Journal of Hydrology 434–435 (2012) 36–45

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff

Hua Chen [a,b,*], Chong-Yu Xu [a,b], Shenglian Guo [a]

[a] State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, 430072, China
[b] Department of Geosciences, University of Oslo, PO Box 1047 Blindern, NO-0316 Oslo, Norway

ARTICLE INFO

Article history:
Received 7 May 2011
Received in revised form 8 February 2012
Accepted 18 February 2012
Available online 28 February 2012
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Martin Beniston, Associate Editor

Keywords:
Climate change
Statistical downscaling
GCM
Hydrological models
Hanjiang basin
China

SUMMARY

In this study a rigorous evaluation and comparison of the difference in water balance simulations resulted from using different downscaling techniques, GCMs and hydrological models is performed in upper Hanjiang basin in China. The study consists of the following steps: (1) the NCEP/NCAR reanalysis data for the period 1961–2000 are used to calibrate and validate the statistical downscaling techniques, i.e. SSVM (Smooth Support Vector Machine) and SDSM (Statistical Downscaling Model); (2) the A2 emission scenarios from CGCM3 and HadCM3 for the same period are used as input to the statistical downscaling models; and (3) the downscaled local scale climate scenarios are then used as the input to the Xin-anjiang and HBV hydrological models. The results show that: (1) for the same GCM, the simulated runoffs vary greatly when using rainfall provided by different statistical downscaling techniques as the input to the hydrological models; (2) although most widely used statistics in the literature for evaluation of statistical downscaling methods show SDSM has better performance than SSVM in downscaling rainfall except the Nash–Sutcliffe efficiency (NSC) and root mean square error-observations standard deviation ratio (RSR), the runoff simulation efficiency driven by SDSM rainfall is far lower than by SSVM; and (3) by comparing different statistics in rainfall and runoff simulation, it can be concluded that NSC and RSR between simulated and observed rainfall can be used as key statistics to evaluate the statistical downscaling models' performance when downscaled precipitation scenarios are used as input for hydrological models.

1. Introduction

Studies of climate change impacts on water resources have become a hot topic recently. Climate models (global and regional) and hydrological models are the important tools used in these studies (Boe et al., 2007; Chen et al., 2007; Gleick, 1987; Guo et al., 2002; Xu, 2000; Xu et al., 2005). However, there exist many key challenges in the application of GCMs and hydrological models (Fowler et al., 2007; Xu, 1999). First, the spatial scales of GCMs and hydrological models are inconsistent. Therefore, the output of GCMs cannot be directly used as input to hydrological models. Second, the accuracy of precipitation simulations from GCMs cannot meet the requirements of hydrological simulations. The dynamic downscaling and statistical downscaling are the most commonly used methods in the one-way coupling of GCMs and hydrological models (Bergstrom et al., 2001; Fowler et al., 2007; Pinto et al., 2010; Schoof et al., 2009; Wilby et al., 1999). Dynamic downscaling models, i.e., regional climate models (RCMs) have clear physical meanings, however they are computationally expensive. Statistical downscaling models, on the other hand, are based on statistical relationship and hence require less computational time.

It is a well accepted fact that considerable errors exist in all three steps in assessment of climate change impacts, knowing as uncertainty of climate modeling, uncertainty of downscaling techniques, and uncertainty of hydrological modeling. Some scholars have already compared and analyzed the uncertainties of climate change impacts on runoff by using different downscaling techniques and hydrological models under different scenarios. Dibike and Coulibaly (2005) applied two types of statistical (a stochastic and a regression based) downscaling techniques and two hydrological models to simulate the corresponding future flow regime in a catchment. Prudhomme and Davies (2009) used a lumped conceptual rainfall–runoff model, three GCMs and two downscaling techniques to investigate the climate change impacts on river flows. Chiew et al. (2010) assessed the runoff simulated by the SIMHYD rainfall–runoff model with daily rainfall which was downscaled from three GCMs using five downscaling models. Segui et al. (2010) evaluated the uncertainty related to climate change

* Corresponding author at: State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, 430072, China. Tel./fax: +86 2768773568.
E-mail address: chua@whu.edu.cn (H. Chen).

0022-1694/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2012.02.040

impacts on water resources by applying a distributed hydrological model and three different downscaling techniques. These literatures are basically the application of GCM scenarios for driving the hydrological models to obtain runoff series of different scenarios. But so far, not a single GCM or hydrological model is found to be dramatically better than others. This is partly because evaluation of statistical downscaling models has been focused on the comparison of rainfall statistics obtained from GCMs or downscaling models, rather than on evaluation of the usefulness of the rainfall data series in driving hydrological models for water balance calculations. We argue that one of the ultimate purposes of downscaling methods is to provide rainfall (or other meteorological) input for hydrological models for simulation/prediction of discharge and other water balance components. Therefore, the evaluation of downscaling methods should also include the comparison of the results of hydrological simulations in the study of climate change impacts on runoff.

In this study, the application of statistical downscaling techniques in providing rainfall input for hydrological models for climate change impacts studies in Hanjiang River basin is assessed through one-way coupling of multiple GCMs, multiple downscaling methods and multiple hydrological models. The main aim of the study is to evaluate and compare the uncertainty in the simulated runoffs resulting from different impact assessment steps, including the use of NCEP/NCAR reanalysis data, GCMs predictions, downscaling models' results, and hydrological models' simulations. The technique route of this study was drawn in Fig. 1. The aim is achieved through following steps: (1) establishment and evaluation of the statistical downscaling methods and hydrological models in the Hanjiang basin; (2) evaluate and compare the uncertainty in the simulated runoffs resulting from using two statistical downscaling methods, two GCMs and two hydrological models; (3) discuss the usefulness of evaluation criteria for downscaling methods with respect to the performance of the hydrological models in simulating runoff using different rainfall inputs provided by the combinations of GCMs and downscaling methods.

2. Study area and data

2.1. Study area

The upper Hanjiang River basin with an area of 59,115 km^2 is selected as the study region (Fig. 2). Hanjiang River is the largest tributary of the Yangtze River, which rises on the southern slope of the Qin Ling Mountains and flows through Shaanxi and Hubei provinces, falls into the Yangtze River at Wuhan, with a length of 1577 km and a drainage area of about 159,000 km^2 (Chen et al., 2007). The longitude and latitude of Hanjiang basin is 106°15′–114°20′E and 30°10′–34°20′N respectively. For the base period of 1961–2000, the annual average temperature is 12–16 °C; the water surface evaporation is 700–1100 mm, and land surface evaporation is 400–700 mm, increasing from southwest to northeast. The Hanjiang basin locates in the East Asian subtropical monsoon region with annual average rainfall of 873 mm, mainly comes from the southeast and southwest warm air. The rainfall is unevenly distributed during the year with the maximum rainfall of 4 consecutive months accounts for 55–65% of the annual rainfall,

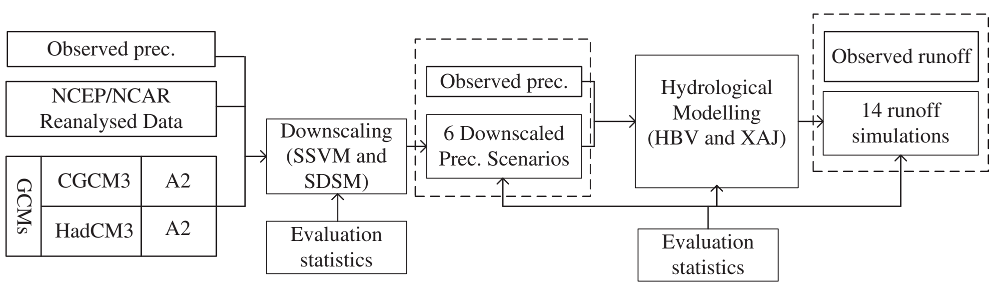

Fig. 1. The technique route in this study.

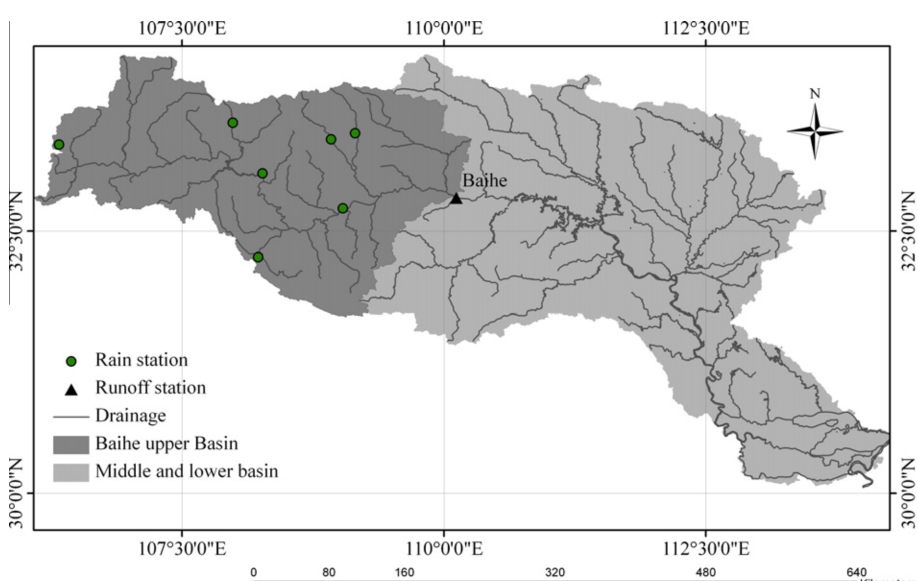

Fig. 2. The location of the study area.

Table 1
The selection of large-scale climate factors for the statistical downscaling methods in this study.

Daily variable	Code
Precipitation (mm)	prec
Maximum temperature (K)	tmax
Minimum temperature (K)	tmin
Mean temperature	temp
Mean sea level pressure	mslp
500 hPa geopotential height	p500
850 hPa geopotential height	p850
Near surface relative humidity	rhum
Relative humidity at 500 hPa height	r500
Relative humidity at 850 hPa height	r850
Near surface specific humidity	shum
Geostrophic airflow velocity	**_f
Vorticity	**_z
Zonal velocity component	**_u
Meridional velocity component	**_v
Wind direction	**th

decreasing from south to north and from west to east. During these months the rainfall of the upper Hanjiang River accounts for 60–65% of the annual rainfall. The flood season is from May to October in the upper river, the rainfall of flood season accounts for 75–80% of the annual rainfall.

2.2. Data

Seven hydro-metrological stations with daily air temperature, rainfall, and pan evaporation data and discharge data from Baihe runoff station (as shown in Fig. 1) are selected in this study. The quality of the data has been checked and verified in terms of homogeneity and consistency in earlier studies (Chen et al., 2007, 2010) and no missing data is found in the period of 1961–2000. The calibration and validation periods of the hydrological models are from 1976 to 1986 and from 1987 to 1988, respectively. The daily NCEP/NCAR reanalysis data are used as the observed large-scale climate data to calibrate and validate the downscaling models during the period of 1961–2000. CGCM3 SRES A2 (DAI CGCM3 Predictors, 2008) and HadCM3 SRES A2 (Pope et al., 2000) are used to produce regional climate scenarios through downscaling, which in turn are used to simulate hydrological responses. The grid spatial resolution of large-scale climate factors is 2.5×2.5 degrees, covering four grids over the upper Hanjiang basin. Sixteen large-scale climate factors are used in this study (Table 1), which will be screened further during the establishment of statistical downscaling methods.

3. Statistical downscaling methods and hydrological models

Statistical downscaling methods have been widely used in recent years. This study chooses two typical statistical downscaling methods: Smooth Support Vector Machine (SSVM) (Chen et al., 2010) and Statistical Downscaling Model (SDSM) (Wilby et al., 2002). The first technique is a relatively new learning method in statistical learning theory, while the second statistical downscaling technique is based on regression analysis. To investigate the potential difference between different hydrological models for climate impact study, two widely-used hydrological models are utilized in this study, which are the Xin-anjiang model (Zhao, 1992) and the HBV model (Bergstrom, 1975).

3.1. Statistical downscaling methods

3.1.1. Smooth Support Vector Machine (SSVM)

Support Vector Machine (SVM) is a new supervised learning method proposed by Vapnik (1998), based on the Vapnik–Chervonenkis (VC) dimension and Structural Risk Minimization (SRM). It can find the best compromise between the model complexity and learning ability through the limited sample information. It has a good ability of prediction and can address the small sample, nonlinear, high dimension and local minimum points and other practical issues. It has become one of the research focuses in machine learning research field and has been successfully applied into classification and regression. Lee and Mangasarian (2001) proposed a new Smooth Support Vector Machine (SSVM) to simplify the SVM training and further reduce the computational complexity. The constrained quadratic optimization problem is converted into the unconstrained convex quadratic optimization problem by using the smoothing method. The experiments show that, SSVM performs better than SVM algorithm in solving problems. Chen et al. (2010) established the statistical relationship between the GCM atmospheric predictors and the observed rainfall in Hanjiang basin and evaluated the simulation ability of the model.

3.1.2. Statistical Downscaling Model (SDSM)

Statistical Downscaling Model (SDSM) is a decision support tool for assessing local climate change impacts, established by Wilby et al. (2002) basing on Windows interface. The model, which incorporates the weather generator and the multiple linear regression technique, is a hybrid statistical downscaling method. After nearly 10 years of development, SDSM has grown to the forth generation, and has been widely used in the climate change studies. SDSM's workflow includes two parts: First, establish the statistical relationship between the predictand and predictors and to determine the required parameters for the weather generator, including data quality control and transformation, screening of predictor variables, model calibration and weather generation (using observed predictors); second part is to simulate the future series of predictand by using the predicted data from GCMs and the parameters generated in the first step.

3.1.3. Evaluation statistics of statistical downscaling

Maraun et al. (2010) summarized that according to the application of the impact study, different statistics of the downscaled precipitation may be of interest, including intensity metrics, temporal and spatial characteristics as well as metrics characterizing relevant physical processes. Statistics regarding precipitation intensity are mean, variance and quantiles or parameters of the precipitation distribution. Temporal statistics are the autocorrelation function, the annual cycle, inter-annual and decadal variability and trends, or measures focusing on the precipitation occurrence such as wet day probabilities, transition probabilities (wet–wet) and the length of wet and dry spells (Maraun et al., 2010, 2011). Extreme measures for temporal statistics include the maximum number of consecutive dry days. The statistics in Table 2 are commonly used in evaluation of statistical downscaling methods in downscaling rainfall (Khan and Coulibaly, 2010; Wetterhall et al., 2006), and are chosen to evaluate the statistical downscaling methods in different seasons (Winter: DJF; Spring: MAM; Summer: JJA; Autumn: SON) in this study. The wet-threshold is chosen as 1.0 mm in this study.

3.2. Hydrological models

3.2.1. Xin-anjiang model

The Xin-anjiang (XAJ) model (Zhao, 1992; Zhao et al., 1980) was first used in prediction of Xin-anjiang Reservoir inflow, and later on became a rainfall–runoff model for general use. Its major feature is the concept of runoff formation as a dependent variable of repletion of storage, i.e., runoff is not produced until the soil moisture content of the aeration zone reached field capacity, and thereafter, runoff is equal to the rainfall excess without further loss. XAJ model, with three runoff components, has been widely used in humid

Table 2
Selection and definition of indicators for evaluation of statistical downscaling methods in this study.

Indicators	Definition
Mean	Average of all values
Variance	Variance of all values in each time period
Percentile	Value of the User specified percentile
Maximum 5-day total	Maximum total accumulated over 5-days
Percentage wet	Percentage of days that exceed the threshold
Maximum dry spell length	Longest spell with amounts less than the wet-day threshold
Maximum wet spell length	Longest spell with amounts greater than or equal to the wet-day threshold
Peaks over threshold	Count of peaks over User specified threshold (defined as a percentile of all data)

or semi-humid regions in China as well as in many other countries (Jiang et al., 2007; Yang et al., 2010; Zhang and Lindstrom, 1996).

3.2.2. HBV model

The HBV model is a conceptual hydrological model and it was originally developed at the Swedish Meteorological and Hydrological Institute (SMHI) for runoff simulation and hydrological forecasting in the early 1970s (Bergstrom, 1975). It consists of routines for snow accumulation and melting, soil moisture accounting, runoff response, and finally a flow routing procedure. The model is based on a sound scientific foundation and can meet its data demands in most areas, which has the scope of applications in more than 40 countries (Ashagrie et al., 2006; Bergstrom et al., 2001; Jin et al., 2009; Seibert et al., 2010; Yu and Wang, 2009).

3.2.3. Evaluation criteria for hydrological models

The Percent Bias (PBIAS) between the observed and simulated discharge, Nash–Sutcliffe efficiency (NSC) (Nash and Sutcliffe, 1970) and Root Mean Square Error (RMSE)-observations standard deviation ratio (RSR) were selected to evaluate the merits of the hydrological models in this study.

PBIAS, expressed in percentage, measures the average tendency of the simulated data to be larger or smaller than their observed counterparts (Gupta et al., 1999) and is calculated with the following equation:

$$\mathrm{PBIAS} = \left[\frac{\sum_{i=1}^{n}(Q_i^{obs} - Q_i^{sim}) * 100}{\sum_{i=1}^{n} Q_i^{obs}}\right] \tag{1}$$

where Q_i^{obs} is the ith day observed discharge, Q_i^{sim} is the ith day simulated discharge and n is the total number of days in the runoff series. The optimal value of PBIAS is 0.0, with positive values indicating model underestimation bias, negative values indicating model overestimation bias and low-magnitude values indicating accurate model simulations.

NSC is a normalized statistic that determines the relative magnitude of the residual variance ("noise") compared to the measured data variance ("information") (Nash and Sutcliffe, 1970). It is computed by the following equation:

$$\mathrm{NSC} = 100 * \left(1 - \left[\frac{\sum_{i=1}^{n}(Q_i^{obs} - Q_i^{sim})^2}{\sum_{i=1}^{n}(Q_i^{obs} - Q_{mean}^{obs})^2}\right]\right) \tag{2}$$

NSC ranges from $-\infty$ to 1. A value of 1% or 100% corresponds to a perfect match of the simulation to the observation. An efficiency of 0 indicates that the model simulations are as accurate as the mean of the observation, whereas an efficiency less than zero occurs when the observed mean is a better predictor than the model simulation or, in other words, when the residual variance is larger than the data variance.

Moriasi et al. (2007) developed a model evaluation statistic, named the RMSE-observations standard deviation ratio (RSR). RSR standardizes RMSE using the observations' standard deviation and is calculated as the ratio of the RMSE and the standard deviation of measured data, as shown in the following equation:

$$\mathrm{RSR} = \frac{\mathrm{RMSE}}{\mathrm{STDEV}_{obs}} = \frac{\sqrt{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_i^{sim}\right)^2}}{\sqrt{\sum_{i=1}^{n}\left(Q_i^{obs} - Q_{mean}^{obs}\right)^2}} \tag{3}$$

RSR incorporates the benefits of error index statistics and includes a scaling/normalization factor, so that the resulting statistic and reported values can apply to various constituents. RSR varies from the optimal value of 0 to a large positive value and the lower the RSR, the lower the RMSE, and the better the model simulation performance. In general, hydrological model simulation can be judged as "satisfactory" if NSC > 0.50 and RSR < 0.70, and if PBIAS ±25% for streamflow (Moriasi et al., 2007).

4. Results

4.1. The establishment of SSVM and SDSM

The statistical relationship between large-scale circulation factors and the rainfall of hydrological stations above the upper Hanjiang basin is established by using SSVM and SDSM. The calibration period is from 1961 to 1990, and the evaluation period is from 1991 to 2000. The selection of large-scale climate factors is a very important and crucial step in the study of statistical downscaling methods. In order to obtain the most relevant factors with the rainfall of the basin, the Screen Variables of SDSM are used to select the large-scale climate factors. All the factors in Table 1 are selected by setting the measured rainfall, and the correlation coefficient 0.3 of large-scale factors is used as the threshold. The selected factors are shown in the shadow part of Table 1. There are four factors been selected in each grid, and 16 large-scale climate factors are used as predictors in the establishment of statistical downscaling models.

SSVM and SDSM were established by utilizing the NCEP/NCAR reanalysis data and the observed rainfall data of each hydrological station. To comprehensively evaluate the performance of SSVM and SDSM, the mean values of the observations and simulations of seven hydrological stations were calculated and compared. The evaluation of SSVM and SDSM was provided in Table 3, where shadowed values show the better skills between the two models. It is evident that SDSM has better performance than SSVM in simulating rainfall as most statistics' values of SDSM are in shadow area in the calibration and validation periods. In the calibration period, SDSM has an overwhelming advantage to SSVM, because most statistics of SDSM are better than those of SSVM. In the validation period, there are only a few statistics of SSVM which perform better than those of SDSM, such as standard deviation and 5-day maximum rainfall in summer, autumn and annual, maximum dry spell length in spring and winter, peaks over threshold in spring and autumn, annual percentile and autumn' mean. Fig. 3 shows the monthly statistics' values of the simulated rainfall by SSVM and SDSM from NCEP/NCAR reanalysis data in the validation period, reflecting similar conclusions with the above analysis. It can be seen from the above analysis that SDSM has better capacity than SSVM to downscale and simulate rainfall from large scale climatic predictors in the region as far as the commonly used statistics are concerned. In general, the simulation of rainfall of the both methods seems acceptable; however, their usefulness and weakness in driving hydrological models will be further evaluated in the following sections.

Table 3
Mean values of evaluation criteria of SSVM and SDSM in calibration (1961–1990) and validation (1991–2000) period computed from the meteorological seven stations.

Indictors	Seasons	Calibration			Validation		
		Observed	SSVM	SDSM	Observed	SSVM	SDSM
Mean (mm/d)	DJF	0.28	0.40	0.25	0.26	0.36	0.23
	MAM	2.11	2.01	1.93	1.78	1.63	1.89
	JJA	4.53	4.69	4.35	4.47	4.88	4.30
	SON	2.84	2.75	2.71	2.14	2.48	2.80
	Annual	2.45	2.47	2.32	2.17	2.35	2.31
Standard deviation (mm/d)	DJF	0.92	0.61	0.65	0.91	0.67	0.67
	MAM	4.70	3.79	4.24	3.93	3.49	4.30
	JJA	7.95	7.27	7.36	8.18	8.57	7.45
	SON	6.17	4.81	6.11	5.10	4.18	7.23
	Annual	5.79	5.02	5.45	5.45	5.36	5.83
Percentile (95)	DJF	1.54	1.47	1.43	1.30	1.40	1.24
	MAM	11.11	9.23	9.82	10.36	7.34	10.12
	JJA	19.76	17.45	18.56	22.73	19.74	19.93
	SON	15.03	12.88	13.22	12.45	10.29	11.61
	Annual	13.43	11.97	12.04	12.50	12.52	11.88
5-day maximum rainfall (mm)	DJF	24.10	12.65	10.82	17.16	10.76	13.14
	MAM	75.16	63.15	72.51	61.51	65.87	58.94
	JJA	163.08	123.52	136.26	121.32	124.30	124.95
	SON	143.36	107.88	106.25	101.06	101.47	158.05
	Annual	163.08	123.52	136.26	121.32	124.30	158.05
Percentage wet	DJF	0.08	0.11	0.08	0.07	0.11	0.07
	MAM	0.31	0.51	0.34	0.30	0.50	0.32
	JJA	0.53	0.72	0.54	0.48	0.70	0.53
	SON	0.35	0.58	0.38	0.29	0.57	0.37
	Annual	0.32	0.48	0.34	0.29	0.47	0.32
Maximum dry spell length	DJF	77	79	78	60	48	91
	MAM	21	17	33	15	12	22
	JJA	12	8	10	17	7	21
	SON	28	18	29	30	9	27
	Annual	56	72	53	64	48	74
Maximum wet spell length	DJF	4	6	3	3	4	2
	MAM	9	20	9	5	18	6
	JJA	18	28	12	13	26	11
	SON	14	21	10	9	22	7
	Annual	18	28	12	13	26	11
Peaks over threshold (>90%)	DJF	0	0	0	0	0	0
	MAM	66	68	64	11	16	19
	JJA	177	305	191	69	132	69
	SON	106	154	113	25	23	30
	Annual	349	527	368	105	171	118

4.2. Hydrological models' calibration and parameter optimization

The calibration period is from 1976 to 1986 and the validation period is from 1987 to 1988 for the two hydrological models. The parameters of the hydrological models are optimized with three algorithms, namely Rosenbrock (Rosenbrock, 1960), simplex (Nelder and Mead, 1965; Spendley et al., 1962) and genetic (Wang, 1991). Model calibration and evaluation results are shown in Table 4, which shows that both XAJ and HBV models have high performance in reproducing historical flow data for the study basin. The NSC is about 85%, PBIAS is equal to zero and less than 3% and RSR is less than 0.50 in the calibration and validation periods, respectively. Fig. 4 shows the measured and simulated discharge series during the validation period from May to November in 1987 for illustrative purpose. Both Table 4 and Fig. 4 reveal that the two models work well in the study basin in reproducing the historical flow and in simulation of flood peaks.

5. Evaluation of the uncertainty of the GCMs, downscaling methods and hydrological models in the study basin

In order to compare and analyze the uncertainty of the GCMs, downscaling methods and hydrological models, observed rainfall and six rainfall scenarios simulated by using SSVM and SDSM with the NCEP/NCAR reanalysis data, CGCM3 and HadCM3 SRES A2 scenarios during 1961–2000 as large scale predictors, were used as inputs for XAJ and HBV models. Figs. 5–7 and Table 5 show the runoff simulation results of different rainfall scenarios. These results are discussed in the following sections.

5.1. Evaluation of different downscaling methods with same hydrological models and NCEP/NCAR reanalysis data

Monthly mean and standard deviation (STD) of runoff simulated by using precipitation scenarios downscaled from the NCEP reanalysis data were drawn in Fig. 5. It can be seen from Fig. 5 that the monthly means and STD of the simulation results are close to those of observations, and it is difficult to judge which model performs better results based on these two criteria. Three other statistics (PBIAS, NSC and RSR) were therefore used to evaluate the performance of the runoff simulations driven by precipitation scenarios downscaled from the NCEP reanalysis data by using SSVM and SDSM. The results were listed in Table 5 where it is evident that the error of runoff simulation by using the rainfall inputs obtained from SDSM is significantly greater than that of SSVM. For XAJ model, the NSC values (11.62% monthly, −28.19% daily) by SDSM are

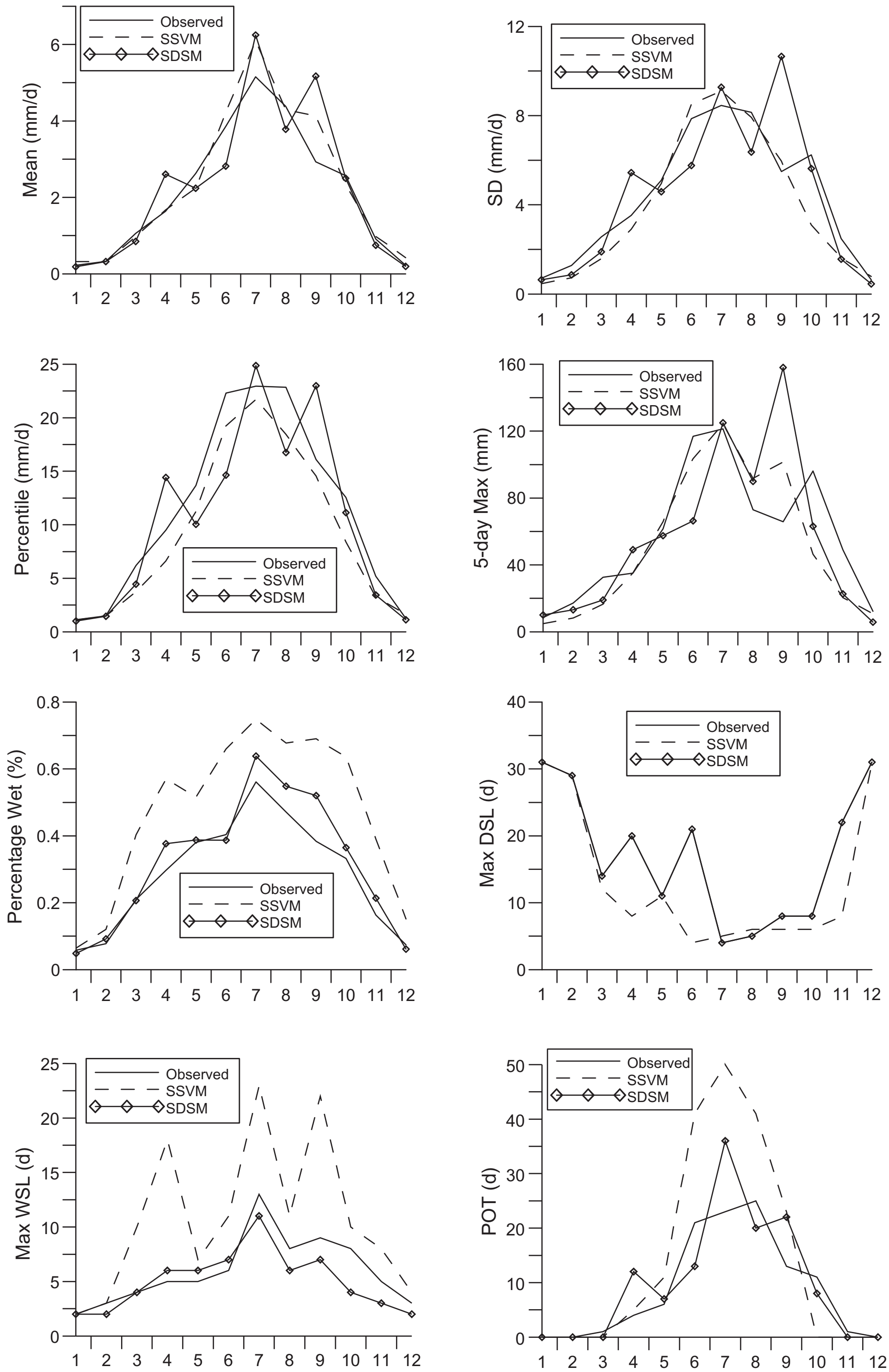

Fig. 3. Comparison of mean monthly rainfall statistics downscaled from NCEP/NCAR reanalysis data using SSVM and SDSM.

significantly lower than those by SSVM, which are 68.35% and 54.42% for monthly and daily time steps respectively. The values of RSR by SDSM are also higher than those by SSVM, while there is no big difference between the values of PBIAS. The results of HBV are similar to those of XAJ. This result is different from what was concluded based on the commonly used statistics as showed in Table 3, where SSVM has lower skills compared with SDSM. Keeping in mind that one of the main uses of rainfall data should be for water balance calculations and for serving as input to hydrological models for flow simulation, the contradiction results of Tables 5 and 3 mean that the commonly used statistics for evaluation of downscaling methods are not determinative in terms of their usefulness in providing rainfall data for hydrological analysis. There is a need for reconsideration and selection of criteria for evaluation of downscaling methods that are to be used for providing rainfall input to hydrological models.

Table 4
Values of evaluation criteria of hydrological models in calibration (1976–1986) and evaluation (1987–1988) periods.

Hydrological model	Calibration			Validation		
	PBIAS (%)	NSC (%)	RSR	PBIAS (%)	NSC (%)	RSR
HBV	0	85.91	0.37	2.84	85.72	0.40
XAJ	0	84.58	0.37	0.27	85.38	0.39

5.2. Evaluation of different hydrological models with same scenario and downscaling methods

As many hydrological models have been applied in the study of climate change impacts on hydrology and water resources, it is desirable to compare their performance for this purpose and to choose more suitable downscaling models in a certain region. It can be found from Table 5 that there is no obvious difference between NSC and RSR statistics by XAJ model and HBV model, however Fig. 6 shows that when using rainfall downscaled by SSVM from CGCM3 A2 scenario as input, both monthly mean values and STD of model simulated discharge by XAJ show better agreement with observations than those simulated by HBV model in this region. Similar conclusions can be drawn for other precipitation scenarios.

5.3. Evaluation of different GCMs and scenarios with same hydrological models and downscaling methods

It is easily acceptable that the uncertainty of GCMs is greater than that of NCEP/NCAR reanalysis data, which is also a key barrier in the climate change impact studies. To compare the uncertainty

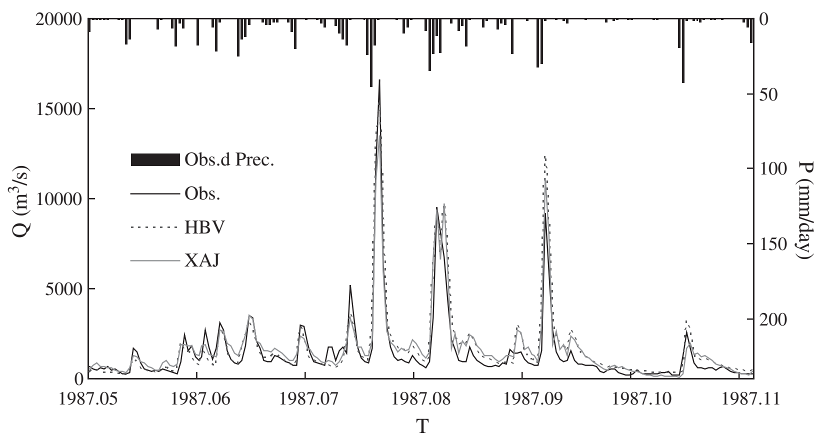

Fig. 4. Observed and simulated discharge hydrograph in the period of May to November in 1987.

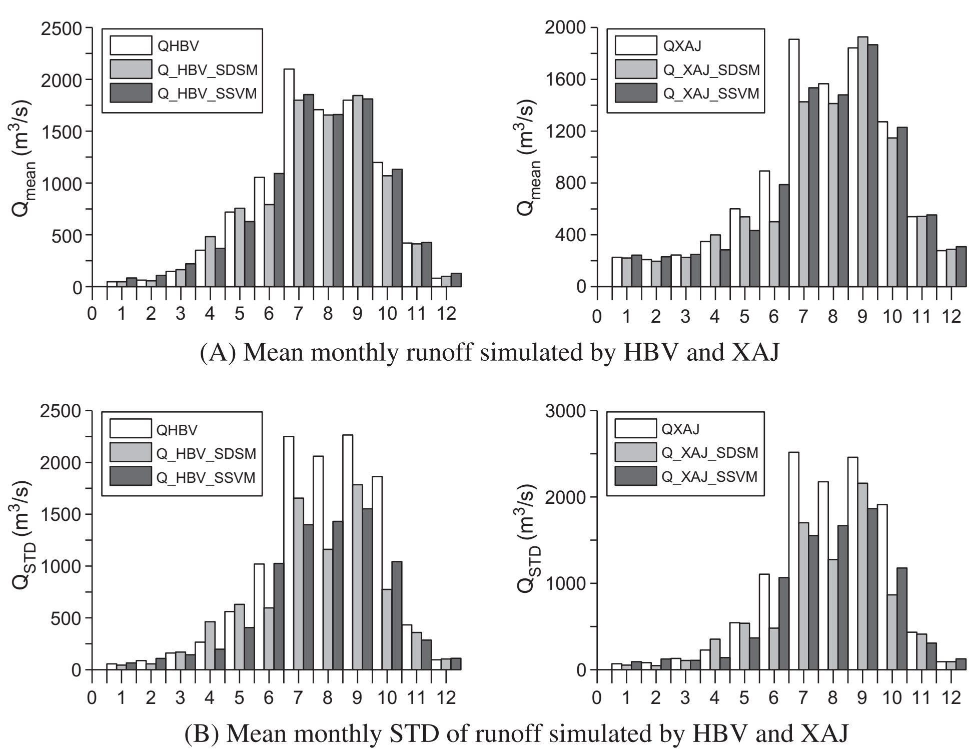

Fig. 5. The comparison of runoff simulation driven by NCEP/NCAR reanalysis data downscaled by SSVM and SDSM.

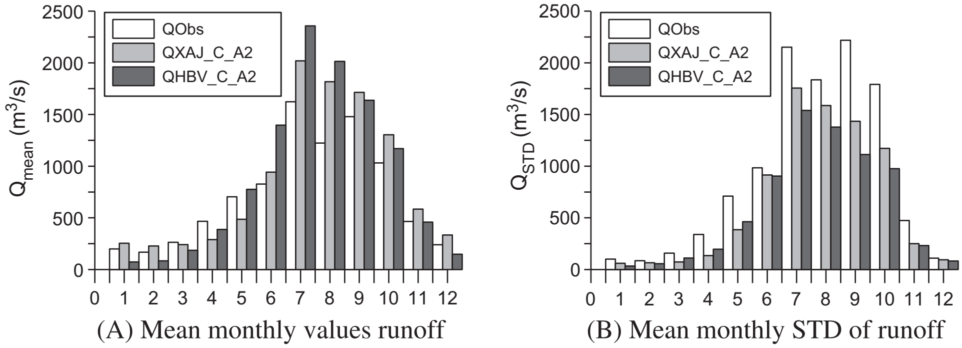

Fig. 6. The comparison of runoff simulation driven by CGCM3 A2 scenario downscaled by SSVM.

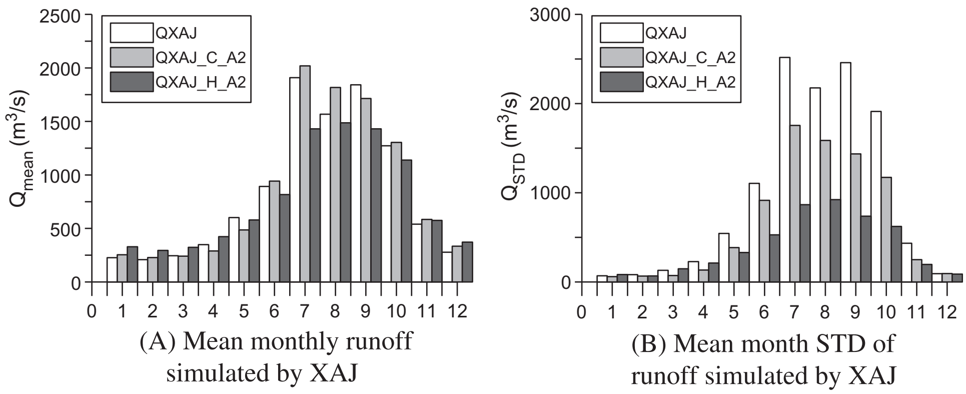

Fig. 7. The comparison of runoff simulation driven by CGCM3 and HadCM3 A2 scenario downscaled by SSVM.

Table 5
Runoff simulation results of different rainfall scenarios for the period of 1961–2000.

Hydrological models	Prec. scenarios	Daily			Monthly		
		PBIAS (%)	NSC (%)	RSR	PBIAS (%)	NSC (%)	RSR
XAJ	P_Observed	−9.76	78.80	0.46	−9.76	85.50	0.38
	P_SSVM-NCEP	−4.32	54.42	0.68	−4.32	68.35	0.56
	P_SDSM-NCEP	6.56	−28.19	1.13	6.56	11.62	0.94
HBV	P_Observed	−7.58	79.33	0.45	−7.58	81.91	0.43
	P_SSVM-NCEP	−8.18	54.05	0.68	−8.18	68.45	0.56
	P_SDSM-NCEP	0.91	−25.83	1.12	0.91	6.68	0.97

of CGCM3 and HadCM3 in terms of the results of runoff simulation, the monthly mean values and STD of runoff simulated by XAJ, driven by precipitation downscaled by SSVM from CGCM3 and HadCM3 A2 scenario and by observed precipitation, were plotted in Fig. 7. It is easily found in Fig. 7 that the monthly means of runoff simulations driven by observed precipitation agree better with those driven by CGCM3 A2 scenario than by HadCM3 A2 scenario. The same conclusion can be drawn by comparing their monthly STD. When comparing results obtained from the other combinations, like SSVM and HBV, SDSM and HBV, and SDSM and XAJ, the similar findings can be obtained that CGCM3 is more appropriate than HadCM3 for studying the climate change impact in the Hanjiang basin.

5.4. Comparison of rainfall evaluation statistics in terms of performance of runoff simulations

It was shown in Table 3 and Fig. 2 that most statistics used in evaluation of the performance of downscaling methods are better for SDSM than those for SSVM. However, Table 5 shows that the NSC of runoff simulations by using the rainfall downscaled from SDSM is much lower than that from SSVM. This study clearly shows that most of the statistics used in this study and in the literature for evaluation of downscaling methods cannot fulfill the need of hydrological modeling study. Therefore, it is necessary to reconsider and select more appropriate statistics for the assessment of the accuracy of rainfall simulation in the downscaling methods when the study of hydrological response to climate change is conducted.

The PBIAS, NSC and RSR of the downscaling methods for rainfall simulations from NCEP/NCAR reanalysis data were calculated and listed in Table 6. It can be seen from Table 6 that the NSC values for the SSVM method are positive with the highest value of 52.38% (daily) and 82.81% (monthly), while the NSC values for the SDSM method are much lower than those for SSVM, even negative value is obtained for daily simulation. Daily and monthly RSR values by SSVM are lower than those by SDSM in Table 6. Although absolute value of PBIAS of SSVM is higher than that of SDSM, the difference between them is small. As a consequence, according to the NSC and RSR values in Table 5, the runoff simulations driven

Table 6
The PBIAS, NSC and RSR of rainfall simulations by using SDSM and SSVM during the period of 1961–2000.

Precipitation scenarios	Daily			Monthly		
	PBIAS (%)	NSC (%)	RSR	PBIAS (%)	NSC (%)	RSR
P_SDSM-NCEP	2.62	−77.41	1.33	2.62	41.46	0.77
P_SSVM-NCEP	−5.01	52.38	0.69	−5.01	82.81	0.41

by precipitation scenarios from SSVM have a better performance than those from SDSM. The comparison results reveal that NSC and RSR may be proper statistics for evaluation of downscaling methods in terms of their usefulness in providing rainfall data for hydrological simulations. The results also reflect that rainfall simulations from larger scale climate predictors by using downscaling methods are still a challenge in the hydrometeorology research.

From above analysis it can be concluded that the statistics in Table 2 can only describe parts of the statistical characteristics of rainfall data, but cannot reflect well the dynamics of the process of simulated rainfall as compared with that of observed rainfall. The NSC and RSR values from the rainfall simulation and runoff simulation are coherent in all scenarios, which indicate that the NSC and RSR may be used as key statistics to evaluate the rainfall simulation of downscaling methods for climate change impact studies, at least before an even better statistic is being defined. Future study needs to define and verify more useful statistics for evaluation of downscaling methods in order to determine the best downscaling method for providing most useful rainfall data as input to drive hydrological models.

6. Conclusions

This study focuses on the comparison and evaluation of the skills and competences of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. The following conclusions can be drawn:

(1) According to the evaluation of calibration and validation performance using observed rainfall and discharge data, the XAJ model and HBV model have similar performance in simulation of historical streamflow in the Hanjiang catchment. However, when applying rainfall downscaled from NCEP, CGCM3 and HadCM3 as inputs to both hydrological models, the accuracy of simulation by XAJ is slightly higher than that of HBV. It indicates that the performance of XAJ model is more superiority than HBV in responding to climate change impact on runoff in this region.
(2) By using the same scenario, downscaling technique and hydrological model, the results showed that CGCM3 is more suitable than HadCM3 to investigate the climate change impact on runoff in this region. It also indicates that if only single GCM was used to analyze the impact of climate change, the conclusions would be not reliable and robust. If there is no limit of GCMs' data acquirement, more GCMs and emission scenarios should be used in the study of climate change impact on hydrology.
(3) For the same GCM and scenario, the simulation results of runoff vary greatly by using rainfall provided by different statistical downscaling techniques as the input to hydrological models. SSVM performed better than SDSM in studying climate change impact on runoff in the Hanjiang basin. Therefore, it is recommended to use more than one statistical downscaling method to study the climate change impacts on runoff.
(4) Most statistics used in this study as well as in the literature for evaluation of the performance of downscaling methods show SDSM has better performance than SSVM in downscaling rainfall, with an exception of the NSC and RSR values. However, the runoff simulation efficiency as measured by NSC and RSR driven by SDSM rainfall is far lower than by SSVM. It can be concluded that NSC and RSR commonly used for evaluation of hydrological models can be used as key statistics of the assessment of statistical downscaling methods as well in assessing climate change impact on hydrology. This study also reveals that more useful statistics for evaluation of downscaling methods are to be defined and verified in order to determine the best downscaling method for providing most useful rainfall data as input to drive hydrological models.

Acknowledgements

The study was supported financially by the National Key Technologies R&D Program of China (2009BAC56B01) and the National Natural Science Fund of China (50809049). The second author was also supported by the Program of Introducing Talents of Discipline to Universities—the 111 Project of Hohai University. The authors greatly appreciate the Canadian Climate Change Scenarios Network (CCCSN) for providing the reanalysis products of the NCEP and HadCM3 outputs for the downscaling tool and would like to acknowledge the Data Access Integration (DAI, see http://quebec.ccsn.ca/DAI/) Team for providing CGCM3 data and technical support. Last but not least, the authors wish to thank the reviewers for their comments and suggestions which greatly improved the quality of the paper.

References

Ashagrie, A.G., de Laat, P.J.M., de Wit, M.J.M., Tu, M., Uhlenbrook, S., 2006. Detecting the influence of land use changes on discharges and floods in the Meuse River Basin – the predictive power of a ninety-year rainfall–runoff relation? Hydrol. Earth Syst. Sci. 10 (5), 691–701.
Bergstrom, S., 1975. The development of a snow routine for the HBV-2 model. Nordic Hydrol. 6 (2), 73–92.
Bergstrom, S., Carlsson, B., Gardelin, M., Lindstrom, G., Pettersson, A., Rummukainen, M., 2001. Climate change impacts on runoff in Sweden – assessments by global climate models, dynamical downscaling and hydrological modelling. Climate Res. 16 (2), 101–112.
Boe, J., Terray, L., Habets, F., Martin, E., 2007. Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. Int. J. Climatol. 27 (12), 1643–1655.
Chen, H., Guo, S.L., Xu, C.Y., Singh, V.P., 2007. Historical temporal trends of hydro-climatic variables and runoff response to climate variability and their relevance in water resource management in the Hanjiang basin. J. Hydrol. 344, 171–184.
Chen, H., Guo, J., Xiong, W., Guo, S.L., Xu, C.Y., 2010. Downscaling GCMs using the Smooth Support Vector Machine method to predict daily precipitation in the Hanjiang Basin. Adv. Atmos. Sci. 27 (2), 274–284.
Chiew, F.H.S., Kirono, D.G.C., Kent, D.M., Frost, A.J., Charles, S.P., Timbal, B., Nguyen, K.C., Fu, G., 2010. Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates. J. Hydrol. 387 (1–2), 10–23.
DAI CGCM3 Predictors, 2008. Sets of Predictor Variables Derived from CGCM3 T47 and NCEP/NCAR Reanalysis. Version 1.1, April 2008, Montreal, QC, Canada, pp. 1–16.

Dibike, Y.B., Coulibaly, P., 2005. Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. J. Hydrol. 307 (1–4), 145–163.
Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27 (12), 1547–1578.
Gleick, P.H., 1987. The development and testing of a water-balance model for climate impact assessment – modeling the Sacramento basin. Water Resour. Res. 23 (6), 1049–1061.
Guo, S.L., Wang, J.X., Xiong, L.H., Ying, A.W., Li, D.F., 2002. A macro-scale and semi-distributed monthly water balance model to predict climate change impacts in China. J. Hydrol. 268 (1–4), 1–15.
Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J. Hydrol. Eng. 4 (2), 135–143.
Jiang, T., Chen, Y.Q.D., Xu, C.Y.Y., Chen, X.H., Chen, X., Singh, V.P., 2007. Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. J. Hydrol. 336 (3–4), 316–333.
Jin, X.L., Xu, C.Y., Zhang, Q., Chen, Y.D., 2009. Regionalization study of a conceptual hydrological model in Dongjiang basin, south China. Quaternary Int. 208, 129–137.
Khan, M.S., Coulibaly, P., 2010. Assessing hydrologic impact of climate change with uncertainty estimates: bayesian neural network approach. J. Hydrometeorol. 11 (2), 482–495.
Lee, Y.J., Mangasarian, O.L., 2001. SSVM: a smooth support vector machine for classification. Comput. Optim. Appl. 20 (1), 5–22.
Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themessl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M., Thiele-Eich, I., 2010. Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev. Geophys. 48 (RG3003), 1–38.
Maraun, D., Osborn, T.J., Rust, H.W., 2011. The influence of synoptic airflow on UK daily precipitation extremes. Part I: observed spatio-temporal relationships. Clim. Dynam. 36 (1–2), 261–275.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.
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.
Nelder, J.A., Mead, R., 1965. A simplex-method for function minimization. Comput. J. 7 (4), 308–313.
Pinto, J.G., Neuhaus, C.P., Leckebusch, G.C., Reyers, M., Kerschgens, M., 2010. Estimation of wind storm impacts over Western Germany under future climate conditions using a statistical–dynamical downscaling approach. Tellus Ser. A – Dynam. Meteorol. Oceanogr. 62 (2), 188–201.
Pope, V.D., Gallani, M.L., Rowntree, P.R., Stratton, R.A., 2000. The impact of new physical parameterizations in the Hadley Centre climate model-HadCM3. Climate Dynam. 16, 123–146.
Prudhomme, C., Davies, H., 2009. Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate. Climatic Change 93 (1–2), 177–195.
Rosenbrock, H.H., 1960. An automatic method for finding the greatest or least value of a function. Comput. J. 3 (3), 175–184.
Schoof, J.T., Shin, D.W., Cocke, S., LaRow, T.E., Lim, Y.K., O'Brien, J.J., 2009. Dynamically and statistically downscaled seasonal temperature and precipitation hindcast ensembles for the southeastern USA. Int. J. Climatol. 29 (2), 243–257.
Segui, P.Q., Ribes, A., Martin, E., Habets, F., Boe, J., 2010. Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. J. Hydrol. 383 (1–2), 111–124.
Seibert, J., McDonnell, J.J., Woodsmith, R.D., 2010. Effects of wildfire on catchment runoff response: a modelling approach to detect changes in snow-dominated forested catchments. Hydrol. Res. 51 (5), 378–390.
Spendley, W., Hext, G.R., Himsworth, F.R., 1962. Sequential application of simplex designs in optimisation and evolutionary operation. Technometrics 4 (4), 441–461.
Vapnik, V.N., 1998. Statistical Learning Theory. Wiley, New York.
Wang, Q.J., 1991. The genetic algorithm and its application to calibrating conceptual rainfall–runoff models. Water Resour. Res. 27 (9), 2467–2471.
Wetterhall, F., Bardossy, A., Chen, D.L., Halldin, S., Xu, C.Y., 2006. Daily precipitation-downscaling techniques in three Chinese regions. Water Resour. Res. 42 (11).
Wilby, R.L., Hay, L.E., Leavesley, G.H., 1999. A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. J. Hydrol. 225 (1–2), 67–91.
Wilby, R.L., Dawson, C.W., Barrow, E.M., 2002. SDSM – a decision support tool for the assessment of regional climate change impacts. Environ. Model. Software 17 (2), 147–159.
Xu, C.Y., 1999. From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Prog. Phys. Geogr. 23 (2), 229–249.
Xu, C.-Y., 2000. Modelling the effects of climate change on water resources in central Sweden. Water Resour. Manage. 14, 177–189.
Xu, C.Y., Widen, E., Halldin, S., 2005. Modelling hydrological consequences of climate change – progress and challenges. Adv. Atmos. Sci. 22 (6), 789–797.
Yang, W., Andreasson, J., Graham, L.P., Olsson, J., Rosberg, J., Wetterhall, F., 2010. Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. Hydrol. Res. 41 (3–4), 211–229.
Yu, P.S., Wang, Y.C., 2009. Impact of climate change on hydrological processes over a basin scale in northern Taiwan. Hydrol. Process. 23 (25), 3556–3568.
Zhang, X.N., Lindstrom, G., 1996. A comparative study of a Swedish and a Chinese hydrological model. Water Resour. Bull. 32 (5), 985–994.
Zhao, R.J., 1992. The Xinanjiang model applied in China. J. Hydrol. 135 (1–4), 371–381.
Zhao, R.J., Zhang, Y.L., Fang, L.R., Liu, X.R., Zhang, Q.S., 1980. The Xinanjiang Model, Hydrological Forecasting Proceedings Oxford Symposium. IASH, pp. 351–356.