

## Effect of length of the observed dataset on the calibration of a distributed hydrological model

XINGQI CUI<sup>1</sup>, WENCHAO SUN<sup>1</sup>, JIANBIAO TENG<sup>2</sup>, HAO SONG<sup>1</sup> & XIAOLEI YAO<sup>1</sup>

<sup>1</sup> College of Water Sciences, Beijing Normal University, Beijing, 100875, China

<sup>2</sup> South China Institute of Environmental Sciences, Guangzhou, 510655, China

[tengjianbiao@scies.org](mailto:tengjianbiao@scies.org)

**Abstract** Calibration of hydrological models in ungauged basins is now a hot research topic in the field of hydrology. In addition to the traditional method of parameter regionalization, using discontinuous flow observations to calibrate hydrological models has gradually become popular in recent years. In this study, the possibility of using a limited number of river discharge data to calibrate a distributed hydrological model, the Soil and Water Assessment Tool (SWAT), was explored. The influence of the quantity of discharge measurements on model calibration in the upper Heihe Basin was analysed. Calibration using only one year of daily discharge measurements was compared with calibration using three years of discharge data. The results showed that the parameter values derived from calibration using one year's data could achieve similar model performance with calibration using three years' data, indicating that there is a possibility of using limited numbers of discharge data to calibrate the SWAT model effectively in poorly gauged basins.

**Key words** Heihe basin; length of river discharge data; SWAT; model calibration

### 1 INTRODUCTION

There is a consensus that hydrological modelling is essential for improve the understanding of the hydrological cycle at basin scale (Li *et al.* 1992). Model calibration is a key process for deriving reasonable model parameter values that reflect the characteristics of the water cycle. Only after parameter calibration, from which the best parameter values are determined, is the model regarded as a reliable tool to make predictions. Usually continuous discharge measurements of several years are used for model calibration. However, there are many basins lacking continuous observations for the calibration of hydrological models. How to calibrate models effectively in such poorly gauged or ungauged basins becomes a challenge.

At present, the methods of estimating hydrological model parameters in ungauged basins mainly include regionalization, calibration using remote sensing information and calibration using limited numbers of river discharge data. The regionalization method is commonly used (Blöschl & Sivapalan, 1995; Chai *et al.*, 2005; Young, 2006; Li *et al.*, 2011), which infers parameter values in ungauged basins from gauged ones. Generally, the uncertainty of regionalization methods is considerable. Calibration using satellite observations of river hydraulic information (Sun *et al.*, 2010, 2012), which works as a surrogate of river discharge data in the model calibration, is only applicable to middle to large basins, due to the limitation in the resolution of the remote sensing data. Considering the fact that in the real world, limited numbers of river discharge data may be available in many basins, some researchers have tried to use such low numbers of data for model calibration: Perrin *et al.* (2007) calibrated two rainfall-runoff models using different numbers of observations in 12 American basins. The results demonstrate that in some cases, the model could be calibrated effectively with ten observations. Seibert and Beven (2009) also used a limited number of river discharge data during one year in a Swedish basin to calibrate the HBV model. The research indicated that a few river discharge data could contain the same amount of information for hydrological model parameter identification as long data records.

Distributed hydrologic models based on physical mechanisms can describe the heterogeneity in the climate conditions, land cover and soil type within a basin, which make them attractive for predicting the influence of climate change on the water cycle and analysing the impact of land cover change (Wang *et al.*, 2004; Xu and Cheng, 2010). In previous studies, evaluations of the influence of the amount of discharge measurements used on model calibration were largely focused on conceptual hydrologic models. However, due to the high time demand in the calibration, such evaluation is hard to be carried out for distributed hydrological model.

In this study, the influence of the amount discharge measurements used on the calibration of a distributed hydrological model, the Soil and Water Assessment Tool (SWAT) is evaluated, for the purpose of exploring the minimum number of river discharge observations that can derive reasonable model parameter values. The upper Heihe Basin in China is selected as the case study. For SWAT, usually daily discharge observations over several years are used for model calibration. In this study, whether using only one year data can calibrate the SWAT model effectively is examined.

## 2 STUDY AREA

Heihe basin is located in the northwest of Gansu province, the river originates from the Qilian Mountains and flows successively through Qinghai, Gansu and Inner Mongolia. The elevation of the basin decreases from the southern high mountains area to the north high-plain area. Based on differences in geomorphology, the basin can be divided into three regions: the upper-reach region belonging to the Qilian Mountain area, the middle-reach region belonging to the Hexi Corridor Plain area and the lower-reach region belonging to the Alxa Plateau area. The upper-reach region is the upstream area of the Yingluo Gorge located on the main stream, covering an area of about 10 000 km<sup>2</sup>. It is in the Qinghai-Tibet Plateau climatic region. The annual average temperature of the whole watershed is lower than 2°C and annual precipitation ranges from 300 to 700 mm.

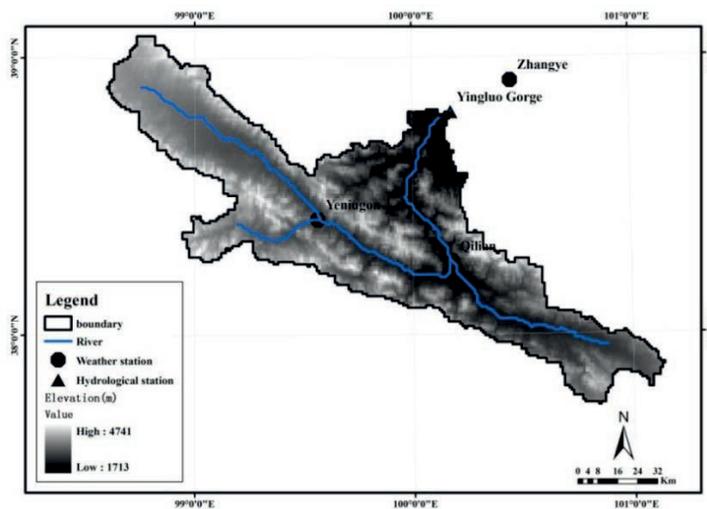


Fig. 1 Topography, weather stations and hydrological station of the upper Heihe basin.

## 3 HYDROLOGICAL MODEL AND CALIBRATION METHOD

### 3.1 SWAT model and sensitivity analysis

The Soil and Water Assessment Tool (SWAT) is a physically-based distributed hydrological model which is suitable for hydrological simulations of middle to large size basins with different soil types, land use and management practices.

In this study, model parameter sensitivity is evaluated by a multiple regression system calculation. The relation between parameter values generated by Latin hypercube sampling and the corresponding value of the objective function were analysed by the following equation:

$$g = \alpha + \sum_{i=1}^m \beta_i b_i \quad (1)$$

where  $g$  is the objective function value,  $\alpha$  and  $\beta_i$  are the regression equation coefficients,  $b_i$  is the parameter value, and  $m$  is the number of parameters. The sensitivity of  $b_i$  was determined using the  $t$ -test, and the significance of parameter sensitivity was determined by the value of  $p$ . The closer to 0 the value of  $p$  was, the more significant the sensitivity was. A model parameter is identified as a sensitive one when the value of  $p$  was less than or equal to 0.05.

### 3.2 Automatic calibration method

The Sequential Uncertainty Fitting, ver.2 (SUFI-2) is one of the automatic calibration programs in the model calibration tool SWAT Calibration and Uncertainty Programs (SWAT-CUP). SUFI-2 combines calibration and uncertainty analysis that accounts for all uncertainties. In SUFI-2, model output uncertainty is quantified by the 95% prediction uncertainty (95PPU) calculated at the 2.5% and 97.5% levels of the cumulative distribution of output variables obtained through Latin hypercube sampling (Yang *et al.*, 2008).

Here, we used the Nash-Sutcliffe efficiency (*NSE*) to evaluate the performance of SWAT model calibration. The value of *NSE* ranges from 0 to 1 and a high value indicates a high degree of fit between the observed and simulated data (Nash & Sutcliffe, 1970). The goodness of fit and the degree to which the calibrated model accounts for the uncertainties are assessed by the *P-factor* and *R-factor*. The *P-factor* is the percentage of measured data bracketed by the 95PPU. The *R-factor* is the average width of the 95PPU band divided by the standard deviation of the measured data. Theoretically, a simulation that exactly corresponds to measured data results in a *P-factor* of 1 and *R-factor* of zero. As a larger *P-factor* can be achieved at the expense of a larger *R-factor*, a balance must be reached between the two. When acceptable values of *R-factor* and *P-factor* are reached, then the parameter uncertainties are the desired parameter ranges. Further goodness of fit can be quantified by the *NSE* between the observations and the final “best” simulation (Abbaspour *et al.*, 2004, 2011; Schuol *et al.*, 2008).

## 4 DATA AND EXPERIMENT DESIGN

### 4.1 Dataset

In this paper, the DEM data (1:250 000) and land-use data were obtained from the Environmental and Ecological Science Data Center for West China, and the soil type data was obtained from ISSAS (Institute of Soil Science, Chinese Academy of Sciences). The daily meteorological data from 2003 to 2008 of three weather stations, and river discharge data from Yingluo Gorge station for the same period, were used.

### 4.2 Experimental design

Step one, the applicability of SWAT to the study area was assessed by calibration using discharge data for the period 2003–2005 for calibration and the data of 2006–2008 for model validation. Step two, the model was calibrated using river discharge data in the years of 2003, 2004 and 2005, respectively. And the validation period was set to be same with step one (i.e. the period of 2006–2008). Finally, the model performances for the four calibrations (i.e. calibration using the data for 2003–2005, and single year data of 2003, 2004 and 2005) in the validation period were compared using the Nash-Sutcliffe efficient (*NSE*), *P-factor* and *R-factor*, as assessment criteria.

**Table 1** Parameters selected for calibration and initial ranges.

Parameters	Definition	Initial
ALPHA_BF	Baseflow alpha factor (days)	0–1
ALPHA_BNK	Alpha factor for bankstorage	0–1
CN2	Soil Conservation Service runoff curve number for moisture condition II	70–99
ESCO	Soil evaporation compensation factor	0.8–1.5
GW_DELAY	Groundwater delay time (days)	30–450
GW_REVAP	Groundwater revap coefficient	0–0.2
GWQMN	Depth of water in the shallow aquifer required for return flow to occur (mm)	0–1.5
SMFMN	Minimum temperature-index snowmelt factor	0–7
SMFMX	Maximum temperature-index snowmelt factor	0–7
SMTMP	Snow melt base temperature (°C)	–5 to 5
SURLAG	Surface runoff lag coefficient	0–8
TIMP	Snow pack temperature lag factor	0.01–1

To reduce the impact of the automatic optimization method on the results of model calibration, the settings of SUFI-2 were made exactly same for the four calibrations. The same initial parameter ranges (Table 1) were set. The number of random parameter sets being generated was set at 1000. The iteration time for each SUFI-2 optimization was set to be three.

## 5 RESULTS AND DISCUSSION

### 5.1 Model performance using three years of calibration data

In this study, the SWAT model was applied in the hydrologic simulation for upper Heihe Basin, which used the discharge data of 2003 to 2005 for calibration and the data of 2006 to 2008 for validation. Results of the model performance are given in terms of *NSE*, *P-factor* and *R-factor* indices for both calibration and validation periods (Table 2). In the calibration, the value of *NSE* was 0.65, and 71% of the data were bracketed by the 95PPU with the *R-factor* equalling 1.08. In the validation, the value of *NSE* reached 0.67, 95PPU captured 67% of the observed data with an *R-factor* of 0.99. That was a satisfactory result for the simulation of daily discharge. The result shows that the SWAT model can be used in hydrologic simulation at Yingluo Gorge station.

**Table 2** Calibration and validation results for the four data sets.

Year	<i>P-factor</i>		<i>R-factor</i>		<i>NSE</i>	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
2003–2005	0.71	0.67	1.08	0.99	0.65	0.67
2003	0.58	0.48	0.52	0.52	0.74	0.67
2004	0.78	0.72	1.37	1.1	0.59	0.67
2005	0.66	0.5	0.66	0.54	0.66	0.67

### 5.2 Sensitivity analysis of model parameters

Sensitivity analysis of the model parameters using different discharge datasets is shown in Table 3. At the significance level of 0.05, the sensitive calibration parameters found using the 2003–2005 discharge data are ALPHA\_BNK, GW\_DEAY and SURLAG; whereas the sensitive parameters of calibration using one year of data, 2003, are ALPHA\_BNK, CN<sub>2</sub>, ESCO, GW\_DELAY and SURLAG; the sensitive parameters of calibration using year 2004 data are ALPHA\_BF, ALPHA\_BNK, and SURLAG; sensitivity parameters of calibration using only year 2005 data are ALPHA\_BNK, ESCO, GW\_REVAP and SURLA. The sensitive parameters derived from the four model calibrations are different, which indicates the information content in the four calibration datasets is different.

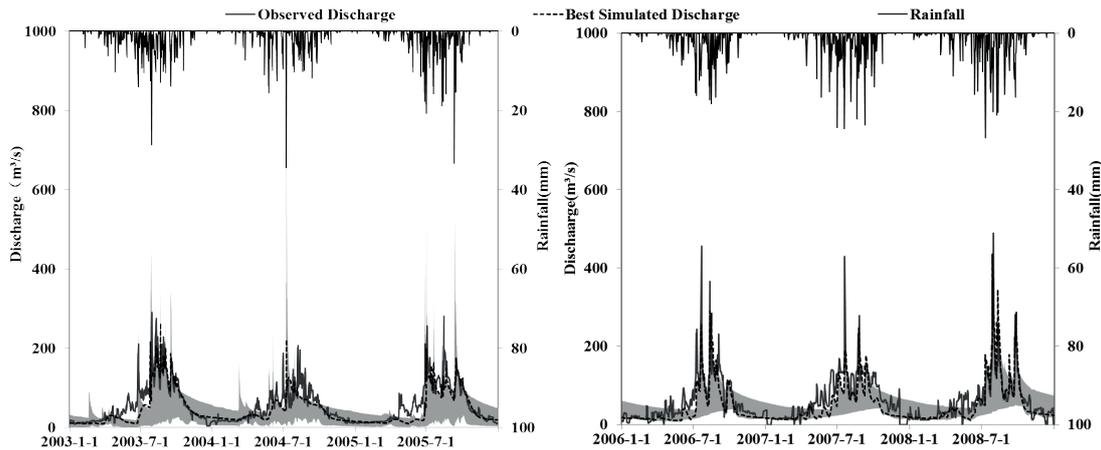
**Table 3** Results of parameter sensitivity analysis for the four data sets.

Parameters	2003–2005		2003		2004		2005	
	<i>t</i>	<i>P</i>	<i>t</i>	<i>P</i>	<i>t</i>	<i>P</i>	<i>t</i>	<i>P</i>
ALPHA_BF	0.14	0.89	-1.31	0.19	-7.13	0	-0.13	0.9
ALPHA_BNK	32.21	0	9.17	0	28.58	0	5.84	0
CN <sub>2</sub>	1.28	0.2	3.6	0	-1	0.32	0.54	0.59
ESCO	0.05	0.96	43.12	0	0.9	0.37	7.62	0
GW_DELAY	2.6	0.01	2.47	0.01	0.03	0.97	-1.29	0.2
GW_REVAP	-0.49	0.63	0.49	0.62	1.4	0.16	-2.28	0.02
GWQMN	-0.32	0.75	-0.08	0.94	-0.55	0.58	-20.67	0
SMFMN	0.86	0.39	-0.11	0.91	-1.78	0.07	-0.37	0.72
SMFMX	-0.32	0.75	0.07	0.94	-0.97	0.33	0.69	0.49
SMTMP	0.77	0.44	1.34	0.18	0.17	0.87	-0.67	0.51
SURLAG	2.02	0.04	2.19	0.03	-2.56	0.01	-3.27	0
TIMP	-0.15	0.88	-0.9	0.37	0.63	0.53	-0.63	0.53

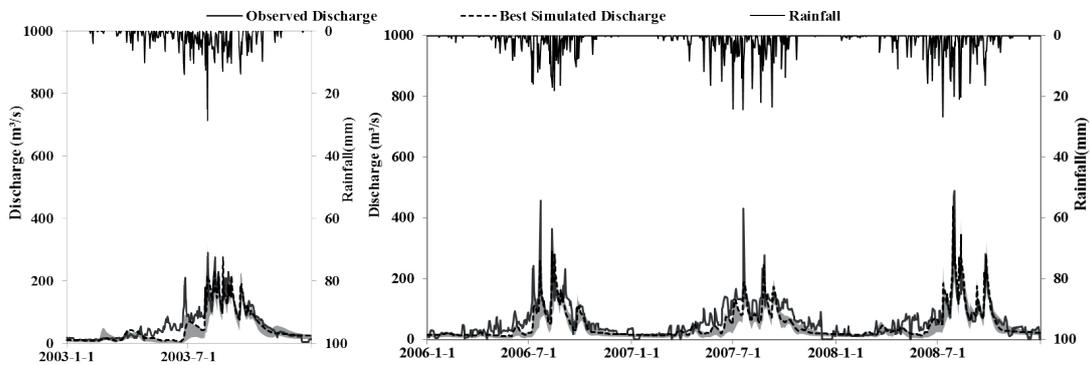
Note: *t* provides a measure of sensitivity; larger absolute values are more sensitive; *p* determines the significance of the sensitivity. A value close to zero has more significance.

### 5.3 Influence of length of calibration data on model performance

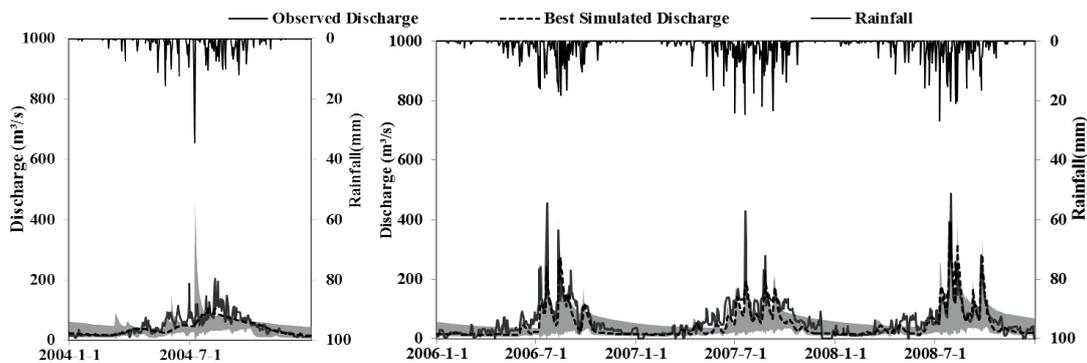
The simulation results for river discharge using four calibration data sets are shown in Figs 2–5 and Table 2. For the calibration period, the optimized *NSE* for 2003 and 2005 was higher than that using three year period data, 2003–2005. One possible reason is that the one-year data are easier to fit to the model than three-year data. The *NSE* obtained from calibration using only 2004 data is the lowest of the four calibrations, which is possibly due to the fact of the relatively poor quality of the precipitation data. *NSE* values obtained from validation using all the datasets were 0.67,



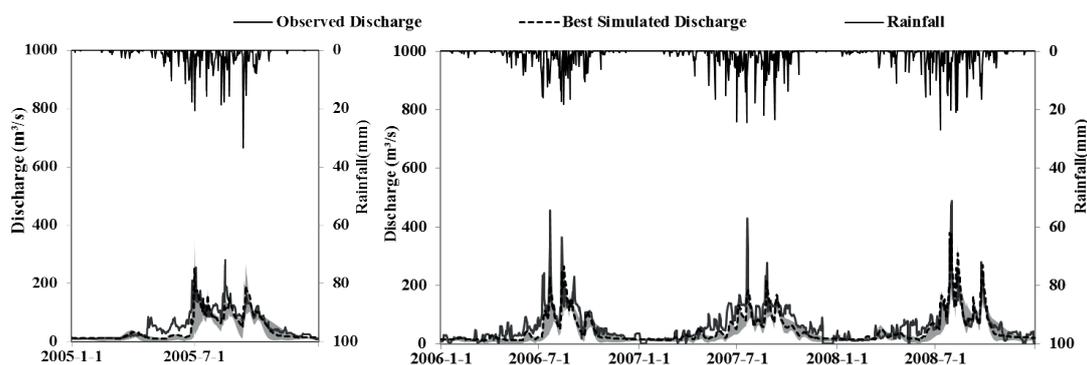
**Fig. 2** Rainfall, observed and simulated discharge and 95PPU (shaded area) for calibration using river discharge data of year 2003–2005 and validation period (2006–2008).



**Fig. 3** Rainfall, observed and simulated discharge and 95PPU (shaded area) for calibration using river discharge data of year 2003 and validation period (2006–2008).



**Fig. 4** Rainfall, observed and simulated discharge and 95PPU (shaded area) for calibration using river discharge data of year 2004 and validation period (2006–2008).



**Fig. 5** Rainfall, observed and simulated discharge and 95PPU (shaded area) for calibration using river discharge data of year 2005 and validation period (2006–2008).

indicating that the best parameter values derived from model calibration using only one year of discharge data can achieve similar model performance with that using data of three years.

In both the calibration and validation period, the values of *P-factor* obtained from calibration using data of years 2003 and 2005 are lower than calibration using the three-year data, which means that the number of observations falling into the uncertainty interval reduces, implying that the simulation uncertainty increases using the data of 2003 and 2005, respectively, compared with using three-year data. Although the value of *P-factor* for calibration using data of 2004 is higher than calibration using three-year data, the value of *R-factor* is also higher, i.e. the uncertainty band is wider, which is a sign of increased uncertainty. In general, using only one year of data for model calibration, the simulation uncertainty will be higher than calibration using three-year data.

## 6 SUMMARY AND CONCLUSION

This paper compared the results of SWAT model calibration using three-year (2003 to 2005) and one single year (2003, 2004 and 2005) of discharge data in the upper Heihe Basin. The best parameter set obtained from the SUFI-2 automatic optimization method performs similarly for the four calibrations, which indicates that using only one year of data is possible for calibration of SWAT effectively. At the same time, the simulation uncertainty for using only one year of data is higher. In general, the results of this study demonstrate that calibration of distributed hydrologic models using a smaller number of discharge measurements data than commonly used is feasible. To analyse the general applicability of using limited numbers of discharge data for the calibration of distributed type hydrological models, this method needs to be tested in more basins under different climate and runoff generation mechanisms.

**Acknowledgements** This study was supported by the National Natural Science Foundation of China (Grants: 91125015, 41201018), the National Key Technology R&D Program (Grant: 2013BAB05B04) and the Fundamental Research Funds for the Central Universities.

## REFERENCES

- Abbaspour, K.C. (2011) *User manual for SWAT-CUP, SWAT calibration and uncertainty analysis programs*. Eawag: Swiss Federal Institute of Aquatic Science and Technology, Duebendorf, Switzerland.
- Abbaspour, K.C., Johnson, A and Van Genuchten, M.T. (2004) Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone Journal* 3(4), 1340–1352.
- Arnold, J.G., et al. (1998) Large area hydrologic modeling and assessment. Part I: Model development. *JAWRA Journal of the American Water Resources Association* 34, 73–89. doi: 10.1111/j.1752-1688.1998.tb05961.x.
- Blöschl, G. and Sivapalan, M. (1995) Scale issues in hydrological modelling: A review. *Hydrol. Processes* 9, 251–290. doi: 10.1002/hyp.3360090305.
- Chai Xiaoling, et al. (2005) Runoff analysis and calculation methods for the regions un-gauged. *China Rural Water and Hydropower* 5, 20–22.
- Li Hongxia, et al. (2011) Review of hydrological model parameter regionalization for ungauged catchments. *Journal of China Hydrology* 31(3), 13–17.
- Li Zhijia, Li Jiren and Han Congshang, (1992) A synthetic method for simulating both discharge and stage of flood in a channel. *Yellow River* 2, 16–20.

- Nash, J.E. and Sutcliffe, J.V. (1970) River flow forecasting through conceptual models Part I – A discussion of principles. *Journal of Hydrology* 10(3), 282–290.
- Perrin, C., *et al.* (2007) Impact of limited streamflow data on the efficiency and the parameters of rainfall–runoff models. *Hydrological Sciences Journal* 52(1), 131–151.
- Schuol, J., *et al.* (2008) Modeling blue and green water availability in Africa. *Water Resour. Res.* 44, W07406, doi:10.1029/2007WR006609.
- Seibert, J. and Beven, K.J. (2009) Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci.* 13(6), 883–892.
- Sun, W.C., Ishidaira, H. and Bastola, S. (2010) Towards improving river discharge estimation in ungauged basins: calibration of rainfall-runoff models based on satellite observations of river flow width at basin outlet. *Hydrol. Earth Syst. Sci.* 14, 2011-2022, doi:10.5194/hess-14-2011-2010.
- Sun, W., Ishidaira, H. and Bastola, S. (2012) Calibration of hydrological models in ungauged basins based on satellite radar altimetry observations of river water level. *Hydrol. Processes* 26, 3524–3537. doi:10.1002/hyp.8429.
- Wang Shugong, Kang Ersi and Li Xin. (2004) Progress and perspective of distributed hydrological models. *Journal of Glaciology and Geocryology* 26(1), 61–65.
- Xu Zongxue and Cheng Lei. (2010) Progress on studies and applications of the Distributed Hydrological Models. *Journal of Hydraulic Engineering* 41(9), 1009–1017.
- Yang, J., *et al.* (2008) Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *Journal of Hydrology* 358, 1–23.
- Young, A.R. (2006) Stream flow simulation within UK ungauged catchments using a daily rainfall–runoff model. *Journal of Hydrology* 320, 155–172.