A comparison of single- and multi-gauge based calibrations for hydrological modeling of the Upper Daning River Watershed in China's Three Gorges Reservoir Region

Yongwei Gong, Zhenyao Shen, Ruimin Liu, Qian Hong and Xing Wu

ABSTRACT

Many parameters in hydrological models cannot be measured directly and should be estimated through calibration. In this study, four separate calibration schemes, including a multi-gauge and three single-gauge ones, were conducted to simulate the daily stream flow of the Upper Daning River in China's Three Gorges Reservoir Region using the Soil and Water Assessment Tool (SWAT). A highly efficient calibration technique – the Sequential Uncertainty Fitting version-2 (SUFI-2) procedure – was utilized in the calibration and validation. Results indicated that all four schemes could obtain satisfactory parameter sets, and there were no obvious differences between the simulation results obtained by the multi- and single-gauge calibration schemes in terms of Nash–Sutcliffe efficiency value (E_{NS}). This might be attributed to the uniformly distributed land uses, similar topography variations and the low resolution of soil type map of the study area.

Key words | calibration, parameter extrapolating, Sequential Uncertainty Fitting version-2, Soil and Water Assessment Tool, Upper Daning River Watershed

Yongwei Gong
Zhenyao Shen (corresponding author)
Ruimin Liu
Qian Hong
Xing Wu
State Key Joint Laboratory of Environmental Simulation and Pollution Control,
School of Environment,
Beijing Normal University,
Beijing, 100875,
China
E-mail: *zyshen@bnu.edu.cn*

Yongwei Gong
Key Laboratory of Urban Stormwater System and Water Environment (Beijing University of Civil Engineering and Architecture),
Ministry of Education,
Beijing, 100044,
China

INTRODUCTION

Hydrological models are frequently used for assessing water resource problems across the globe. These models can be classified into lumped, semi-lumped, semi-distributed and distributed ones. One important factor affecting their performances is the input data, including precipitation, spatial information, observed discharges, etc. (Krysanova *et al.* 1999; Andersen *et al.* 2001; Moussa *et al.* 2007; Das *et al.* 2008; Feyen *et al.* 2008).

Semi-distributed and distributed hydrological models usually have a large number of parameters to describe the hydrological process. However, due to temporal and spatial variability, measurement error, and because many parameters have physical meanings but cannot be measured or even have no physical meaning, etc., the values of many parameters cannot be exactly known. Therefore, in most cases a model calibration is necessary (Eckhardt & Arnold 2001).

Calibration can be conducted manually or automatically. To some extent, manual calibration is subjective and time consuming (Kannan *et al.* 2008), so automated calibration techniques, such as the Shuffled Complex Evolution (SCE) algorithm (Duan *et al.* 1992), the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb 2001), the Parameter Estimation (PEST) software (Doherty 2005), the Soil and Water Assessment Tool (SWAT) Calibration and Uncertainty Programs (SWAT-CUP) (Abbaspour 2008), have been developed to overcome these shortcomings. A calibration technique should be selected according to the efficiency of the algorithm and the demands of the study.

Single-gauge calibration has commonly been used for hydrological modeling (Hörmann *et al.* 2009; Li *et al.* 2010b; Sudheer *et al.* 2011). This is inevitable in a region where there is only one gauge or little hydrological data.

doi: 10.2166/nh.2012.021

However, the data of only one gauge may not be capable of characterizing the spatial and temporal variability over a large watershed. Therefore, researchers have attempted to use the data of multi-gauges for calibration (Cao *et al.* 2006; Bekele & Nicklow 2007; Migliaccio & Chaubey 2007; Feyen *et al.* 2008; Zhang *et al.* 2008; Li *et al.* 2010a). Previous comparisons between these two calibration methods conducted by Bekele & Nicklow (2007), Zhang *et al.* (2008) and Li *et al.* (2010a) showed better performances of multi-gauge calibration than single-gauge calibration for hydrological modeling. However, this conclusion is site specific and may not suit the conditions of other areas.

The goal of this study was to investigate whether there were significant differences between multi-gauge and single-gauge calibrations in hydrological modeling for the Upper Daning River Watershed in the Three Gorges Reservoir Region in China. Minor differences would suggest that the calibrated parameters could be extrapolated to nearby watersheds in the Three Gorges Reservoir Region with high confidence, while major differences indicated that parameter extrapolating should be cautious. The results from this study may help estimate the discharges of the tributaries without gauges to the reservoir, and thereby provide a scientific basis for the operation of the Three Gorges Reservoir.

The remainder of this paper is organized as follows. The Materials and Methods section presents the description on the study watershed and data used in the study. A description of the model, calibration technique and procedure is also provided in the same section. The multi-gauge and single-gauge calibration results are presented in the Results section. The reasons for the different results of this study are explained in the Discussion section. The final section provides the Conclusions.

MATERIALS AND METHODS

Watershed and data

The drainage area controlled by the Wuxi hydrological gauge (WX) on the Daning River was selected as the study watershed (the Upper Daning River Watershed). This area is located in Wuxi County in the Three Gorges Reservoir Region of China (Figure 1) and covers an area of 2,010 km^2. It is characterized by the north subtropical monsoon climate and has an annual mean precipitation of 1,182 mm. The WX gauge has an annual discharge of 956 mm. The primary land uses in the watershed include 61.9% forest, 24.9% arable land, and 12.5% pasture. The primary soil types include 30.5% yellow brown soil, 20.3% yellow cinnamon soil and 15.6% purple soil. The geology of the area consists of two major components: a pre-Sinian crystalline basement, and a Sinian–Jurassic sedimentary cover. The former is composed of magmatic and metamorphic rocks, and outcrops only sporadically in the area. The latter is widespread and comprises interbedded carbonate, sandstone and shale formations. The altitude of this region is from 200 to 2,588 m and decreases from northeast to southwest.

Rainfall data were collected at six rainfall gauges in the watershed (Figures 1 and 2) for 2000–2007. The other required climate data were collected at Wuxi County Weather Station. The 1:100,000 land use map for the year 2000 developed by the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, the 1:250,000 digital elevation model (DEM) map developed by National Geomatics Center of China and the 1:1,000,000 soil type map developed by the Institute of Soil Science, Chinese Academy of Sciences were used in this study (Figure 3). The daily stream flow data for 2000–2007 of three hydrological gauges – Ningqiao (NQ), Ningchang (NC), and WX (Figure 1) – were collected from Changjiang Water Resources Commission, China (Table 1). Additionally, the database for planting, harvest, and tillage operations was built based on field investigations.

Watershed model and calibration technique description

The SWAT model (version ArcSWAT 2.1.1beta) was used in this study. SWAT is a process-based distributed-parameter simulation model, operating on a daily step (Arnold *et al.* 1998). It was developed to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying land uses, soil types and management conditions over long periods (Neitsch *et al.* 2005b). It has been widely used for assessing water resource and non-point source pollution

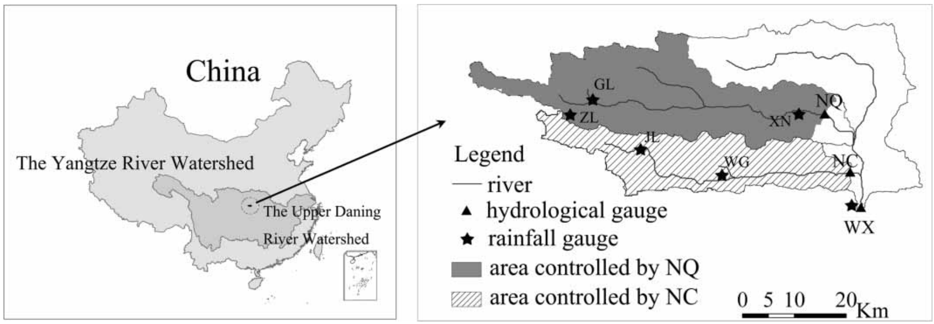

Figure 1 | Location map of the study watershed and the three hydrological gauges.

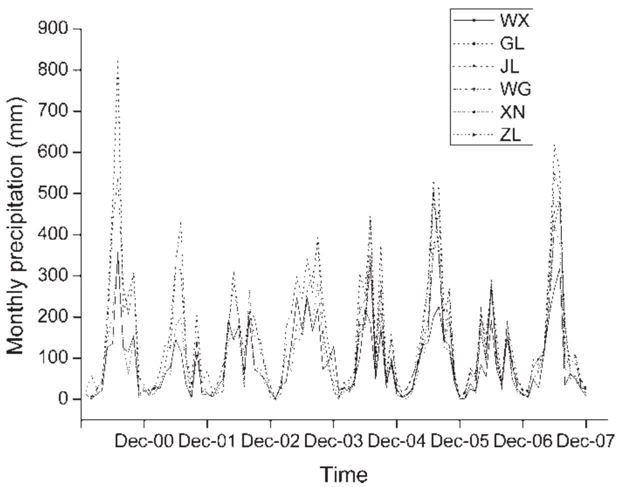

Figure 2 | Monthly precipitation of the rainfall gauges for 2000–2007.

problems across the globe (Gassman *et al.* 2007). SWAT partitions a watershed into a number of sub-watersheds based on a given DEM. Each sub-watershed is divided into hydrologic response units (HRUs) that are particular combinations of land use, soil type and slope range. Runoff volume is calculated using the soil conservation service Curve Number (CN) method or the Green and Ampt method and is calculated continuously as a function of time. Water is routed through the channel network using the variable storage routing method or the Muskingum river routing method. In the present study, the CN method and the variable storage method were employed.

Parameters for SWAT are defined at three different levels: watershed, sub-watershed and HRU. Watershed level parameters are used to model processes throughout the watershed. Sub-watershed level parameters are set at the same values for all HRUs in the sub-watershed. HRU level parameters can be set to unique values for each HRU in the watershed.

The Sequential Uncertainty Fitting version-2 (SUFI-2) procedure was used to calibrate and validate the SWAT model in this study. It is an inverse optimization approach that uses the Latin hypercube sampling (LHS) procedure along with a global search algorithm to examine the behavior of objective functions. The LHS method is a stratified sampling technique where the random variable distributions are divided into equal probability intervals. Firstly, parameters are divided into the indicated number of simulations. Secondly, parameter segments are randomized. Finally, a random sample is taken in every segment and the combination forms a parameter set. The initial parameter ranges can be updated for every iteration, and the recommended new parameter ranges are always centered on the current best estimate (Abbaspour *et al.* 2004). The procedure has been incorporated into the SWAT-CUP software (Abbaspour 2008), which can be downloaded for free from the Eawag website (Abbaspour 2009).

Evaluation criteria

In this study, the Nash–Sutcliffe efficiency values (E_{NS}) was chosen to assess the SWAT performances. E_{NS} is expressed as (Nash & Sutcliffe 1970):

$$E_{NS} = 1 - \sum_{i=1}^{n} (O_i - P_i)^2 / \sum_{i=1}^{n} (O_i - \bar{O})^2 \tag{1}$$

where O_i and P_i are the observed and simulated values for the ith pair, $\bar{O}$ is the mean of the observed values, and n is the total number of paired values.

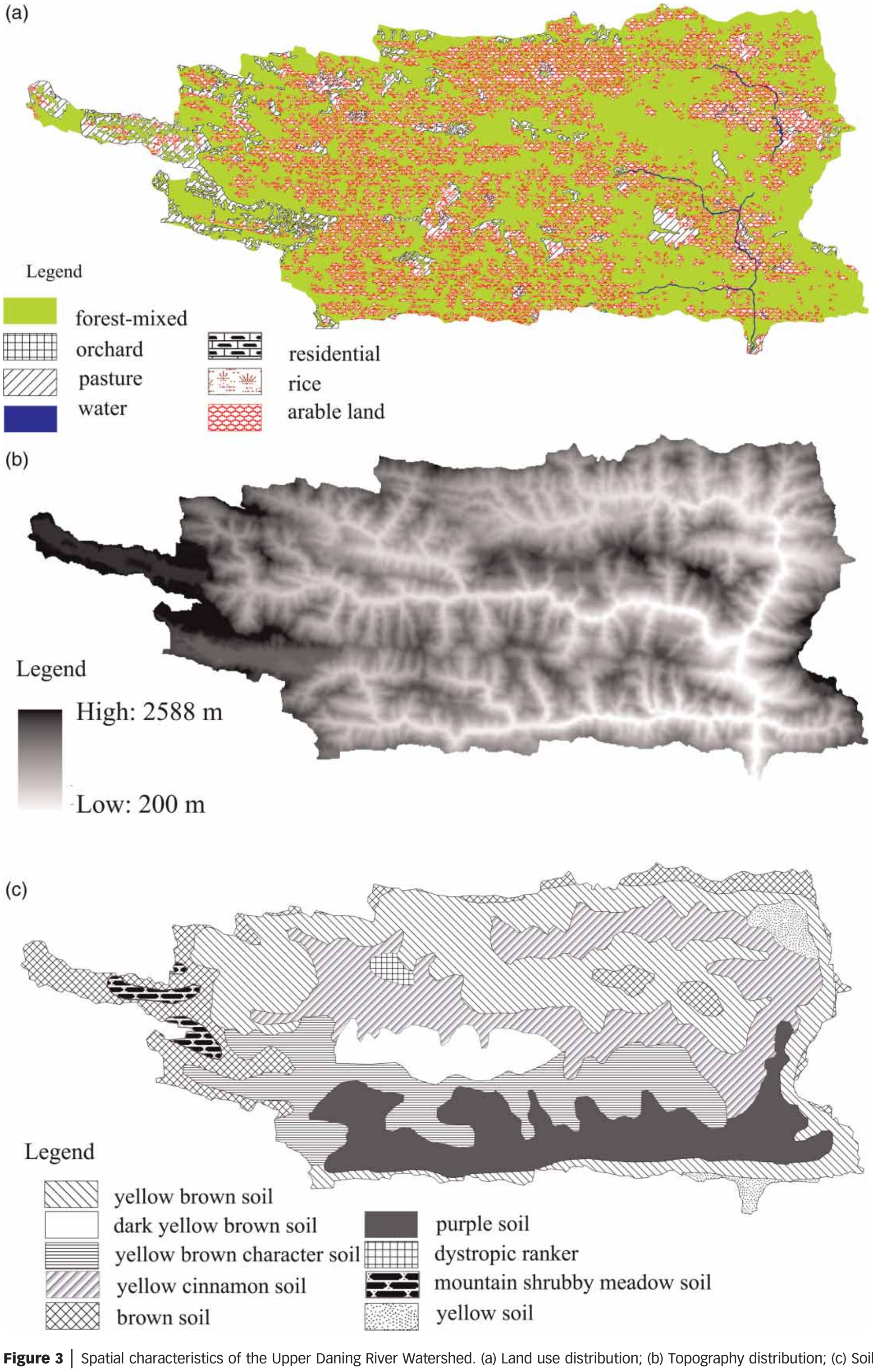

Figure 3 | Spatial characteristics of the Upper Daning River Watershed. (a) Land use distribution; (b) Topography distribution; (c) Soil type distribution.

Table 1 | Daily flow statistics of the NQ, NC, and WX gauges (m^3/s)

Period	Gauge	Mean	Minimum	Maximum	Standard deviations
2000–2003	NQ	26.8	3.0	858.0	50.5
	NC	15.1	0.5	623.0	37.8
	WX	68.4	10.6	2,230.0	125.1
2004–2007	NQ	25.8	11.7	1,050.0	92.2
	NC	13.8	1.0	576.0	37.8
	WX	60.1	10.1	1,730.0	118.1

The range of the E_{NS} value is from $-\infty$ to 1, with 1 indicating a perfect fit. According to Moriasi *et al.* (2007), satisfactory to very good of E_{NS} fall in the range of 0.5 to 1.

Model calibration and validation

Sub-watershed boundaries and HRUs were delineated using the ArcSWAT interface. In SWAT application, the threshold configuration for sub-watersheds and HRUs delineation has a great impact on simulated results. In a number of initial tests, different threshold drainage areas were specified for the Upper Daning River Watershed to obtain various watershed configurations. The percentage threshold levels for land use, soil type and slope range were all set to 0, which allowed all land uses, soil types and slope ranges to be included in the simulations. The configuration with the 50 km^2 threshold drainage area had the highest E_{NS} (Gong *et al.* 2010).

HRU delineation is subdividing each sub-watershed by determining the thresholds for land use, soil and slope. The user may specify sensitivities for the land use, soil, and slope data that will be used to determine the number and type of HRUs in each sub-watershed. The land uses that cover a percentage of the sub-watershed area less than the threshold level are eliminated. After the elimination process, the area of the remaining land uses is re-apportioned so that 100% of the land area in the sub-watershed is modeled. The thresholds for soil and slope have similar meanings. Previously we tested the simulation performances using the 0–0–0, 5–5–5 and 10–10–10% (land use–soil–slope) area thresholds for HRU delineation. Results indicated there were no obvious differences in the modeling results. To enhance the modeling efficiency, the 10–10–10% area thresholds for HRU delineation was used in this study.

Four parameterization schemes were conducted for daily stream flow simulations. One was the multi-gauge calibration, and the chosen parameters were calibrated simultaneously using all the hydrological data of the NQ, NC and WX gauges. The other three were single-gauge calibrations, and each one used the data of one gauge. Hereafter, these four schemes are referred to as Multi-cal, NQ-cal, NC-cal and WX-cal, and the stream flow simulations of the three gauges were referred to as Q_{NQ}, Q_{NC} and Q_{WX}, respectively. For each scheme, the daily stream flow data of 2000–2003 were used for calibration, and the data of 2004–2007 were used for validation, with the exception of the years 2005–2007 when no data were available for the NQ gauge.

Identifying the most sensitive input parameters was not an essential component of the calibration procedure in this study because parameterization was applied to the parameter set rather than to individual parameters and all parameters were calibrated simultaneously in the SUFI-2 procedure. Seventeen parameters that govern hydrological process were calibrated (Table 2). The initial range of each parameter was determined according to the recommendations from Neitsch *et al.* (2005a) and Abbaspour (2008). During the calibrating process, SWAT-CUP changed the parameters based on their existing values. In Table 2, v__, a__ and r__ are the codes that indicate the types of change to be applied to the existing parameters. v__ means the existing parameter value is to be replaced by the given value, a__ means the given value is added to the existing parameter value, and r__ means the existing parameter value is multiplied by (1 + a given value).

After a number of tests, it was found that four iterations with 1,000 model runs in each iteration were sufficient to provide a relatively small and stable range for the parameter set. The E_{NS} values of the observed and simulated data were calculated for all iterations.

For the multi-gauge calibration scheme, the weighted sum of E_{NS} was used to determine the best set of parameter values in SUFI-2 (Abbaspour 2008). It was calculated by the following formula:

$$f = w_1 f_{NQ} + w_2 f_{NC} + w_3 f_{WX} \quad (2)$$

Table 2 | Selected parameters of the SWAT model for calibration

No.	Parameter	Description	Lower bound	Upper bound
1	r__CN2.mgt	Curve number II	–0.25	0.15
2	v__SFTMP.bsn	Snow melt base temperature (°C)	–5	5
3	v__CH_N2.rte	Manning's 'n' value for the main channel	0	0.5
4	v__CH_K2.rte	Effective hydraulic conductivity in main channel alluvium (mm h^{-1})	0	150
5	v__ALPHA_BNK.rte	Base flow alpha factor for bank storage (days)	0	1
6	v__SOL_AWC(1-2).sol	Available soil water capacity	0	1
7	r__SOL_K(1-2).sol	Soil hydraulic conductivity (mm/h)	–0.2	300
8	a__SOL_BD(1-2).sol	Moist bulk density (mg/m^3)	0.1	0.6
9	r__OV_N.hru	Overland manning roughness	–0.1	0.1
10	v__CANMX.hru	Maximum canopy storage (mm)	0	100
11	v__ESCO.hru	Soil evaporation compensation factor	0.01	1
12	r__SLSUBBSN.hru	Average slope length (m)	–0.1	0.1
13	v__Rchrg_Dp.gw	Deep aquifer percolation fraction	0	1
14	v__GWQMN.gw	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	0	5,000
15	v__REVAPMN.gw	Threshold depth of water in the shallow aquifer for re-evaporation to occur (mm)	0	500
16	v__ALPHA_BF.gw	Base flow recession constant	0	1
17	v__GW_DELAY.gw	Groundwater delay (days)	1	45

where f was the weighted sum of E_{NS}, f_{NQ}, f_{NC} and f_{WX} were the E_{NS} values for Q_{NQ}, Q_{NC} and Q_{WX}, respectively, and w_1, w_2 and w_3 were the weights of f_{NQ}, f_{NC} and f_{WX}, respectively.

The values of w_1, w_2 and w_3 were determined according to the drainage areas controlled by each gauge. Each of the NQ and NC gauges controls one tributary, while the WX gauge is on the main reach and controls all tributaries, and the corresponding control areas are 777, 504 and 2,010 km^2, respectively (Figure 1). Therefore the values of w_1, w_2 and w_3 were set to be 0.24, 0.15 and 0.61 (normalized), respectively (Gong *et al.* 2010).

By using the SUFI-2 algorithm, the interactions between the multiple gauges were not considered during the calibration processes in this study. The differences between the multi- and single-gauge calibrations lie in that the parameters calibrated by the former scheme could reflect the characteristics of the whole watershed more averagely. When using the watershed outlet as the only gauge in single-gauge schemes, the computational time for SWAT-CUP running is the same as multi-gauge ones. However, when the gauge is in the upper reach of the watershed, the computational time of single-gauge schemes is often much less than that of multi-gauge ones. Therefore, the location of the single gauge determines the running time, and thus impacts the choice of single- or multi-gauge schemes.

RESULTS

The objective function values of the four parameterization schemes in the calibration and validation periods are listed in Table 3. For each scheme, the weighted sum of E_{NS} was calculated using Equation (2).

Single-gauge calibration

In the calibration period, the E_{NS} values for all gauges were high for the three single-gauge schemes. For NQ-cal and NC-cal, there were highest E_{NS} values for Q_{NQ} and Q_{NC}, respectively. It was not the same for WX-cal, but the differences of

Table 3 | The E_{NS} values of different schemes in the calibration and validation periods

Scheme	Variable	E_{NS} (calibration)	E_{NS} (validation)
NQ-cal	Q_{NQ}	0.77	0.80
	Q_{NC}	0.68	0.44
	Q_{WX}	0.73	0.52
	Weighted sum	0.73	0.58
NC-cal	Q_{NQ}	0.74	0.73
	Q_{NC}	0.85	0.53
	Q_{WX}	0.86	0.64
	Weighted sum	0.83	0.65
WX-cal	Q_{NQ}	0.73	0.69
	Q_{NC}	0.84	0.53
	Q_{WX}	0.85	0.65
	Weighted sum	0.82	0.64
Multi-cal	Q_{NQ}	0.77	0.76
	Q_{NC}	0.83	0.54
	Q_{WX}	0.85	0.66
	Weighted sum	0.83	0.67

E_{NS} between the best simulation (Q_{NC}) and Q_{WX} were minor under this scheme.

In the validation period, the E_{NS} values for all gauges were satisfactory for the three single-gauge schemes. The E_{NS} values of Q_{NC} were lower than those of Q_{NQ} and Q_{WX}. However, all the E_{NS} values were >0.5 except for the E_{NS} value of Q_{NC} under the NQ-cal scheme.

Multi- versus single-gauge calibration

When using the multi-gauge scheme for calibration, the E_{NS} values were high in the calibration period and satisfactory in the validation period. There were no apparent differences for E_{NS} between it and the other three single-gauge schemes. Compared to NQ-cal, NC-cal and WX-cal, the changes in E_{NS} values were 13.1–22.1, –2.4–4.1 and –1.2–5.5% for calibration and –5.0–26.9, 1.9–4.1 and 1.5–3.8% for validation, respectively. The results indicated that the multi-gauge calibration scheme could not always obtain better model performances than single-gauge schemes. However, some single-gauge schemes, such as NC-cal, and WX-cal, provided similar simulation performances with the multi-gauge scheme. Moreover, the calibration work was conducted on an area of 15% of the total watershed for NC-cal, which needed relatively less computational time than the multi-gauge scheme. While considering the modeling performances and computational time together, the NC-cal had the absolute advantage of the Multi-cal.

The overall results indicated that all schemes yielded acceptable simulation results, both for the calibration and validation periods. The calibrated parameter values were normalized using the initial range of each parameter and are shown in Figure 4. Graphical analysis shows that most parameters had significantly varied values and this revealed the phenomenon of equifinality (different sets of parameter values are equally likely as simulators of a watershed) (Beven 1993).

DISCUSSION

For all three single-gauge calibrations, the E_{NS} values were high when the calibrated parameter values of one gauge were used to simulate the stream flow of the other two gauges. The finding differs from Zhang *et al.* (2008) where the E_{NS} values were much lower when the calibrated parameter values of one gauge were applied to the other region of the study watershed. The better performances of parameter extrapolating in this study may be attributable to the similar rainfall trends within the watershed (Figure 2).

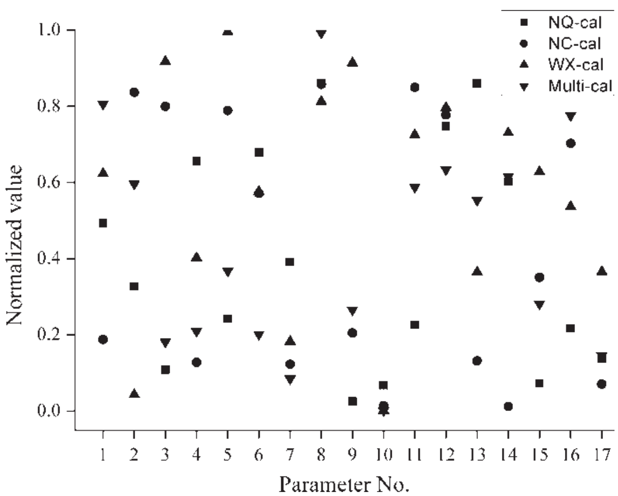

Figure 4 | Normalized parameters obtained by four schemes (the parameter no. is listed in Table 2).

Another reason may be the high stream flow of the Daning River (Table 1). The coefficients of variation (CV) for the three data sets were large. However, a shortcoming of the Nash–Sutcliffe statistic is that it does not perform well in periods of low flow (Schaefli & Gupta 2007; Pandey et al. 2008). If the daily observed flow approaches the average value, the denominator of Equation (1) tends to be zero and E_{NS} approaches negative infinity. The Nash–Sutcliffe statistic works well when the CV value for the data set is large, and this is in accordance with the Daning River study.

For both the calibration and validation periods, there were no apparent differences between the simulated results provided by multi-gauge and the other three single-gauge calibration schemes. This could be explained in terms of the SWAT modeling mechanism. In SWAT, the hydrological simulation is conducted on the HRU basis; therefore many important parameters, such as the ones in the HRU general input files, the HRU management files and the groundwater input files, are determined on the basis of combinations of land use, topography and soil. Land uses can impact on several components of the water balance: plant physiology regulates transpiration; canopy structure determines interception storage and throughfall; and rooting depth, density and structure affect plant water uptake and infiltration capacity (Breuer et al. 2009). The topography variations and the physical properties of the soils are also important and have strong influences on the hydrological process (Dunn & Mackay 1995; Shivakoti et al. 2008). In the Upper Daning River Watershed, there were relatively uniformly distributed land uses across the watershed and similar slope variations between the areas controlled by NQ and NC gauges (Figure 3). Therefore, land uses and slope variations had little impact in this study. The resolution of soil type information map used in this study was rather coarse (1:1,000,000). As a result, even though the soil type distribution was different between the areas controlled by NQ and NC gauges, this variation had little influence on parameter extrapolating in this study. However, possible errors and uncertainties in the input and output data and in the model structure may be masked or compensated by adjusting the parameters. This might lead to a good model performance, but can also result in unrealistic parameter values. The parameter uncertainty was also analyzed in a previous study (Gong et al. 2011).

For the same reason described above, the weight determination for the three gauges in the Multi-cal was not very important in this study. Although multi-objective function optimization algorithms are recommended in multi-gauge calibration (Migliaccio & Chaubey 2007; Zhang et al. 2008), it is not essential to a good calibration of the physically based distributed model (Beven 2006; Cao et al. 2007). Clearly there must be some level of subjectivity in using the selected weights. However, these weights were reasonable since the relative importance of the WX gauge in the watershed should be greater than that of the NQ and NC gauges, and this could be well reflected by using the present values. It is possible to achieve good modeling performances for single-gauge calibration if the land uses and soil types are uniformly distributed across the watershed and the topography variations are similar for different sub-watersheds of the study watershed.

During the calibration in this study, one disadvantage concerning the SWAT model was found: some parameters can only have single values across the whole watershed, such as the parameters in the basin input file (the basin input file contains inputs for physical processes modeled or defined at the watershed level). However, in a large watershed these parameters may vary considerably, and so this restriction would affect the modeling performance. If these parameters could be assigned by different values for different sub-watersheds, there might be better simulation results using the SUFI-2 procedure for the multi-gauge calibration.

There are a limited number of hydrological gauges within the Three Gorges Reservoir Region, and so the discharges from some tributaries of the Yangtze River to the reservoir are not exactly known. Usually, hydrology prediction for ungauged areas is necessary (Sivapalan et al. 2003; Bocchiola et al. 2010; Castiglioni et al. 2010; Makungo et al. 2010). There are many methods to regionalize hydrological parameters (Dawson et al. 2006; Widén-Nilsson et al. 2007; Bastola et al. 2008; Dornes et al. 2008; Castiglioni et al. 2010). The current set-up of the study is one relatively simple way to test whether parameters can be extrapolated to other watersheds. It can be inferred from the results of this study that the model parameterization results calibrated by one gauge could be used on nearby watersheds with similar land use distributions, soil types and geology. However, because the observed data were available for only three gauges, the conclusion of this study is site specific and may be different if the watershed conditions are significantly different from those discussed here. Therefore, to provide scientific basis for the operation of the Three Gorges Reservoir, more tests may be needed to determine whether the discharges of the ungauged watersheds in the Three Gorges Reservoir Region could be estimated in this way.

CONCLUSIONS

One multi-gauge and three single-gauge calibration schemes were adopted for parameterization of the SWAT model in the Upper Daning River Watershed using the SUFI-2 procedure. The results indicated that all schemes yielded acceptable simulation results, both for the calibration and validation periods. The parameters calibrated for any single-gauge calibration scheme could be applied to the other gauges with relatively high E_{NS} values. There were no apparently different performances between the multi-gauge and the other three single-gauge schemes, and this might be attributed to the uniformly distributed land uses, similar slopes variations and the low resolution of the soil type map. For the same reason, the weight determination for the three gauges in the Multi-cal was not very important in this study. However, this conclusion is site specific and may be different if the watershed conditions are significantly different from those discussed in this study. Additionally, it is not appropriate that some parameters have only single values across the watershed. Therefore, the rule of SWAT parameter setting deserves further improvement in future studies.

ACKNOWLEDGEMENTS

The study was supported by the Nonprofit Environment Protection Specific Project of China (No. 200709024), the National Science Foundation for Distinguished Young Scholars (No. 51025933) and the National Natural Science Foundation of China (No. 40771193). The authors are grateful to Dr Karim C. Abbaspour at Eawag, Swiss Federal Institute of Aquatic Science and Technology for help with the SWAT-CUP application.

REFERENCES

Abbaspour, K. C. 2008 *SWAT-CUP2: SWAT Calibration and Uncertainty Programs – A User Manual*. Department of Systems Analysis, Integrated Assessment and Modelling (SIAM), Eawag, Swiss Federal Institute of Aquatic Science and Technology, Switzerland.

Abbaspour, K. C. 2009 SWAT-CUP Programme Version 2.1.5.

Abbaspour, K. C., Johnson, C. A. & van Genuchten, M. T. 2004 Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J.* **3** (4), 1340–1352.

Andersen, J., Refsgaard, J. C. & Jensen, K. H. 2001 Distributed hydrological modelling of the Senegal River Basin – model construction and validation. *J. Hydrol.* **247** (3–4), 200–214.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modeling and assessment: part I. Model development. *J. Am. Water Resour. As.* **34** (1), 73–89.

Bastola, S., Ishidaira, H. & Takeuchi, K. 2008 Regionalisation of hydrological model parameters under parameter uncertainty: a case study involving TOPMODEL and basins across the globe. *J. Hydrol.* **357** (3–4), 188–206.

Bekele, E. G. & Nicklow, J. W. 2007 Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* **341** (3–4), 165–176.

Beven, K. 1993 Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.* **16**, 41–51.

Beven, K. 2006 A manifesto for the equifinality thesis. *J. Hydrol.* **320** (1–2), 18–36.

Bocchiola, D., Mihalcea, C., Diolaiuti, G., Mosconi, B., Smiraglia, C. & Rosso, R. 2010 Flow prediction in high altitude ungauged catchments: a case study in the Italian Alps (Pantano Basin, Adamello Group). *Adv. Water Resour.* **33** (10), 1224–1234.

Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H. G., Gräffe, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindströmk, G., Seibert, J., Sivapalan, M. & Viney, N. R. 2009 Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: model intercomparison with current land use. *Adv. Water Resour.* **32** (2), 129–146.

Cao, W., Bowden, W., Davie, T. & Fenemor, A. 2006 Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Process.* **20** (5), 1057–1073.

Cao, W., Davie, T., Fenemor, A. & Bowden, W. 2007 Reply to comment on Cao W, Bowden BW, Davie T, Fenemor A. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Process.* **20** (5), 1057–1073. *Hydrol. Process*. **21** (23), 3229–3230.

Castiglioni, S., Lombardi, L., Toth, E., Castellarin, A. & Montanari, A. 2010 Calibration of rainfall-runoff models in ungauged basins: a regional maximum likelihood approach. *Adv. Water Resour.* **33** (10), 1235–1242.

Das, T., Bardossy, A., Zehe, E. & He, Y. 2008 Comparison of conceptual model performance using different representations of spatial variability. *J. Hydrol.* **356** (1–2), 106–118.

Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y. & Wilby, R. L. 2006 Flood estimation at ungauged sites using artificial neural networks. *J. Hydrol.* **319** (1–4), 391–409.

Deb, K. 2001 *Multi-objective Optimization Using Evolutionary Algorithms*. Wiley, Chichester, UK.

Doherty, J. 2005 *PEST: Model Independent Parameter Estimation. Fifth Edition of User Manual*. Watermark Numerical Computing, Brisbane, Australia.

Dornes, P. F., Tolson, B. A., Davison, B., Pietroniro, A., Pomeroy, J. W. & Marsh, P. 2008 Regionalisation of land surface hydrological model parameters in subarctic and arctic environments. *Phys. Chem. Earth, Parts A/B/C* **33** (17–18), 1081–1089.

Duan, Q., Sorooshian, S. & Gupta, V. K. 1992 Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **28** (4), 1015–1131.

Dunn, S. M. & Mackay, R. 1995 Spatial variation in evapotranspiration and the influence of land use on catchment hydrology. *J. Hydrol.* **171** (1–2), 49–73.

Eckhardt, K. & Arnold, J. G. 2001 Automatic calibration of a distributed catchment model. *J. Hydrol.* **251** (1–2), 103–109.

Feyen, L., Kalas, M. & Vrugt, J. 2008 Semi-distributed parameter optimization and uncertainty assessment for large-scale streamflow simulation using global optimization. *Hydrol. Sci. J.* **53** (2), 293–308.

Gassman, P., Reyes, M., Green, C. & Arnold, J. 2007 The soil and water assessment tool: historical development, applications, and future research directions. *T. ASAE* **50** (4), 1211–1250.

Gong, Y., Shen, Z., Hong, Q., Liu, R. & Liao, Q. 2011 Parameter uncertainty analysis in watershed total phosphorus modeling using the GLUE methodology. *Agric. Ecosyst. Environ.* **142** (3–4), 246–255.

Gong, Y., Shen, Z., Liu, R., Wang, X. & Chen, T. 2010 Effect of watershed subdivision on SWAT modeling with consideration of parameter uncertainty. *J. Hydrol. Eng.* **15** (12), 1070–1074.

Hörmann, G., Köplin, N., Cai, Q. & Fohrer, N. 2009 Using a simple model as a tool to parameterise the SWAT model of the Xiangxi river in China. *Quatern. Int.* **208** (1–2), 116–120.

Kannan, N., Santhi, C. & Arnold, J. G. 2008 Development of an automated procedure for estimation of the spatial variation of runoff in large river basins. *J. Hydrol.* **359** (1–2), 1–15.

Krysanova, V., Bronstert, A. & Muller-Wohlfeil, D. I. 1999 Modelling river discharge for large drainage basins: from lumped to distributed approach. *Hydrol. Sci. J.* **44** (2), 313–331.

Li, X., Weller, D. E. & Jordan, T. E. 2010a Watershed model calibration using multi-objective optimization and multi-site averaging. *J. Hydrol.* **380** (3–4), 277–288.

Li, Z., Shao, Q., Xu, Z. & Cai, X. 2010b Analysis of parameter uncertainty in semi-distributed hydrological models using bootstrap method: a case study of SWAT model applied to Yingluoxia watershed in northwest China. *J. Hydrol.* **385** (1–4), 76–83.

Makungo, R., Odiyo, J. O., Ndiritu, J. G. & Mwaka, B. 2010 Rainfall-runoff modelling approach for ungauged catchments: a case study of Nzhelele River sub-quaternary catchment. *Phys. Chem. Earth, Parts A/B/C* **35** (13–14), 596–607.

Migliaccio, K. & Chaubey, I. 2007 Comment on Cao W, Bowden BW, Davie T, Fenemor A. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Process.* **20** (5), 1057–1073. *Hydrol. Process*. **21** (23), 3226–3228.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., Veith, T. & USDA, A. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *T. ASAE* **50** (3), 885–900.

Moussa, R., Chahinian, N. & Bocquillon, C. 2007 Distributed hydrological modelling of a Mediterranean mountainous catchment – Model construction and multi-site validation. *J. Hydrol.* **337** (1–2), 35–51.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *J. Hydrol.* **10** (3), 282–290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Srinivasan, R. & Williams, J. R. 2005a *Soil and Water Assessment Tool Input/Output File Documentation*, version 2005. Texas Water Resources Institute, Temple, Texas, USA.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Srinivasan, R. & Williams, J. R. 2005b *Soil and Water Assessment Tool Theoretical Documentation*, version 2005. Texas Water Resources Institute, Temple, Texas, USA.

Pandey, A., Chowdary, V. M., Mal, B. C. & Billib, M. 2008 Runoff and sediment yield modeling from a small agricultural watershed in India using the WEPP model. *J. Hydrol.* **348** (3–4), 305–319.

Schaefli, B. & Gupta, H. V. 2007 Do Nash values have value? *Hydrol. Process.* **21** (15), 2075–2080.

Shivakoti, B. R., Fujii, S., Boontanon, S. K., Ihara, H., Moriya, M. & Tanaka, S. 2008 Grid size effects on a distributed water quantity-quality model in a hilly watershed. *Water Sci. Technol.* **58** (9), 1829–1836.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. & Zehe, E. 2003 IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012:

shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **48** (6), 857–880.

Sudheer, K. P., Lakshmi, G. & Chaubey, I. 2011 Application of a pseudo simulator to evaluate the sensitivity of parameters in complex watershed models. *Environ. Modell. Softw.* **26** (2), 135–143.

Widén-Nilsson, E., Halldin, S. & Xu, C.-Y. 2007 Global water-balance modelling with WASMOD-M: parameter estimation and regionalisation. *J. Hydrol.* **340** (1–2), 105–118.

Zhang, X., Srinivasan, R. & Van Liew, M. 2008 Multi-site calibration of the SWAT model for hydrologic modeling. *T. ASAE* **51** (6), 2039–2049.

First received 5 March 2010; accepted in revised form 18 August 2011. Available online 16 February 2012