New Approaches to Hydrological Prediction in Data-sparse Regions (Proc. of Symposium HS.2 at the Joint IAHS & IAH Convention, Hyderabad, India, September 2009). IAHS Publ. 333, 2009.

Coupling VIC with GCM models to predict climate change impact in the Hanjiang basin, China

SHENGLIAN GUO, JING GUO, JUN ZHANG & HUA CHEN

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China
slguo@whu.edu.cn

Abstract A Smooth Support Vector Machine (SSVM) is proposed for statistical downscaling of daily precipitation and temperature from GCM output. The Variable Infiltration Capacity (VIC) distributed hydrological model with a 9 × 9 km^2 grid resolution is established and calibrated in the Hanjiang basin of China. Validation results show that SSVM can approximate observed precipitation and temperature data reasonably well, and the VIC model can simulate the runoff hydrograph with high model efficiency and low relative error. By applying the SSVM model, the trends of precipitation and temperature projected from CGCM2 under the A2 and B2 scenarios will decrease in the 2020s, and increase in the 2080s. However, in the 2050s, the precipitation will decrease under the A2 scenario and there will be no significant changes under the B2 scenario, but the temperature will be not obviously change under either scenario. Under both scenarios, the impact analysis of runoff made with the downscaled precipitation and temperature time series as input to the VIC distributed model, resulted in a decreasing trend for the 2020s and 2050s, and an overall increasing trend for the 2080s.

Key words climate change; statistical downscaling; GCM; SSVM; VIC model; impact study

INTRODUCTION

Recently, there has been a growth in the scientific evidence that global climate has changed, and that the changing will continue. Such changes in climate will also have significant impact on local and regional hydrological regimes, which will in turn affect ecological, social and economic systems. There is an urgent need to improve our understanding of the global climate system to assess the possible impact of a climate change on hydrological processes. General circulation models (GCMs), which describe atmospheric processes by mathematical equations, are one of the most important tools for studying the impact of climate change. Statistical downscaling aims to derive empirical relationships that transform large-scale features of the GCM (predictors) to regional-scale variables (predictands), such as precipitation and temperature (Tripathi *et al.*, 2006).

Distributed hydrological models have the ability to produce simulations of spatial patterns of hydrological response due to soil, vegetation, land use, precipitation, evaporation and runoff; furthermore, the gridded structure of the models can be easily coupled with GCMs. Therefore, the statistical downscaling of GCM outputs as input to distributed hydrological models to assess the effects of climate change has been recognized by many authors (Xu, 1999; Wilby *et al.*, 2006; Manoj *et al.*, 2006; Ghosh & Mujumdar, 2008), and will become the best approach for climate change impact studies.

Here, a smooth support vector machine (SSVM) is proposed for statistical downscaling of daily precipitation and temperature from GCM output. The variable infiltration capacity (VIC) distributed hydrological model is established and calibrated in the Hanjiang basin of China. Then the downscaled data from CGCM2 outputs was used as input to the VIC model to assess the impact of climate change on future precipitation, temperature and runoff in the Hanjiang basin.

STUDY AREA

The Hanjiang basin is the largest tributary of the Yangtze River; it passes through the provinces of Shannxi and Hubei of China, and merges into the Yangtze River at Wuhan city. The river's length is 1570 km and the basin area is 159 000 km^2. The basin has a sub-tropical monsoon climate and has, as a result, dramatic diversity in its water resources. Annual precipitation varies from 700 to 1100 mm, of which 70–80% of the total amount occurs in the wet season from May to October.

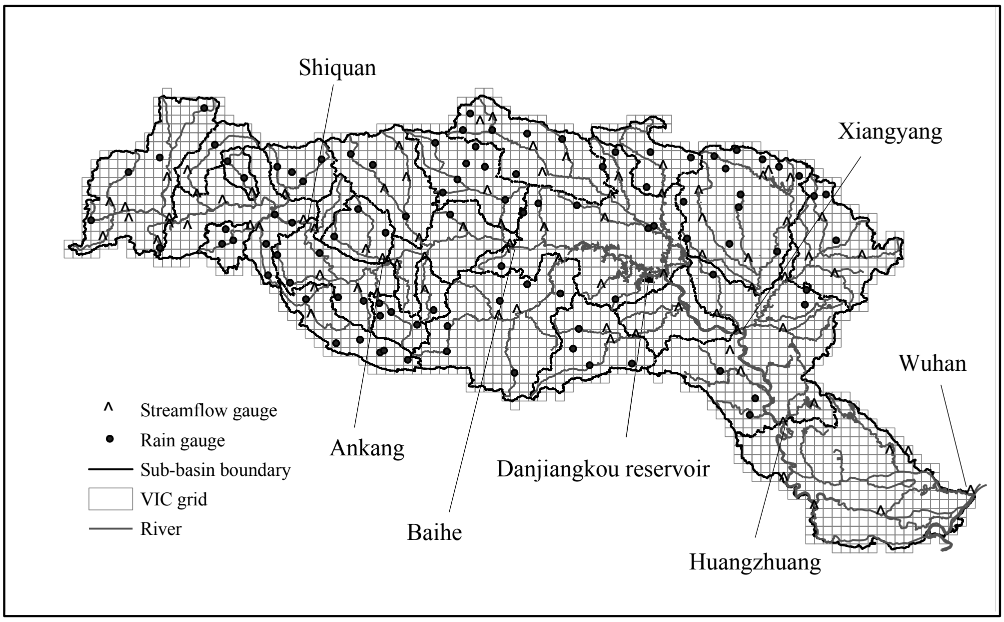

Fig. 1 Characteristics of the Hanjiang basin for the VIC model.

The Hanjiang basin plays critical roles in the flood control and water supply in China. The Danjiangkou Reservoir, located in the middle reach of the Hanjiang basin (Fig. 1), is the source of water for the middle route of the South–North Water Diversion Project, and the Jianghan plain in the lower basin is one of the most important bases for commodity grain production. The available water resources in Hanjiang basin and the impact of water diversion have been discussed by many authors (Guo *et al*., 2002; Chen *et al.*, 2007).

SSVM STATISTICAL DOWNSCALING MODEL

Support vector machine (SVM) is a new machine study method based on statistical learning theory and stresses for studying statistical learning rules under small sample conditions (Vapnik, 1998). SVM solves many practical problems, such as small-sample, non-linear, high dimension number and global minimum points, by using a structural risk minimization principle. Recently, SVM has been widely applied in the fields of hydrological classification and regression analysis (Tripathi *et al*. 2006; Chen & Yu, 2007). However, it has some drawbacks in dealing with the large-sample data, such as slow training speed, low implementation efficiency and inadaptability to noise. To overcome the drawbacks of the SVM for large-sample data, Lee *et al*. (2005) proposed a new smoothing strategy for solving regression of the large-scale training data, called smooth support vector machine (SSVM), which has been verified as more efficient than the SVM algorithm mentioned above. The inequality constraint problem of SVM is replaced by an unconstrained problem and the SSVM has a unique global optimal solution. The detailed introduction of SVM and SSVM algorithms has been described by Lee *et al*. (2005), Tripathi *et al*. (2006) and Chen & Yu (2007). A tuning procedure which can automatically optimize parameters (Lee *et al*., 2005) is applied in this study to estimate the parameters of the SSVM.

Predictands and predictors

It is one of the most important steps in a downscaling exercise to select appropriate predictors, or characteristics from GCMs. The mean sea level pressure (MSLP), surface air temperature (2 m), 500-hPa geopotential height (GH) and specific humidity (SH), and 850-hPa GH and SH were selected as the predictors for precipitation; 850-hPa temperature (TEM) and MSLP were considered as the predictors for temperature (Wilby *et al.*, 1999). The observed daily data of large-scale predictor variables representing the current climate condition (1960–2000) were derived from the re-analysis data set of the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR). The simulated daily data of the predictor variables by GCMs are available from the second generation coupled global climate model of the Canadian Centre (CGCM2). The geographical extent, 102.5°E–115°E, 27.5°N–37.5°N was chosen to include all areas with noticeable influence in the circulation patterns that govern weather in the Hanjiang basin. The inverse distance weighting method (IDW), which is based on the assumption that the interpolating surface should be influenced most by nearby points and less by more distant points, is used to spatially interpolate CGCM2 grids (3.75° lat. × 3.75° long.) into the NCEP/NCAR grids.

Verification of SSVM model

In order to evaluate the performance of SSVM, the data set was divided into two parts, the calibrating set, 1961–1990, and validating set, 1991–2000. The calibrating data is used to establish the regression function that is learnt from calibrating data; the validating set is used to assess the prediction ability of the resulting regression function.

Prior to downscaling of the predictors, the NCEP/NCAR reanalysis data and GCM data are standardized to reduce systematic biases in the mean and variances of GCM outputs. Principal component analysis (PCA) has been widely used to reduce dimensions and compress data while keeping most of the information content of the original data set (Tripathi *et al.*, 2006). Table 1 show that the first eight PCs (principal components) for precipitation and the first two PCs for temperature have represented 90% and 91% of the information content of the original predictors, respectively. Therefore, the use of these PCs as input to SSVM, enables the dimension of the data set to be reduced without decreasing the performance of the statistical downscaling model.

The differences of the mean and standard deviation between the observed and simulated daily precipitation are considered as the most important criteria to evaluate the downscaling model. It was observed that the relative bias (RB) between mean observed and downscaled values are all within 6%, the standard deviation values downscaled by SSVM are much lower than those of observations which means that high rainfall and concentrated periods of precipitation cannot be captured well by the downscaling model. It is also shown that SSVM performs well, with a good agreement between the observed and downscaled outputs, for the mean values of daily mean temperature, daily maximum temperature and daily minimum temperature. Their relative biases are all within 3% except for daily minimum temperature in the winter season, while the standard deviation of downscaled temperature is smaller than the observed, especially in summer and winter.

Table 1 The percent variance and cumulative values of the PCs by using the PCA method to process the NCEP predictors daily data set.

Predictand	PCs No.	1	2	3	4	5	6	7	8
Precipitation	Eigenvalue	77.34	25.58	10.33	6.16	3.5	3.12	1.76	1.5
	Percent variance	54	18	7	4	3	2	1	1
	Cumulative percent	54	72	79	83	86	88	89	90
Temperature	Eigenvalue	41.38	2.62						
	Percent variance	86	5						
	Cumulative percent	86	91						

VIC DISTRIBUTED MODEL

The Variable Infiltration Capacity (VIC) distributed hydrological model is a macro-scale hydrological model based on a soil–vegetation–atmosphere transfer scheme, which is designed to describe the land surface in numerical weather prediction and climate, and describe the variation and transfer of water and energy (Liang *et al.*, 1994).

Establishing the VIC model for the Hanjiang basin

The VIC model has one kind of bare soil and different vegetation types in each grid cell. It includes both the saturation and infiltration excess runoff processes in a grid cell with consideration of the sub grid-scale soil heterogeneity, and the frozen soil processes for cold climate conditions. Three types of evaporation: evaporation from wet canopy, evapotranspiration from dry canopy and evaporation are considered. The one-dimensional Richards equation is used to describe the vertical soil moisture movement, and the moisture transfer between soil layers obeys the Darcy law. The ARNO method is used to describe baseflow which takes place only in the lowest layer. The routing model represented by the unit hydrograph method for overland flow and the linear Saint-Venant method for channel flow, allow runoff to be predicted (Liang *et al.*, 1994).

The hydrological information, DEM, forcing, soil and vegetation data, etc., is required for VIC model calibration. DEM data of 0.009 degree (around 1×1 km^2 cell size) spatial resolution for the Hanjiang basin were derived and used to delineate the sub-basin boundary and stream network. Figure 1 presents the information of the Hanjiang basin, including hydrological and rain stations, sub-basin boundary and 9×9 km^2 grid. Vegetation type data were taken from the global land cover classification generated by the University of Maryland, USA, with a 1-km pixel resolution. Vegetation parameters were based on the vegetation from the Land Data Assimilation System. The soil parameters were derived from the soil classification information of the global 5-min data provided by the National Atmospheric and Oceanic Administration, USA.

Calibration and validation results

The VIC model has six parameters that need to be calibrated. Daily hydrological and meteorological data from 1980–1986 and 1987–1990 are used for calibration and verification, respectively. The Nash-Sutcliffe efficiency (R^2) and the relative error (RE) of the volumetric fit criteria are used to justify the performance of the model. The main parameterization procedures are described as follows: the model parameters were first calibrated for gauged sub-basins, and then the calibrated parameters were used as the initial values for the corresponding grids; finally, hydrological control stations were selected in the main streams of the Hanjiang basin to test and optimize the grid parameters through a trial-and-error method. Table 2 lists the simulated results of six control stations in the main stream of the Hanjiang basin during the calibration and validation periods. The mean value of R^2 is 90.4% in the calibration period and 81.98% in the validation period. The mean RE values are 2.88% and 4.32% during the calibration and validation period, respectively. These results show that the VIC model can simulate daily runoff hydrograph well in the Hanjiang basin.

Table 2 The calibration and validation results of six hydrological stations in the Hanjiang River.

Hydrological Station	Area (km^2)	Calibration		Validation	
		R^2 (%)	RE (%)	R^2 (%)	RE (%)
Shiquan	24 629	89.27	11.00	77.15	16.00
Ankang	35 600	87.94	12.00	77.75	12.00
Baihe	59 115	89.27	4.00	80.70	3.00
Danjiangkou	95 220	88.43	–3.00	78.22	–8.00
Xiangyang	103 261	98.69	–1.08	97.33	3.24
Huangzhuang	142 056	93.10	–5.67	80.71	–0.33
Mean		91.12	2.88	81.98	4.32

HYDROLOGICAL IMPACT OF CLIMATE CHANGE

The large-scale predictor variables derived from the A2 and B2 scenarios of CGCM2 are used as input to the validated SSVM model to downscale the future climate change scenarios. Daily precipitation, daily mean temperature, daily maximum temperature and daily minimum temperature are downscaled by SSVM for four periods, namely: the current (1961–2000), 2020s (2011–2040), 2050s (2041–2070) and 2080s (2071–2100). The monthly mean statistics of downscaling results for precipitation and temperature for different periods are plotted in Fig. 2, showing a decreasing trend for the 2020s and the 2050s, and an increasing trend for the 2080s under scenario A2.

The deviations of simulated precipitation and temperature from the current simulated annual mean values, for the different future periods, are listed in Table 3. Between the current period and the 2080s, there are on average increases in precipitation of about 18.04% under A2 scenario and about 10.17% under B2 scenario. In the same period, there are average increases in daily maximum temperature and daily minimum temperature of about 1.86°C and 1.28°C under A2 scenario and about 1.41°C and 0.93°C under B2 scenario. This implies a corresponding increase in daily mean temperature by about 1.63°C and 0.80°C under the A2 and B2 scenarios, respectively.

The inverse distance weighting method is used to interpolate the downscaled hydrological and meteorological data. These data were input to VIC to simulate the runoff corresponding to future climate change scenarios. Table 3 lists the relative changes of mean annual runoff in the Hanjiang basin. Under A2 scenario, the mean annual runoff changes by about –30.21%, –14.38% and 31.04%, respectively, for the 2020s, 2050s and 2080s; under the B2 scenario, the changes are about –17.04%, –3.75% and 15.52%, respectively. Figure 3 shows the spatial distribution of runoff based on the 9 × 9 km^2 grid in the Hanjiang basin for the current, 2020s, 2050s and 2080s periods. It can be seen that the spatial distribution of runoff simulated by the VIC model is, in general, consistent under both climate change scenarios.

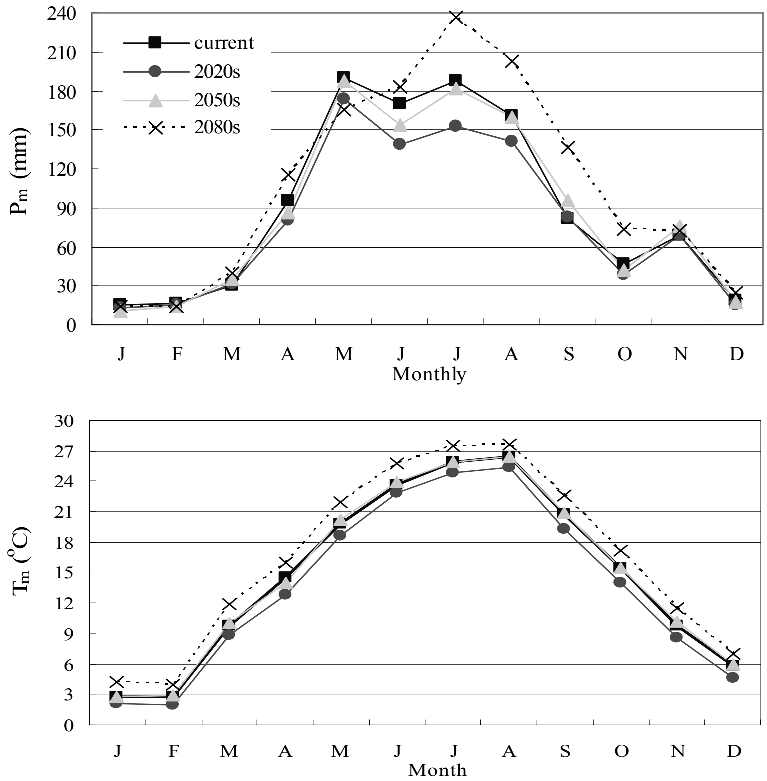

Fig. 2 General trend in mean monthly precipitation (P_m) and temperature (T_m) for scenario A2 downscaled by SSVM in the Hanjiang basin.

SUMMARY AND CONCLUSION

The objective of this study was downscaling of large-scale atmospheric variables from GCM outputs to climate variables at regional and local scale in order to investigate the hydrological

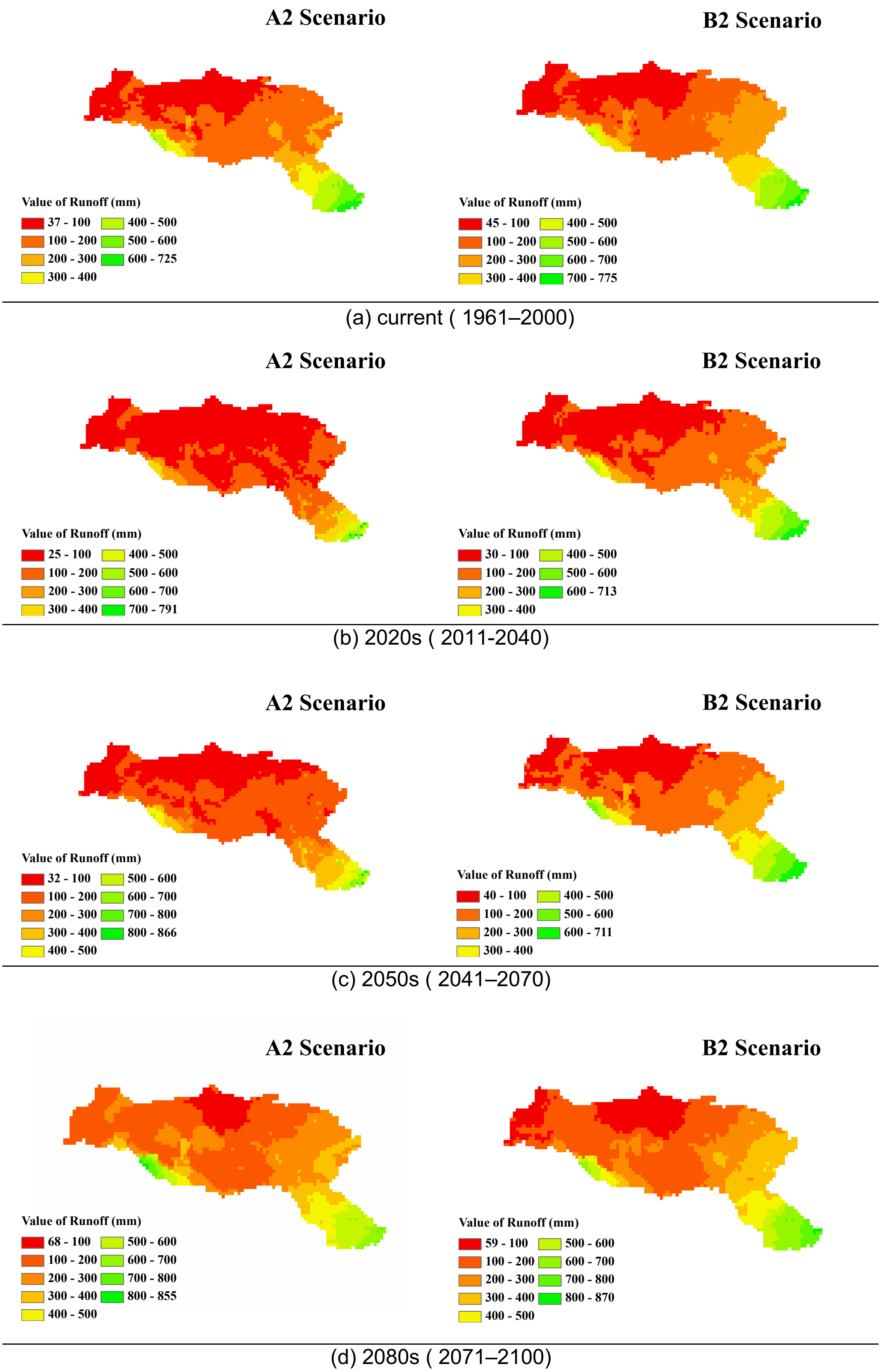

Fig. 3 Spatial distribution of runoff simulated by the VIC distributed model in the Hanjiang basin.

Table 3 Simulated changes in annual mean precipitation, temperature and runoff under the A2 and B2 scenarios of CGCM2.

	A2 Scenario				B2 Scenario			
	Current	2020s	2050s	2080s	Current	2020s	2050s	2080s
Precipitation (mm)	1082.37	952.87	1059.68	1277.67	1117.84	1021.55	1117.09	1231.53
Change (%)		–11.96	–2.10	18.04		–8.61	–0.07	10.17
Mean temperature (°C)	14.80	13.68	14.89	16.43	14.79	14.24	14.91	15.59
Change (%)		–1.20	0.09	1.63		–0.55	0.12	0.80
Runoff ($10^8 m^3$)	571.62	398.96	489.45	749.05	621.20	515.32	597.92	717.63
Changes (%)		–30.21	–14.38	31.04		–17.04	–3.75	15.52

impact of future climate change. A smooth support vector machine (SSVM) was proposed for statistical downscaling of daily precipitation and temperature. The SSVM model approximates the observed climate data reasonably well, except that it has underestimated the variance of precipitation and temperature. The VIC distributed model was used with a 9 × 9 km^2 grid and the results show that it can simulate the runoff hydrograph well in the Hanjiang basin.

For downscaling precipitation and temperature, the results corresponding to both climate change scenarios show that the SSVM model has estimated a decreasing trend in the 2020s, an increase trend in the 2080s, and a mixed trend in the 2050s. The impact analysis of runoff from the Hanjiang basin shows a similar trend in future under both climate change scenarios. Moreover, the VIC distributed model also gives the spatial distribution of changes of runoff and it can be seen that, throughout almost the entire basin in the 2020s and most regions of the basin in the 2050s, runoff will be decreasing under both climate change scenarios, while in the 2080s runoff will be increasing significantly across the whole basin.

Acknowledgements The study is financially supported by the National Natural Science Fund of China (50679063, 50809049).

REFERENCES

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.* **334**, 171–184.

Chen, S. T. & Yu, P. S. (2007) Pruning of support vector networks on flood forecasting. *J. Hydrol.* **347**(1-2), 67–78.

Ghosh, S. & Mujumdar, P. P. (2008) Statistical downscaling of GCM simulations to streamflow using relevance vector machine. *Adv. Water Resour*. **31**(1), 132–146.

Guo, S. L., Wang, J. X., Xiong, L. H. & Ying, A. W. (2002) A macro-scale and semi-distributed monthly water balance model to predict climate change impacts in China. *J. Hydrol.* **268**, 1–15.

Lee, Y., Hsieh, W. & Huang, C. (2005) ϵ-SSVR:A smooth support vector machine for insensitive regression. *IEEE Trans Knowledge and Data Engineering* **17**, 678–685.

Liang, X., Lettenmaier, D. P., Wood, E. F. &, Burges, S. J.(1994) A simple hydrologically based model of land Surface water and energy fluxes for GCMs. *J. Geophys. Res.* **99**(14), 415–428.

Manoj, J., Arnold, J. G., Gassman, P. W., Giorgi, F. & Gu, R. R. (2006) Climate change sensitivity assessment on Upper Mississippi River Basin streamflows using SWAT. *J. Am. Water Resour. Assoc*. **42**(4), 997–1015.

Tripathi, S., Srinivas, V. V. & Nanjundiah, R. S.(2006) Downscaling of precipitation for climate change scenarios: A support vector machine approach. *J. Hydrol*. **330**(3-4), 621–640.

Vapnik, V. N. (1998) *Statistical Learning Theory.* Wiley, New York, USA.

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., Whitehead, P. G., Wade, A. J., Butterfield, D., Davis, R. J. & Watts, G. (2006) Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK. *J. Hydrol*. **330**, 204–220.

Xu, C. Y. (1999) From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progr. Phys. Geogr.* **23**, 229–249.