

基于 SUFI-2 算法和 SWAT 模型的妫水河流域水文模拟及参数不确定性分析

李 谦^{1,2,3}, 张 静^{1,2,3}, 宫辉力^{1,2,3}

(1.首都师范大学城市环境过程与数字模拟国家重点实验室培育基地, 北京 100048;

2.首都师范大学资源环境与地理信息系统北京市重点实验室, 北京 100048;

3.首都师范大学三维信息获取与应用教育部重点实验室, 北京 100048)

摘 要:敏感性分析和不确定性分析是分布式水文模型参数校准和模型应用的基础。本文以妫水河流域为例,基于 SWAT 模型和 SUFI-2 算法进行模型参数敏感性分析,结合手动调参和自动率定,通过SUFI-2 的 P 因子和 R 因子进行模型不确定性分析,构建妫水河流域分布式水文模型。本次妫水河流域月尺度模拟中:率定期,确定系数 $R^2=0.59$,效率系数 $NSE=0.56$;验证期,确定系数 $R^2=0.82$,效率系数 $NSE=0.80$; P 因子均大于 0.5, R 因子均小于 0.5。结果表明,妫水河流域 SWAT 模型水文模拟效果较好。

关键词:SWAT;SUFI-2;敏感性分析;不确定性分析

中图分类号:P338

文献标识码:A

文章编号:1000-0852(2015)03-0043-06

随着社会经济的快速发展,我国水资源短缺问题愈发突出。延庆县以农业为主,是北京市主要的农产品供应区之一。妫水河是延庆县主要河流,是城市生活用水和工农业用水的主要来源之一。近年来,由于人口增加、经济发展、人类活动增强以及持续的干旱气候条件,使得水资源日益减少,水资源短缺严重,供需矛盾突出^[1-2]。精确模拟径流过程是实施水资源管理的重要工作。而基于物理机制的分布式水文模型为径流模拟和流域水文循环研究提供了有效工具^[3]。SWAT 模型(Soil and Water Assessment Tool)是由美国农业部(USDA)农业研究中心(ARS)开发的基于物理过程,并能够模拟不同土地利用和多种农业管理措施对流域水、泥沙、营养物质、杀虫剂等输送迁移影响的分布式流域水文模型^[4],在中国和世界各地得到了广泛应用^[5]。

水文模型的非线性和参数的相关性导致模型解空间可能会存在多个局部最优解,模型率定结果往往存在很大的不确定性。水文模型的不确定性主要包括模型输入的不确定性、系统结构的不确定性和系统输出的不确定性^[6]。由于模型不确定性较为复杂,同时对于模型不确定性的分析并不能直接提高模拟结果的精

度,所以已有研究一般只通过目标函数来检验模型在流域的适用性,并未对模型不确定性进行分析^[7-8]。

本文在北京市延庆县妫水河流域构建 SWAT 分布式水文模型,对 2005~2008 年的月径流量进行连续模拟,并结合 SWAT 和 SUFI-2 算法进行敏感性分析、参数率定和验证以及不确定性分析,探索 SWAT 模型在妫水河流域的适用性和模型参数的不确定性。

1 研究区概况与基本数据

妫水河流域(图 1)位于北京市延庆县,属于海河流域永定河水系,是官厅水库的上游河流之一。流域地势东北高,西南低,东、南、北三面环山,西面为官厅水库。本文研究区选自东大桥水文站以上流域,集水面积 594km²。处北纬 40°22′23″N~40°38′27″N,东经 115°49′3″E~116°20′32″E。

研究区是大陆性季风气候,属温带与中温带、半干旱与半湿润带的过渡地带,气候冬冷夏凉。根据所得资料统计研究区年平均气温 9.8℃,年平均降水 400mm,年最大降水 523mm(2008 年),最小 375mm(2005 年),降水多集中在 5~8 月。

收稿日期:2013-10-10

基金项目:国家自然科学基金(4127004);北京市科技新星项目(2010B046);北京市自然科学基金(8123041)

作者简介:李谦(1983-),男,湖南邵阳人,硕士研究生,主要从事 GIS 和遥感在水文方面的应用研究。E-mail:cnuliqian@gmail.com

本研究所用的数据包括:90m×90m 分辨率 DEM, 2005 年土地利用数据 (图 2), 1995 年的土壤类型图 (图 3) 及土壤属性数据, 研究区 6 个气象站点的日降水、日最高最低温度、日相对湿度、日平均风速、日照时数以及 1 个水文观测站的日径流量数据。各站点数据均从 2005 年 1 月到 2008 年 12 月。各站点位置如图 1 所示。根据现有数据, 确定 2005 年为模型预热期, 2006~2007 年为模型率定期, 2008 年为模型验证期。

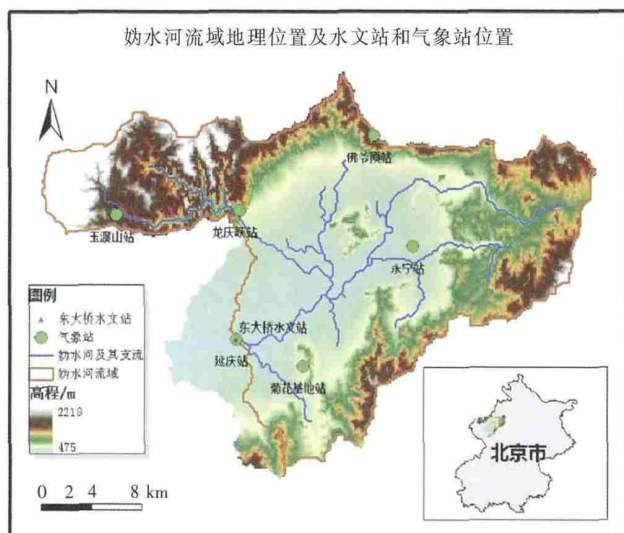


图 1 研究区地理位置及数据站点分布

Fig.1 Location of the study area and distribution of the data sites

2 研究方法

本研究分为三个部分:敏感性分析、率定以及率定

后的不确定性分析。其中, SWAT 的敏感性分析属于率定前就能进行的敏感性分析。SUFI-2 算法的敏感性分析分为两类: 第一类是率定前对单个参数进行分析; 第二类需要进行若干次迭代运算, 它敏感性是本次迭代结果, 作为下次迭代的参考。

2.1 SWAT 模拟及敏感性分析方法

SWAT 通过加载奶水河流域的 DEM 数据, 将流域自动划分为若干个子流域。在子流域的基础上进行土地利用、土壤和坡度的叠加分析, 计算出研究区的水文响应单元。在此基础上, 利用气象数据, 模拟运算出东大桥水文站的径流量。

SWAT 模型采用 LH-OAT 敏感性分析法。该方法具有 LH 采样法 (Latin Hypercube Sampling) 和 OAT (One-At-a-Time) 敏感度分析的优点^[6]。

LH 采样法采用分层式采样, 可以用最少的采样数最优地覆盖采样立方。因此, LH 采样法比随机采样法要求有更高效率的输出统计估算。其采样过程如下: 首先将每个参数空间分成 n 份, 且每个采样范围出现的概率都是 $1/n$; 其次, 生成参数的随机值, 并确保任一采样范围仅采样一次; 最后, 参数随机组合, 模型运行 n 次, 其结果进行多元线性回归分析。

OAT 敏感度分析方法是指模型运行 $m+1$ 次, 以获取 m 个参数中某一特定参数的敏感度。其优点在于模型每运行一次仅一个参数存在变化而其他参数视为不变。因此, 该方法可以清楚地将输出结果的变化明确地归因于某一特定输入参数值的变化^[9-10]。

最后, 参照敏感度取值范围表 (表 1), 可以确定参

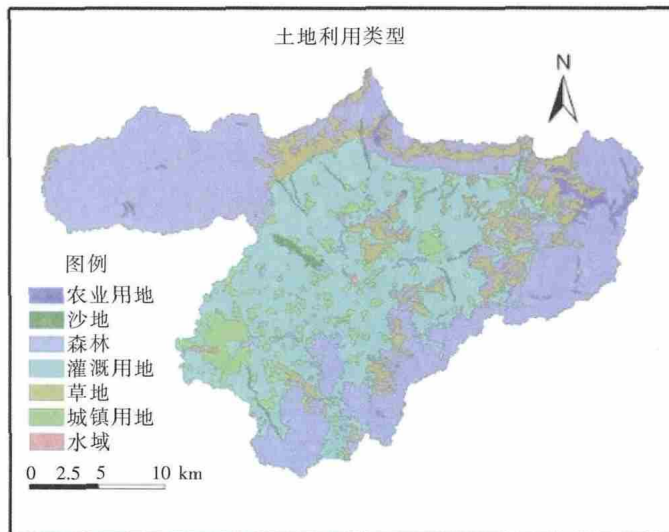


图 2 土地利用类型图

Fig.2 Land use of the study area

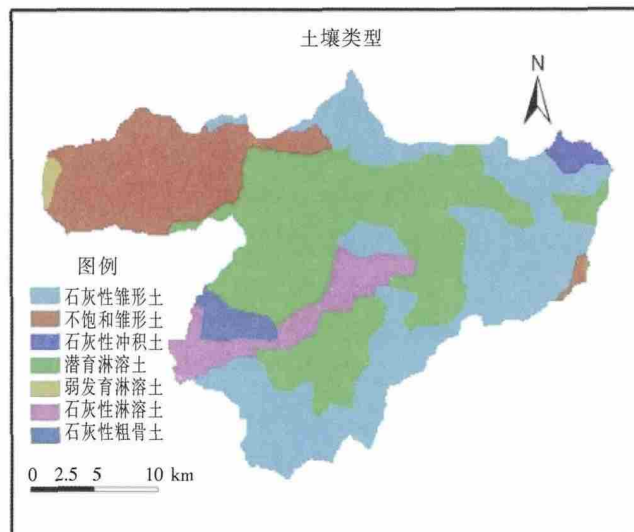


图 3 土壤类型图

Fig.3 Soil type of the study area

表1 敏感度取值表^[11]

Table1 Sensitivity values

分类	敏感度取值	敏感度等级
I	<0.05	低
II	0.05~0.2	中
III	0.2~1.0	高
IV	>1.0	很高

数敏感度高低。

本文采用了 SWAT 自带的敏感性分析功能与 SUFI-2 的敏感性分析结果进行对比参考。

2.2 SUFI-2 敏感性分析方法

在 SUFI-2 中,有两类共 3 种方法得到参数的敏感性^[6]。

第一种方法是“one-at-a-time”(OAT)。每次仅计算一个参数的敏感性,同时把其他参数视为不变化。这个方法有两个优点:一是可以检验 SWAT 的敏感性分析结果;二是可以在整体率定之前就得到参数的敏感程度。

第二种方法是全局敏感度。这种方法是在率定的过程中,计算下次率定时所需要的参数敏感性。在统计学中, T 统计量(T -stat)假设检验样本时用于与临界值相比,越大越好; T 检验值(T -Test)用来确定每个样本的相对显著性; P -value(P 值)是 T 检验值查表对应的 P 概率值, P 值体现了 T 统计量的显著性。该方法采用 T -state 的绝对值作为敏感性的参考,参数 T -stat 的绝对值越大,敏感性越高;同时采用 P -value 来指示 T -state 的显著性。参数的 P 值越接近 0,显著性越大^[6]。

第三种方法是观察散点图。散点图列出了每次模拟的目标函数值,可以把参数的敏感性区间缩小到目标函数最符合要求的区间。下文中将给出本次方法的散点图实例。

2.3 率定结果分析方法

选取基于观测值和模拟值计算的确定性系数 R^2 及其修正参数 bR^2 、Nash-Sutcliffe 效率系数 NSE 和标准误差 $RMSE$ 来综合评价 SWAT 模型模拟效果。

决定系数 R^2 计算公式如下:

$$R^2 = \frac{\left(\sum_{i=1}^n (Q_{s,i} - \bar{Q}_s)(Q_{o,i} - \bar{Q}_o) \right)^2}{\sum_{i=1}^n (Q_{s,i} - \bar{Q}_s)^2 \sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2} \quad (1)$$

由于决定系数 R^2 具有一些缺陷,因此同时采用它的修正公式 bR^2 ^[12]。决定系数 R^2 与回归线系数 b 的乘

积(记作: Φ)

$$\Phi = \begin{cases} |b| R^2 & \text{if } |b| \leq 1 \\ |b|^{-1} R^2 & \text{if } |b| > 1 \end{cases} \quad (2)$$

Nash-Sutcliffe 效率系数 NSE 计算公式如下:

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2} \quad (3)$$

标准误差计算公式如下:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{o,i} - Q_{s,i})^2} \quad (4)$$

式中: n 为观测数据的总数; i 为观测数据的次序; $Q_{s,i}$ 为第 i 次的模拟值; $\bar{Q}_s$ 为模拟值的平均值; $Q_{o,i}$ 为第 i 次观测值; $\bar{Q}_o$ 为观测值的平均值。

对于决定系数, $R^2=1$ 时,表明观测值与模拟值的变化趋势完全一致, R^2 偏离 1 越远,表明观测值与模拟值吻合程度越低。对于其修正系数 Φ , b 为回归系数, bR^2 为确定性系数与回归系数之积,能够解释观测值与模拟值数值大小的差异(由 b 体现)以及变化(由 R^2 体现)。

对于效率系数, NSE 越接近 1,表明模拟值越接近观测值; NSE 越偏离 1,表明模拟值越偏离观测值。

对于标准误差, $RMSE$ 等于 0 时,表明观测值与模拟值完全符合; $RMSE$ 值越大,观测值与模拟值偏差越大。

2.4 SUFI-2 不确定性分析方法

SUFI-2 算法作为参数估计的最优化方法,考虑了输入数据、模型结构、参数及观测数据等的不确定性,并将其反映在率定后的参数范围内。参数率定后 95% 置信水平上的不确定性区间 (95% prediction uncertainty, 95PPU) 包含大多数观测数据,模拟结果总的不确定性,在 2.5% (L95PPU) 和 97.5% (U95PPU) 上的累积分布得到,通过拉丁超立方采样计算输出^[6]。

SUFI-2 算法采用两个指标来衡量参数的不确定性。 P -factor: 95PPU 区间包含观测数据的百分数,其物理意义是 95% 置信区间包含的观测数据量,它可以几乎没有观测数据或者几乎包含全部的观测数据; R -factor: 由观测数据标准差划分的 95PPU 区间的平均宽度,其物理意义是 95% 置信区间内样本的密集程度,宽度越大,样本分布越零散,宽度越小,样本分布越集中。理论上, P -factor 的范围是 0~100%,。同时 R -factor 的范围是 0~ ∞ 。 P -factor 接近 1, R -factor 接近 0 是完全接近于观测数据的模拟^[13]。

表2 部分参数敏感性分析结果

Table2 The results of the parameter sensitivity analysis

参数名称	物理意义	SWAT 敏感度	SWAT 排序	<i>T</i> -stat 绝对值	<i>P</i> -value
Alpha_Bf	基流衰退常数	2.27	2	0.03	0.98
Canmx	最大冠层截留量	1.1	3	0.17	0.87
Cn2	初始 SCS 径流曲线值	4.52	1	0.07	0.95
Esco	土壤蒸发补偿系数	0.77	4	0.73	0.48
Sol_Awc	土壤层有效水容量	0.57	6	0.43	0.67
Sol_Z	土壤表层到底层的深度	0.65	5	0.34	0.74

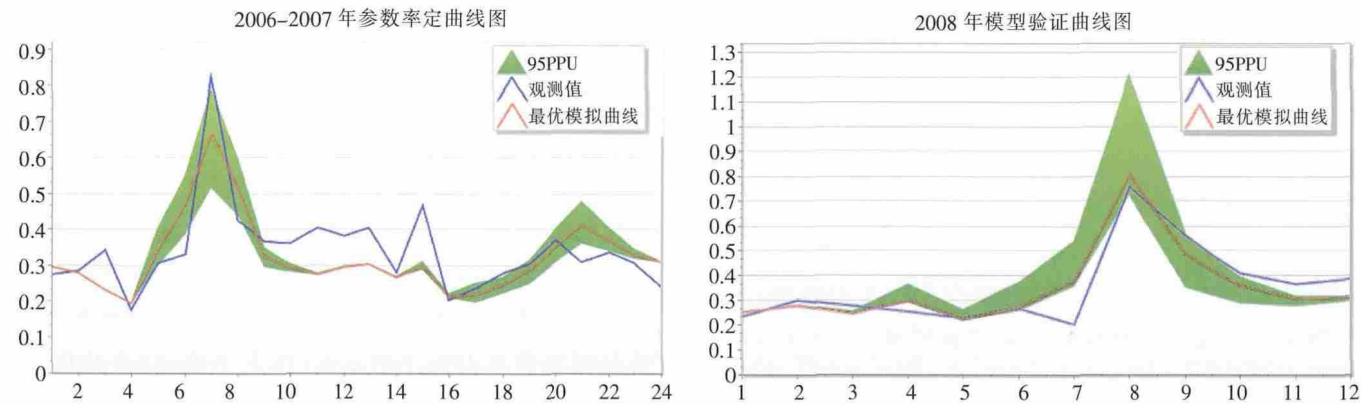


图 4 验证期和率定期东大桥水文站观测值与模拟月径流量过程

Fig.4 The observed and simulated monthly discharges at the Dongdaqiao station

3 结果分析

3.1 参数敏感性分析

SWAT 自带的敏感性分析是在率定之前得到结果的,SUFI-2 的全局敏感性分析是率定之后得到结果的。所以在 SUFI-2 率定之前,对全部参数的敏感性分析,只能参考 SWAT 的敏感性分析结果。

同时,SUFI-2 的 OAT 方法可以在率定之前确定单个参数的敏感性。因此,利用该方法对 SWAT 敏感性分析结果进行分析,可以得到更好的结果。部分参数敏感性分析结果见表 2。

3.2 参数率定及模型验证

本文的率定期为 2006~2007 年,采用 2005 年为预热期,以降低初始条件的影响,预热期在具体分析计算式不予采用。验证期为 2008 年。率定期和验证期月径流量模拟值与观测值过程线总体拟合程度较好,见图 4。率定及验证评价指标通过 SUFI-2 算法实现,结果见表 3。

以上数据显示参数验证以及模型率定结果满足实验要求。

表3 东大桥水文站月径流量模拟结果

Table3 The simulation results for the Dongdaqiao station

	<i>R</i> ²	<i>bR</i> ²	<i>NSE</i>	<i>RMSE</i>
率定期	0.59	0.38	0.56	0.07
验证期	0.82	0.75	0.80	0.05

3.3 不确定性分析

对于水文模拟的不确定性而言,参数取值范围的变化会影响分析结果,较小的参数范围能获取较窄的不确定性区间,并能提高模拟的置信水平,但却会降低参数变异的敏感性,并导致大部分观测数据落在不确定性区间之外;相反,较大的参数范围能够反映出参数对模拟结果的影响,但是由此产生较宽的不确定性区间,降低了模拟的置信水平。

本文的 *P*-factor 和 *R*-factor 如表 4 所示。本文采

表4 不确定性分析结果

Table4 Results of uncertainty analysis

	<i>P</i> -factor	<i>R</i> -factor
率定期	54%	0.07
验证期	67%	0.05

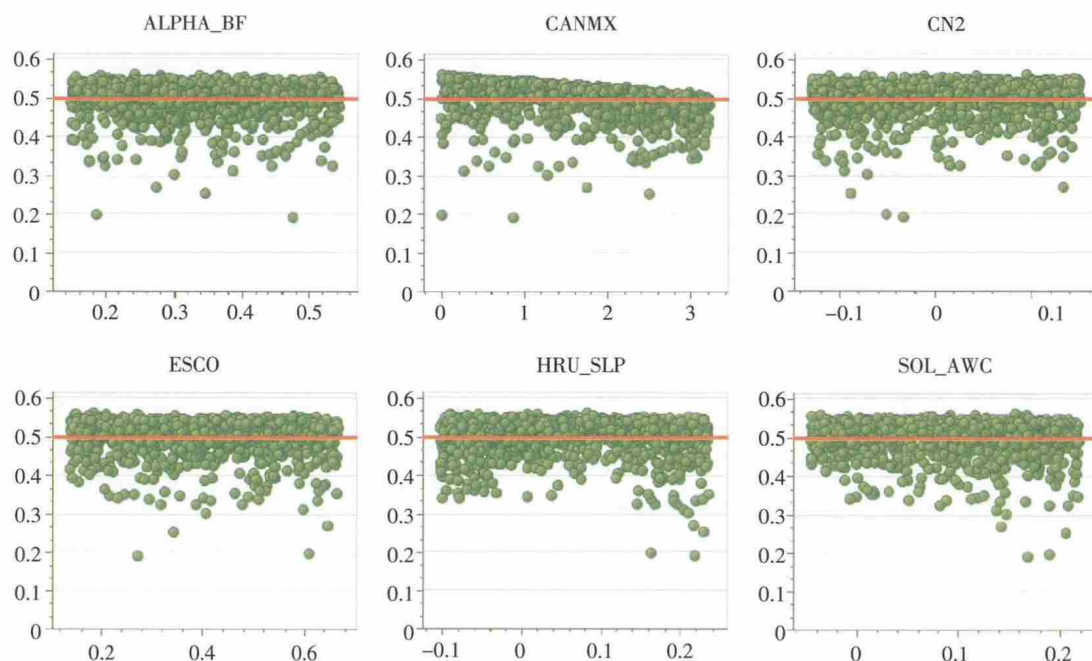


图5 部分参数率定期 1000 次模拟散点图

Fig.5 1000 simulations dot plots for some validation parameters

用的 2.5%~97.5% 区间,是 SUFI-2 默认的 95% 置信区间(95PPU)。较大的 P -factor 意味着较低的不确定性,率定期 0.54 和验证期 0.67 表明本次模拟包含了 95PPU 区间 54% 和 67% 以上的观测数据。同时, R -factor 在率定期和验证期分别为 0.25 和 0.46,表明 95PPU 宽度较窄(同时该宽度也可以从图 4 的绿色部分看出^[6])。

其次, SUFI-2 的散点图,也可以看出每次模拟对应的取值。根据散点分布的集中程度可以看出不确定性的。图 5 是率定期部分参数 1000 次模拟的散点图。

其中,横坐标是各个参数最终的取值范围,纵坐标是 NSE 值。红线是 0.5 的阈值,阈值以上部分越多,置信水平越高,模拟的不确定性越小^[14]。

另外,由 CANMX 的散点图可以看出,若是取小的参数范围(0~1),则能提高置信水平,降低参数的不确定性(R -factor),但是参数变异性(P -factor)就变小了,即参数的可取值范围变小了。

分析结果表明,这个水文站径流模拟的不确定性较小。

4 结论

本文在舂水河流域构建了 SWAT 模型,对研究区的水文过程进行了模拟,并对该模型进行了参数的敏

感性分析、参数率定、模型验证以及不确定分析。主要结论如下:

(1) SUFI-2 的参数敏感性分析必须要依靠 SWAT 的敏感性分析结果,否则初次输入参数具有一定的盲目性,除非一次性输入全部参数。

(2) 参数敏感性分析结果表明,土壤(.sol)参数对径流影响较大,另外 Cn2、Alpha_bf 对径流影响也很敏感。

(3) 东大桥水文站径流量的模拟值与观测值总体拟合程度较好,验证期比率定期高,主要原因是验证年份少,参数拟合相对容易。

(4) 对于不确定性,一般而言,可以通过较大的 95PPU 宽度(大的 R -factor)获取较多的观测数据(大的 P -factor)。而理想状态是较多的观测数据同时只有很小的 95PPU 宽度。SUFI-2 算法正是通过保持较多观测数据的同时减小 95PPU 宽度。

参考文献:

- [1] 王明芹,王海英,张巧芝. 舂水河流域生态清洁小流域建设跃上大流域治理新台阶 [J]. 北京水务, 2010,(5):52-54. (WANG Mingqin, WANG Haiying, ZHANG Qiaozhi. Construction of small eco-clean watershed leapt to a new level of large watershed in Guishui River watershed [J]. Beijing Water, 2010,(5):52-54. (in Chinese))
- [2] 要晋峰,胡加林. 延庆县水务建设管理问题与对策 [J]. 北京水务, 2009,(6):50-52. (YAO Jinfeng, HU Jialin. Water construction man-

- agement problems and solutions in Yanqing county [J]. Beijing Water, 2009,(6):50–52. (in Chinese))
- [3] 王经臣. 分布式水文模型在水资源管理中的作用[J]. 中国城市经济, 2011,(12):283. (WANG Jingchen. The use of distributed hydrological model in water resources management [J]. China Urban Economy, 2011,(12):283. (in Chinese))
- [4] Neitsch S L, Arnold J G, Kiniry J R, et al. Soil and water assessment tool theoretical documentation [A]. Texas Water Resources Institute Technical Report 406 [C]. Texas A&M University, 2011.
- [5] Moriasi D N, Arnold J G, Van Liew M W, et al. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations [J]. Transactions of the ASABE, 2007,50(3):885–900.
- [6] Abbaspour K C. SWAT-CUP4: SWAT Calibration and Uncertainty Programs—A User Manual. Eawag Swiss Federal Institute of Aquatic Science and Technology [Z]. 2011.
- [7] 李慧,靳晟,雷晓云,等. SWAT 模型参数敏感性分析与自动率定的重要性研究——以玛纳斯河径流模拟为例 [J]. 水资源与水工程学报, 2010,21(1):79–82. (LI Hui, JIN Sheng, LEI Xiaoyun, et al. Study on the importance of sensitivity analysis and auto-calibration of SWAT model: taking the case study of Manasi watershed [J]. Journal of Water Resources and Water Engineering, 2010,21(1):79–82. (in Chinese))
- [8] 黄清华,张万昌. SWAT 模型参数敏感性分析及应用[J]. 干旱区地理, 2010,(1):8–15. (HUANG Qinghua, ZHANG Wanchang. Application and parameters sensitivity analysis of SWAT model [J]. Arid Land Geography, 2010,(1):8–15. (in Chinese))
- [9] Nossent J, Bauwens W. Multi-variable sensitivity and identifiability analysis for a complex environmental model in view of integrated water quantity and water quality modeling [J]. Water Science and Technology, 2012,65(3):539–549.
- [10] 徐会军,陈洋波,李昼阳,等. 基于 LH-OAT 分布式水文模型参数敏感性分析 [J]. 人民长江, 2012,(7):19–23. (XU Huijun, CHEN Yangbo, LI Zhouyang, et al. analysis on parameter sensitivity of distributed hydrological model based on LH-OAT method [J]. Yangtze River, 2012,(7):19–23. (in Chinese))
- [11] Lenhart T, Eckhardt K, Fohrer N, et al. Comparison of two different approaches of sensitivity analysis [J]. Physics and Chemistry of the Earth, Parts A/B/C, 2002,27(9): 645–654.
- [12] Krause P, Boyle D P, Base F. Comparison of different efficiency criteria for hydrological model assessment [J]. Advances in Geosciences, 2005,5(5):89–97.
- [13] 左德鹏,徐宗学. 基于 SWAT 模型和 SUFI-2 算法的渭河流域月径流分布式模拟[J]. 北京师范大学学报(自然科学版), 2012,48(5): 490–496. (ZUO Depeng, XU Zongxue. Distributed hydrological simulation using SWAT and SUFI-2 in the Wei river basin [J]. Journal of Beijing Normal University (Natural Science), 2012,48(5): 490–496. (in Chinese))
- [14] Yang J, Reichert P, Abbaspour K C, et al. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe basin in China[J]. Journal of Hydrology, 2008,358(1):1–23.

Hydrological Simulation and Parameter Uncertainty Analysis Using SWAT Model Based on SUFI-2 Algorithm for Guishuihe River Basin

LI Qian^{1,2,3}, ZHANG Jing^{1,2,3}, GONG Huili^{1,2,3}

(1. State Key Laboratory Incubation Base of Urban Environmental Processes and Digital Simulation, Capital Normal University, Beijing 100048, China; 2. Beijing Key Laboratory of Resource Environment and Geographic Information System, Capital Normal University, Beijing 100048, China; 3. Key Laboratory of Dimensional Information Acquisition and Application, Ministry of Education, Capital Normal University, Beijing 100048, China)

Abstract: The application of distributed hydrological model is usually made depending on the sensitivity analysis and the uncertainty analysis. In order to simulate monthly hydrograph for the Guishuihe River Basin, a distributed hydrological model was developed with SWAT (soil and water assessment tool), calibrated and validated with SUFI-2 (sequential uncertainty fitting program) manually and automatically. The sensitivity analysis was processed by LH-OAT using both SWAT and SUFI-2 and the uncertainty analysis was evaluated by SUFI-2. The criteria of the uncertainty results are P -factor and the R -factor. In this study, $R^2=0.59$, $NSE=0.56$ during the calibration period, and $R^2=0.82$, $NSE=0.80$ during the validation period. In the both periods, the P -factor is greater than 0.5 and the R -factor is smaller than 0.5.

Key words: SWAT; SUFI-2; sensitivity analysis; uncertainty analysis