程中,通过对实测资料进行更细致的分析,大大地改善了模型参数的优选问

题。新安江三水源模型的参数可以分为四个层次,蒸散发参数: K 、 WUM 、 WLM 、 C ; 产流参数: WM 、 B 、 IM ; 分水源参数: SM 、 EX 、 KG 、 KI 与汇流参数: CG 、 CI 。上述各层次的顺序,就是模型的计算步骤,根据上一层的计算结果可以作下一层的计算,各层参数有各自的作用与目标,因此相互之间的独立性较好,而同层次中的各参数由于目标相同,相互之间的相关性就大一些,需要进行仔细分析^[7]。

在第一层中,增大 WUM 、 WLM 的值会对 K 产生影响,但由于 WUM 、 WLM 的值有一定的变化范围,因此这种影响是有限的,图 1 绘出了不同的蒸散发折算系数 K 对与不同的上层土壤蓄水量系数 x ($x = WUM / WM, 0 < x < 1.0$) 条件下,模拟结果的水量相对误差 (RE),可以看出 K 对水量相对误差的影响较大,而 WUM 的影响很小,同理可证明 WLM 、 WM 对模拟结果的影响也不显著。在第三层中,对表层土自由水容量 SM 与表层土自由水蓄水库对地下水的出流系数 KG 分别取不同的值作网格交叉计算,结果见图 2,可以看出 SM 与 KG 对模型确定性系数 (R^2) 的影响都较大,属于敏感参数。同理对其他参数进行分析,发现表层土自由水蓄水库对地下水的出流系数 KI 、地下水库的消退系数 CG 、壤中流的消退系数 CI 对模拟结果的影响也较大,需要仔细优选;其他参数的影响不显著,可根据一般经验确定,不必参加仔细优选^[2,7]。

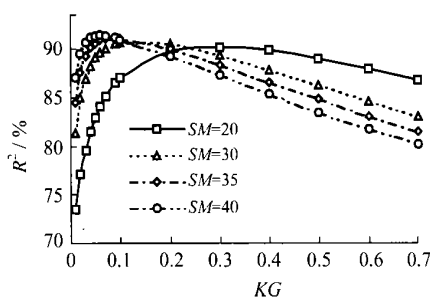


图2 SM 与 KG 对确定性系数的影响

题^[8]。

5.1 多目标问题的求解

目前直接求解多目标问题还有困难,于是人们想办法将其转化为较易求解的单目标问题,由于简化途径不一而形成了多种方法,下面简要介绍常用的几种。

(1) 平方和加权法

给定 m 个规定值,要求 m 个目标函数分别与

规定值相差最小,可采用下述评价函数.

$$U(x) = \sum_{i=1}^m [f_i(x) - f_i^*]^2 \tag{6}$$

式中: $U(x)$ 为综合目标函数; $f_i(x)$ 为第 i 个目标函数; f_i^* 为规定值.

(2) 多目标距离函数法

在利用多目标参数自动优选方法率定模型时, 关键在于如何综合考虑各目标函数间的平衡关系, 令 $i = 1, f_i^* = -A_i$, 对 $U(x)$ 开平方, 则式 (6) 变为

$$F(x) = [F_1(x) + A_1]^2 + [F_2(x) + A_2]^2 + \dots + [F_p(x) + A_p]^2 \tag{7}$$

式中, A_i 对应于每一个目标函数权重, 取决于函数本身, 对不同的目标函数赋不同的 A_i 值进行参数优选, 使式 (7) 中 $[F_i + A_i]$ 项到原点的距离相等, 该方法称为距离函数法^[4,8], A_i 值由下式给出.

$$A_i = \max_{j=1,2,\dots,p} \{F_{j,\min} - F_{i,\min}\} \tag{8}$$

此外还有分层序列法、理想点法、乘除法、功效系数法及直接求非劣解法等, 在此不再介绍.

5.2 多目标距离函数法应用结果分析

对多目标参数优选来说, 不仅要注重各参数之间的相互独立性, 而且要注意各目标函数之间的相互独立性. 选取江西省七一水库 1995 ~ 2000 年汛期遥测雨洪资料 (根据水库流域特征, 以实测流量 100 m³/s 为洪峰流量过程, 以实测流量 50 m³/s 为小水流量过程), 以新安江模型为例, 计算时段长取 1 h. 采用 $F_1(\)$ 、 $F_2(\)$ 、 $F_3(\)$ 、 $F_4(\)$ 为目标函数, 选用基因法, 罗森布瑞克法和单纯形法作为优选方法, 对各种目标函数的不同组合分别进行参数优选, 以水量相对误差 (RE)、确定性系数 (R^2)、洪峰合格率 (DQ) 以及小水过程的 RMSE 四个指标分别评价以上各目标函数的效果. 优选结果见表 1, 表中列出了不同的目标函数组合条件下的模拟结果的统计指标, 从表 1 中可以看出: 采用多目标函数能比较全面的考虑到流量过程的各个要素, 考虑目标函数越多, 就越全面的反映流量过程的各方面特性, 但没有哪一个解的所有评价指标比其他解绝对最优, 同时可以看出采用距离函数法比不采用的情况能更好的协调各目标函数间的平衡约束关系.

表 1 新安江模型多目标函数组合及参数优选结果比较

方案	目标函数	综合目标函数	$R^2/\%$	RE/%	DQ/%	小水均方误
1	$F_1(\)$ 、 $F_2(\)$	$A_i = 0$	92.38	3.07	91.7	12.50
2		(7), (8)	92.06	1.03	100	13.57
3	$F_3(\)$ 、 $F_4(\)$	$A_i = 0$	92.07	8.82	100	12.31
4		(7), (8)	92.27	8.92	100	11.74
5	$F_1(\)$ 、 $F_2(\)$ 、 $F_3(\)$	$A_i = 0$	92.33	3.06	100	12.75
6		(7), (8)	92.20	0.74	100	13.09
7	$F_1(\)$ 、 $F_2(\)$ 、 $F_4(\)$	$A_i = 0$	92.25	4.32	100	11.55
8		(7), (8)	91.76	0.63	100	12.22
9	$F_1(\)$ 、 $F_2(\)$ 、 $F_3(\)$ 、 $F_4(\)$	(7), (8)	92.23	1.60	100	11.97

图 3 和图 4 分别给出了方案 6 和方案 9 的目标函数值的响应面, 从图 3、4 中可以看出目标函数值的响应面是多峰的, 说明各目标函数值之间互相制约, 也就是说并没有一个点比其他所有点绝对最优, 问题的解并非唯一, 而是由许多可行解组成的集合, 即 Pareto 解集, 在参数解空间里寻求 Pareto 解集, 避免了由于人为主观因素选择一个“最优解”, 而损失其他的“最优解”. 对于 Pareto 解集的任意解 i 有以下规律:

- (1) 对于所有支配解 j , Pareto 解集合中至少有一个解 i , 使 $F_k(i) < F_k(j)$ ($k = 1, 2, \dots, p$).
- (2) 在 Pareto 解集合中不存在最优解 j , 使得

$$F_k(j) < F_k(i), (k = 1, 2, \dots, p).$$

图 5 绘出了七一水库 1999 ~ 2000 年 5 ~ 6 月份的实测流量过程以及不同的参数解空间所对应的模拟流量过程, 为了对比明显, 对流量 Q_i 取对数: $\ln(Q_i)$, 这样做的目的是改变流量的相对大小, 并不改变图 5 所要表达的信息. 图 5 中灰色部分对应模型参数解空间所模拟的流量过程外包线, 黑色部分对应 Pareto 解 (表 1 中所有方案) 所模拟的流量过程外包线, 该部分的流量过程线都从某些方面较好的模拟了实测流量过程, 白色实线为实测流量过程. 从图 5 中可以直观的看出 Pareto 解与模型参数空间的关系, 同时可以看到实测流量过程有一部

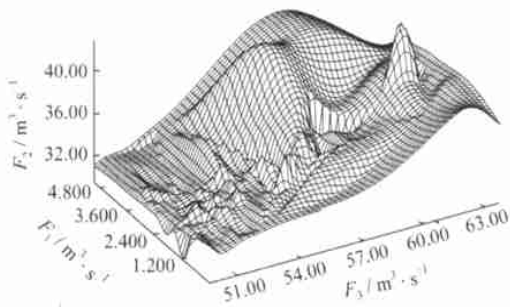


图3 方案6的目标函数值响应面

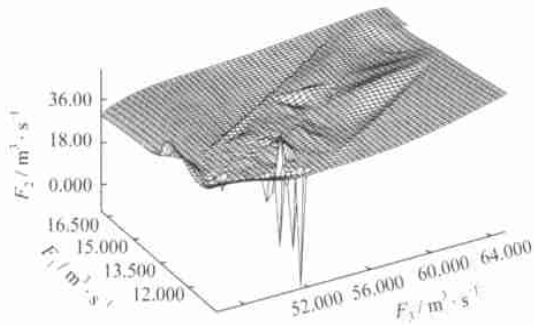


图4 方案9的目标函数值响应面

分并没有包含在 Pareto 解所对应的模拟流量的外包线内,原因可能有三点: 模型的结构存在误差;

选择目标函数时加入的人为因素的干扰; 由于实测资料的误差,所引起的统计值的不确定性.

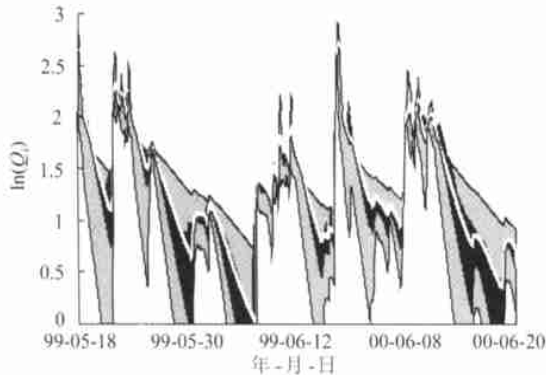


图5 三水源新安江模型参数空间与 Pareto 解集对应的流量过程

对 Pareto 解集中的每个解按优劣进行明确的排列是不可能的,为此引入次等解-优越解的概念,采用 Pareto 解集分类排序法来解决这个问题,该法的思路是: 根据具体的要求给出一定的补充评价标准,确定 Pareto 解集中倾向于该评价标准的所有非支配解,并将这些非支配解的序号定义为 1; 在扣除序号为 1 的非支配解后剩余的群体中确定次优的解,将这些解的序号定义为 2; 重复步骤,直到群体中的所有解都得到一个排序号. Pareto 解集分类排序算法本质上对 Pareto 解集中的解按一定标准分成不同的组,序号越小的组越接近于 Pareto 最优解,序号越大的组越偏离 Pareto 最优解,同一组中的解对于同一个评价标准具有相同的优属度.综上所述,多目标参数优选方法的求解通常由以下两个步骤来实现:

(1) 在参数可行空间里面通过自动优选来确定

Pareto 解空间,同时利用选出的不同的评价标准对模拟与实测数据的不同方面的拟合程度进行评价.

(2) 将 Pareto 解空间不合适的解排除掉,再利用补充的评价标准进一步缩小 Pareto 解空间.

6 结 论

在过去的 20 多年里,概念性流域水文模型的参数优选技术取得了较大进展,综合以上分析,可以得出以下五点结论:

(1) 随着计算机科学的迅猛发展,参数自动优选技术得到了广泛的应用,减少了人们的工作量和由于经验不足所带来的麻烦,提高了参数优选的速度与模拟的精度.

(2) 随着水文遥测系统的建立以及 GIS 技术在水文中的应用,水文资料的准确性、实时性都有所提高,但资料的误差不可避免,必须对资料进行仔细的分析与修正后才可用于参数优选.

(3) 由于多数水文模型的目标函数响应面是多峰的,局部寻优法往往不能确定优选结果是否为全局最优,因此一般需要先利用全局优选法逐步缩小参数空间,再用局部寻优法进一步寻求最优解.

(4) 对模型参数进行灵敏度分析可以减少优选参数的个数,提高参数优选的速度与效率.

(5) 多目标参数优选方法较好的考虑了流量过程的各个方面,能够取得较好的模拟结果,优于传统的单一目标函数优选方法,考虑到模型结构的误差与实测数据存在的系统偏差同时存在,模型参数解区域只能逐步缩小至接近唯一的全局最优解,而不可能取得全局最优解,从应用情况来看,参数解区域一般会维持在一个有限的范围,即 Pareto 解.

(下转第 26 页)

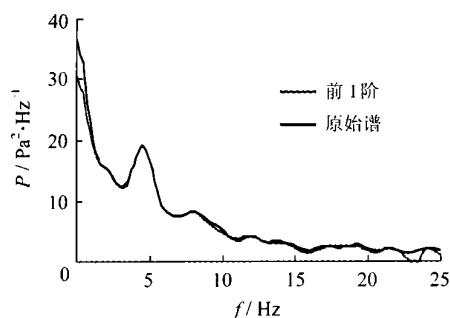


图6 前1阶模态重构谱与原始谱的比较

表现为窄带低频过程。

(3) 冲击射流的脉动壁压场可用正交分解进行描述,其特征谱与特征谱密度函数皆随阶数的增加而迅速衰减。

(4) 对冲击射流的脉动壁压谱密度函数进行低阶近似非常有效,经过前两阶近似所得到的谱密度函数即与原始的谱密度函数差别甚小。

参考文献:

- [1] 许多鸣,余常昭. 平面水射流对槽底的冲击压强及其脉动特性[J]. 水利学报, 1983(5): 52-58.

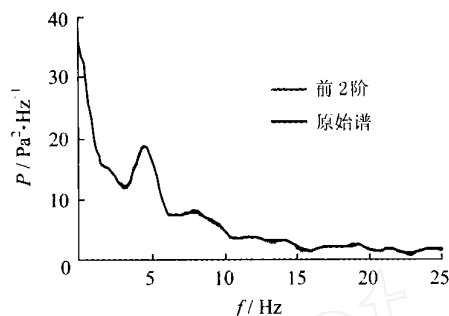


图7 前两阶模态重构谱与原始谱的比较

- [2] 柴华,冬俊瑞,李永祥. 热膜测速技术在挑射水流运动特性及高坝消能机理研究中的应用[J]. 水利学报, 1999(10): 45-51.
- [3] Glauser M N, Leib S J, George W K. Coherent structures in the axisymmetric turbulent jet mixing layer [A]. Turbulent Shear Flows 5 [C], Berlin, Springer-Verlag, 1987. 134-145.
- [4] 刘士和. 湍流结构的正交分解与低阶近似及湍流相干模式的识别[J]. 武汉水利电力大学学报, 2000, 33(2): 1-5.

(上接第22页)

参考文献:

- [1] Singh V P, Frevert D K. Mathematical Models of Large Watershed Hydrology [M]. Water Resources Publications, U. S, 2002.
- [2] 郭生练. 水库调度综合自动化系统[M]. 武汉: 武汉水利电力大学出版社, 2000.
- [3] Gupta V K, Sorooshian S. The relationship between data and the precision of estimate parameters[J]. Journal of Hydrology, 1985, 81: 55-77.
- [4] Madsen H. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives[J]. Journal of Hydrology, 2000, 235: 276-288.
- [5] Wang Q J. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models[J]. Water Resources Research, 1991, 27(9): 2467-2471.
- [6] 谭炳卿. 水文模型参数自动化优选方法的比较分析[J]. 水文, 1996(5): 8-14.
- [7] 赵人俊, 王佩兰. 新安江模型参数的分析[J]. 水文, 1988(6): 2-8.
- [8] 张洪刚, 郭生练, 刘攀, 彭定志. 概念性水文模型多目标参数自动优选方法研究[J]. 水文, 2002(1): 12-16.