



# Runoff of arid and semi-arid regions simulated and projected by CLM-DTVGM and its multi-scale fluctuations as revealed by EEMD analysis

NING Like<sup>1,2</sup>, XIA Jun<sup>3,1\*</sup>, ZHAN Chesheng<sup>1</sup>, ZHANG Yongyong<sup>1</sup>

<sup>1</sup> Key Laboratory of Water Cycle & Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China;

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China;

<sup>3</sup> State Key Laboratory of Water Resources & Hydropower Engineering Science, Wuhan University, Wuhan 430000, China

**Abstract:** Runoff is a major component of the water cycle, and its multi-scale fluctuations are important to water resources management across arid and semi-arid regions. This paper coupled the Distributed Time Variant Gain Model (DTVGM) into the Community Land Model (CLM 3.5), replacing the TOPMODEL-based method to simulate runoff in the arid and semi-arid regions of China. The coupled model was calibrated at five gauging stations for the period 1980–2005 and validated for the period 2006–2010. Then, future runoff (2010–2100) was simulated for different Representative Concentration Pathways (RCP) emission scenarios. After that, the spatial distributions of the future runoff for these scenarios were discussed, and the multi-scale fluctuation characteristics of the future annual runoff for the RCP scenarios were explored using the Ensemble Empirical Mode Decomposition (EEMD) analysis method. Finally, the decadal variabilities of the future annual runoff for the entire study area and the five catchments in it were investigated. The results showed that the future annual runoff had slowly decreasing trends for scenarios RCP 2.6 and RCP 8.5 during the period 2010–2100, whereas it had a non-monotonic trend for the RCP 4.5 scenario, with a slow increase after the 2050s. Additionally, the future annual runoff clearly varied over a decadal time scale, indicating that it had clear divisions between dry and wet periods. The longest dry period was approximately 15 years (2040–2055) for the RCP 2.6 scenario and 25 years (2045–2070) for the RCP 4.5 scenario. However, the RCP 8.5 scenario was predicted to have a long dry period starting from 2045. Under these scenarios, the water resources situation of the study area will be extremely severe. Therefore, adaptive water management measures addressing climate change should be adopted to proactively confront the risks of water resources.

**Keywords:** Community Land Model (CLM); Distributed Time Variant Gain Model (DTVGM); Ensemble Empirical Mode Decomposition (EEMD); decadal variability; arid and semi-arid regions

**Citation:** NING Like, XIA Jun, ZHAN Chesheng, ZHANG Yongyong. 2016. Runoff of arid and semi-arid regions simulated and projected by CLM-DTVGM and its multi-scale fluctuations as revealed by EEMD analysis. *Journal of Arid Land*, 8(4): 506–520. doi: 10.1007/s40333-016-0126-4

Water is one of the most important resources in the arid and semi-arid regions of China, and its value to social-economic department and environmental health is becoming increasingly vital. In addition, multi-scale fluctuations in water resources, such as the interannual and interdecadal changes, are posing serious threats and challenges to water resources management across the arid and semi-arid regions of China. Thus, assessment of water resources on a large and long-term

\*Corresponding author: XIA Jun (E-mail: xiajun6666@gmail.com)

Received 2015-08-14; revised 2015-11-27; accepted 2016-01-06

© Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Science Press and Springer-Verlag Berlin Heidelberg 2016

scale and its characteristics of multi-scale fluctuation are not only essential for sustaining human life and the environment, but also beneficial to water management.

Runoff is one of the most important components in the large scale terrestrial hydrologic cycle (Choi and Liang, 2010). The terrestrial hydrological cycle is linked to the energy cycle and to biogeochemical processes through the Earth's climate system (Oleson et al., 2008). The cycle is composed of interactions between precipitation and the vegetation/soil system, which determines the partitioning of water into various storage reservoirs and the subsequent release of water vapor into the atmosphere. The large-scale hydrological models describing the terrestrial water balance at continental and global scales have increasingly been used for modelling and climate impact assessments (Gudmundsson et al., 2012). The Community Land Model (CLM) is a state-of-the-art soil–vegetation–atmosphere transfer model (Dai et al., 2001). However, the CLM adopted a simple TOPMODEL-based model for surface and subsurface runoff (Niu et al., 2005). Although TOPMODEL-based models can describe the hydro-physical processes, they only consider the spatial variability of topography, and they assume that the hydrological information (such as precipitation) is spatially homogeneous. The Distributed Time Variant Gain Model (DTVGM) is a type of distributed conceptual hydrological model. It simulates the movement of water in the soil–vegetation–atmosphere system, describes the relationships between the cellular grids of the digital elevation model (DEM) in the horizontal direction, and performs mathematical calculations of surface water and ground water on the cellular. Furthermore, DTVGM has few parameters, which can be estimated in terms of the system identification approach to reduce uncertainty (Xia, 2002; Xia et al., 2005).

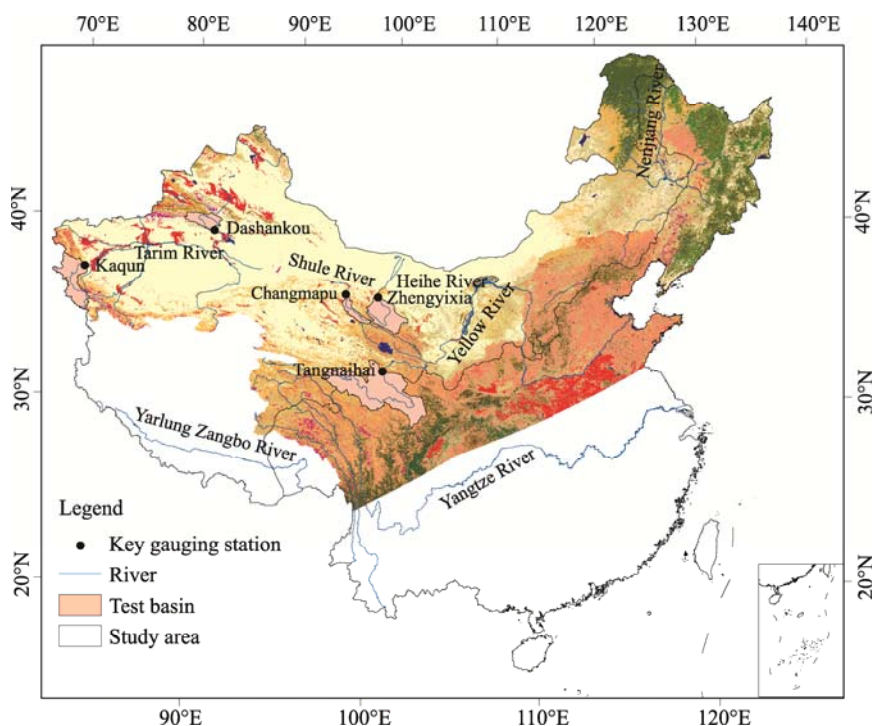
The identification to multi-scale fluctuations is not only the basis for understanding long-term variations, but also the foundation for revealing periodicities and other characteristics of hydrological processes (Kallache et al., 2005). To identify different temporal-scale hydrological components, many methods have been adopted in hydrological time series analysis. Among those methods, the Ensemble Empirical Mode Decomposition (EEMD) method, which can reveal the non-stationary and nonlinear characteristics of a series under multi-temporal scales (Wu and Huang, 2009), has been frequently employed. The EEMD method represents an improvement of the empirical mode decomposition (EMD) method. It not only keeps the strengths of EMD, but also overcomes its drawbacks (Wu and Huang, 2009). Compared with other decomposition methods, EEMD is an adaptive time-frequency data analysis method, and it has proven to be quite versatile for extracting signals from data generated in noisy non-linear and non-stationary processes (Huang and Wu, 2008).

Therefore, the present study coupled DTVGM into CLM, replacing the TOPMODEL-based method to simulate runoff in the arid and semi-arid regions of China. The coupled model was calibrated at five gauging stations for the period 1980–2005, and was validated for the period 2006–2010 for performance assessment. Then, the spatial distributions of the future runoff for different Representative Concentration Pathways (RCP) emission scenarios were explored. Additionally, the multi-scale fluctuation characteristics of the simulated annual runoff for these RCP scenarios were discussed using the EEMD analysis method. Finally, the decadal variability of the future annual runoff was investigated.

## 1 Study area and data description

### 1.1 Study area

This study took the 500-mm isohyet as the boundary between semi-arid and humid regions (Yang et al., 2002). The arid and semi-arid area lies to the left of the Greater Khingan Mountains–Yinshan Mountains–Bayankala Mountains, and covers almost half of China, with an area of  $393.87 \times 10^4 \text{ km}^2$ . This area includes Xinjiang, Qinghai, Inner Mongolia, Ningxia and other administrative regions (Fig. 1). Five predominant land cover types are bare areas (45.99% of the study area), mosaic vegetation/croplands (19.01%), rainfed croplands (10.17%), closed to open grassland (5.76%) and sparse vegetation (4.66%). Additionally, many inland rivers, such as the Tarim, Kaidu and Heihe rivers and the sources of the Yellow and Yangtze rivers, are located in the study area.



**Fig. 1** The study area and locations of the test basins and key gauging stations

## 1.2 Data description

This study took the Princeton's Global Meteorological Forcing Dataset provided by the Department of Civil and Environmental Engineering, Princeton University (Sheffield et al., 2006) to calibrate and validate the model during 1980–2010. This dataset provides near-surface meteorological data for driving land surface models and other terrestrial modeling systems. It has been widely used to force the CLM in China (Li et al., 2011; Wang et al., 2011; Lai et al., 2014). Future forcing climate data were obtained from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project Phase 5 multi-model dataset (CMIP5 dataset). Within this dataset, 21 CMIP5 models were used to calculate the average monthly precipitation and temperature at the resolution of  $1^\circ \times 1^\circ$ , which were collected and provided by National Climate Center of China (Shangguan et al., 2013). Scenarios RCP2.6, RCP4.5, and RCP8.5 were included in the CMIP5 for climate projection (Moss et al., 2010). The land cover data were extracted from GLC2000 China produced by the Institute of Remote Sensing Applications, Chinese Academy of Sciences and the U.S. Geological Study-for the Global Land Cover 2000 Initiative (Xu et al., 2005). The soil texture dataset was extracted from soil texture maps provided by the Institute of Soil Science of the Chinese Academy of Sciences, which accurately reflects the geographical distribution of the soil textures in China. And the data of soil properties were extracted from 1:1,000,000 soil category and distribution dataset, which was compiled based on the Second National Soil Survey of China by the National Soil Survey Office.

The five keys gauging stations are located at the outlets of five test basins (Fig. 1). The observed hydrological data at the gauging stations were provided by the Bureau of Hydrology, Ministry of Water Resources of China. The data of monthly discharge cover 1980–2010, except at the Changmapu Station where data were missing from 2000 to 2005.

## 2 Methodology

### 2.1 Coupling of CLM and DTGVM

In this study, the DTGVM took the place of the TOPMODEL-based method in the Community

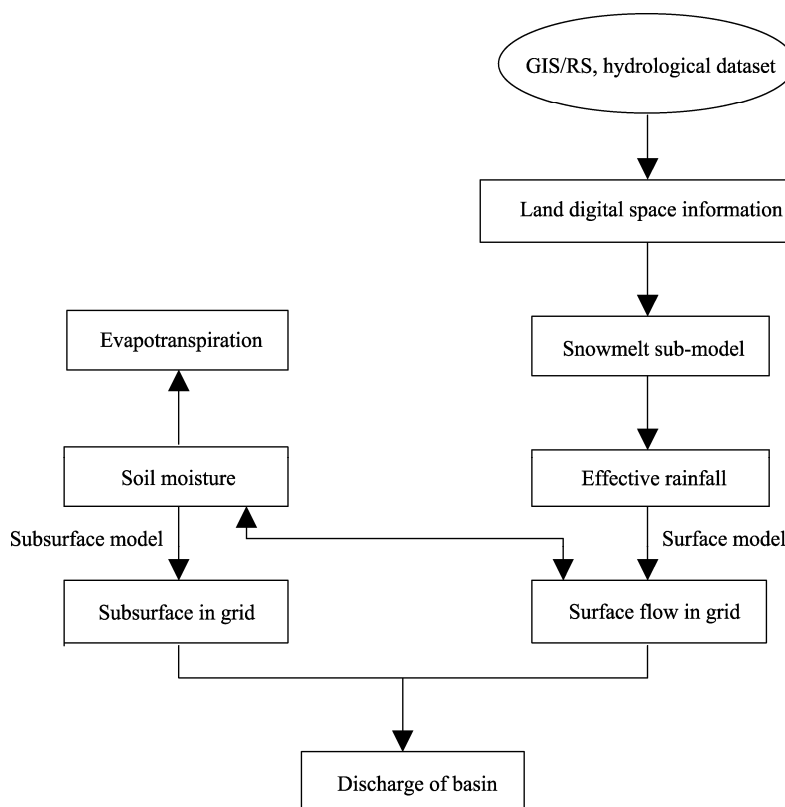
Land Model. Based on the DTGVM, the hydrological cycle of CLM 3.5 was improved, and was adapted to the major geographical and hydrological characteristics of the semi-arid regions of China.

### 2.1.1 CLM 3.5

The CLM 3.5 is the land-surface component of the coupled climate Community Earth System Model (CESM) (Oleson et al., 2004). It is a community-developed global climate system model for investigating interannual and interdecadal climate variabilities, and for projecting future climate change (Collins et al., 2006). CLM 3.5 is designed to simulate the exchange processes of water, energy, and momentum between soil, vegetation and the atmosphere. The land surface is schematized as nested grids representing different land units (glacier, lake, wetland, urban, and vegetation). The vegetated surfaces are represented as a composition of up to 15 plant functional types (PFTs) plus bare soil. All of the PFTs share the same soil column, which was modelled by means of 10 hydraulically active layers vertically distributed according to an exponential law (Cammalleri et al., 2015). Hydrological processes simulated by CLM 3.5 include interception, canopy drip, infiltration, evaporation, surface runoff, plant uptake, subsurface drainage and groundwater storage (Oleson et al., 2004).

### 2.1.2 Distribution Time-Variant Gain Model

DTVGM combines a runoff generation process with a flow routing process, and carries out the hydrological simulation based on grid elements and the stream network (Xia, 2002; Xia et al., 2005). It can describe time-space variation of rainfall based on the DEM grid and spatial digital information. In this model, runoff generation occurs at each grid element, with two or three layers in the vertical direction. For example, in the two-layer model which is divided by the ground surface, the upper is the surface runoff generation layer, and the lower is the subsurface runoff generation layer. The scheme of DTVGM as applied here is shown in Fig. 2.



**Fig. 2** Scheme of the Distributed Time Variant Gain Model (DTVGM) (Xia et al., 2005)

DTVGM includes three major components as follows: (1) Specialization of land surface and hydrological information: watershed features and hydro-meteorological variants can be obtained by spatial interpolation or data assimilation technologies. (2) A runoff generation process on grid elements: surface runoff is calculated by:

$$R_s = g_1 \times S^{g_2} \times P. \quad (1)$$

Where  $R_s$  is surface runoff (mm),  $S$  is soil moisture (mm),  $P$  is precipitation (mm), and  $g_1$  and  $g_2$  are undetermined coefficients. By introducing a non-dimensional variant  $S/W$  instead of  $S$ , the surface runoff of the  $i^{\text{th}}$  grid at time  $t$  can be expressed as:

$$R_{s_{t,i}} = g_1 \times (S_{t,i}/W_i)^{g_2} \times P_{t,i}. \quad (2)$$

Where  $R_{s_{t,i}}$  is surface runoff (mm),  $S_{t,i}$  is soil moisture (mm),  $W_i$  is saturated soil moisture (mm), and  $P_{t,i}$  is precipitation (mm). As for the subsurface runoff generation model, the subsurface runoff can be calculated by coupling the water balance equation to the dynamic storage-outflow function. The subsurface runoff of the  $i^{\text{th}}$  grid at time  $t$  ( $R_{g_{t,i}}$ ) can be calculated as:

$$R_{g_{t,i}} = k_g \times (S_{t+1,i} + S_{t,i})/2. \quad (3)$$

Where  $S_t$  and  $S_{t+1}$  are the soil moistures at time  $t$  and time  $t+1$  respectively, and  $k_g$  is the subsurface runoff generation coefficient. The total runoff on the  $i^{\text{th}}$  grid ( $R_{t,i}$ ) is the sum of surface runoff and subsurface runoff, i.e.  $R_{t,i} = R_{s_{t,i}} + R_{g_{t,i}}$ . (3) A flow routing process based on ranked grids: the flow routing model includes two aspects of the flow route and the routing method. In DTVGM, the flow route can be described as ranked grids (or concentration belts) according to the flow directions of grid elements. The outlet grids can be defined as the first rank. The grids from which water flows into the first ranked grids are defined as the second rank, and the rest may be deduced by analogy. Therefore, the most upstream grids are defined as the highest rank. Flow routing is undertaken from those grids having higher rank to the grids having lower rank. Kinematic wave routing (Thompson, 1999), a more physically-based modeling approach, was adopted as the routing method in DTVGM.

## 2.2 Calibration and validation

An integration method based on the PSUADE platform and SCE-UA method was applied to estimate and optimize the parameters of hydrological models (Song et al., 2012). In this study, the DTVGM model was applied to generate the 3,000 runs of parameters on the PSUADE platform. Efficiency criteria are defined as mathematical measures to determine how well a model simulation fits the observations (Beven, 2011). To assess the performance of the coupled model, the coefficient of determination ( $R^2$ ), Nash-Sutcliffe efficiency (NSE), and percent bias, which are frequently used in hydrologic modeling studies, were adopted.

The coefficient of determination  $R^2$  is defined as the squared value of the coefficient of correlation. It is calculated as:

$$R^2 = \left( \frac{\sum_{i=1}^n (Q_i^{\text{sim}} - \overline{Q_i^{\text{sim}}})(Q_i^{\text{obs}} - \overline{Q_i^{\text{obs}}})}{\sqrt{\sum_{i=1}^n (Q_i^{\text{sim}} - \overline{Q_i^{\text{sim}}})^2 \sum_{i=1}^n (Q_i^{\text{obs}} - \overline{Q_i^{\text{obs}}})^2}} \right)^2. \quad (4)$$

Where obs denotes observed values and sim denotes simulated values. The range of  $R^2$  lies between 0 and 1, and it describes how much of the observed dispersion is explained by the prediction. A value of zero means no correlation at all, whereas a value of 1 means that the dispersion of the prediction is equal to that of the observation.

The efficiency NSE proposed by Nash and Sutcliffe (1970) is defined as one minus the sum of the absolute squared differences between the predicted and observed values normalized to the variance of the observed values during the period under investigation. It is calculated as:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (Q_i^{\text{sim}} - \overline{Q_i^{\text{sim}}})^2}{\sum_{i=1}^n (Q_i^{\text{obs}} - \overline{Q_i^{\text{obs}}})^2}. \quad (5)$$

Nash-Sutcliffe efficiencies can range from  $-\infty$  to 1. An efficiency of 1 (NSE=1) denotes a perfect match of the modelled discharge to the observed data. An efficiency of 0 (NSE=0) indicates that the model predictions match the mean of the observed data, whereas an efficiency of less than zero (NSE<0) signifies that the observed mean is a more effective predictor than the model. Essentially, the closeness of the model efficiency to 1 determines the accuracy of the model.

Percent bias (B) measures if the average tendency of the simulated data is larger or smaller than their observed counterparts (Gupta et al., 1999). The optimal value of B is 0.0, indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias (Gupta et al., 1999). The B value, which is expressed as a percentage, is calculated as:

$$B = \frac{\sum_{i=1}^n (Q_i^{\text{sim}} - Q_i^{\text{obs}})}{\sum_{i=1}^n Q_i^{\text{obs}}} \times 100\%. \quad (6)$$

### 2.3 Ensemble Empirical Mode Decomposition (EEMD)

The Ensemble Empirical Mode Decomposition method (Huang et al., 1998; Wu et al., 2007; Huang and Wu, 2008) decomposes a time series into a finite number of intrinsic mode functions (IMFs) and an instantaneous mean:

$$x(t) = \sum_{j=1}^M \varphi_j(t) + R(t). \quad (7)$$

Where the  $j^{\text{th}}$  IMF  $\varphi_j$  can be written in polar coordinates as  $\varphi_j(t) = r_j(t) \times \sin[\theta_j(t)]$ , where  $r_j$  is the  $j^{\text{th}}$  time-dependent amplitude,  $\theta_j$  is the  $j^{\text{th}}$  time-dependent frequency, and  $R$  is the residual. An IMF is different from Fourier modes for which both  $r_j$  and  $\theta_j$  are time independent, and is defined by the following two properties: (1) each IMF  $\varphi_j$  has exactly one zero crossing between two consecutive local extremes (i.e. a sequence of maxima and minima), and (2) the local mean of each IMF  $\varphi_j$  is zero. The residual can be interpreted as the instantaneous mean of the time series (Huang et al., 1998). In case this instantaneous mean is not constant, we refer to it as a trend, which may be nonlinear on the time scale of the length of the time series (Wu et al., 2007).

### 2.4 The decadal tendency of runoff

The change in a variable over a given period of time is often described by the slope of a linear trend (Tank et al., 2009). The decadal tendency of runoff (DTR) is calculated as follows:

$$y(t) = a_0 + a_1 t, \quad t = 1, 2, \dots, n \text{ (year)}. \quad (8)$$

Where  $y$  is runoff (mm),  $t$  is time (sequence of year), and  $a_1$  is the rate of change in runoff tendency. For example,  $10 \times a_1$  represents the change rate over 10 years,  $a_1 > 0$  means that the climate element increases over time, and  $a_1 < 0$  indicates that the climate element decreases over time. The absolute value of  $a_1$  reflects the degree of climate element change (Qian et al., 2010).

## 3 Results and discussion

### 3.1 Model calibration and validation

The models (CLM 3.5 and CLM-DTVGM) were warmed up following a 5-year period (1980–1985) to neutralize the effect of initial conditions in the study area. Additionally, they were

calibrated at five gauging stations for a 20-year period (1985–2005) and validated for a 5-year period (2006–2010).

Evaluation and comparison of the two models for calibration and validation were carried out through graphical comparisons and statistics.

Table 1 lists the comparison of statistics of goodness of fit for monthly stream flow simulations by two models at the five stations. It reports that the  $R^2$  value of CLM-DTVGM was greater than that of CLM 3.5 for the calibration and validation period except at Changmapu Station. In addition, the NSE and the B values of CLM-DTVGM were better than those of CLM 3.5 for the calibration and validation periods at the gauging stations.

**Table 1** Goodness of two models' fit of for monthly stream flow simulations

Station	CLM 3.5						CLM-DTVGM					
	Calibration (1985–2005)			Validation (2006–2010)			Calibration (1985–2005)			Validation (2006–2010)		
	$R^2$	NSE	B (%)	$R^2$	NSE	B (%)	$R^2$	NSE	B (%)	$R^2$	NSE	B (%)
Dashankou	0.93	0.81	16	0.92	0.75	16	0.94	0.93	5	0.92	0.92	4
Kaqun	0.93	0.83	24	0.96	0.89	25	0.97	0.96	–11	0.98	0.96	–13
Changmapu	0.92	0.68	32	0.88	0.64	23	0.86	0.81	10	0.72	0.70	28
Zhengyixia	0.87	0.73	21	0.88	0.71	23	0.94	0.92	22	0.89	0.88	7
Tangnaihai	0.82	0.69	24	0.87	0.74	20	0.84	0.82	4	0.89	0.88	2

Note: CLM, Community Land Model; CLM-DTVGM, Community Land Model-Distributed Time Variant Gain Model; NSE, Nash-Sutcliffe efficiency; B, percent bias.

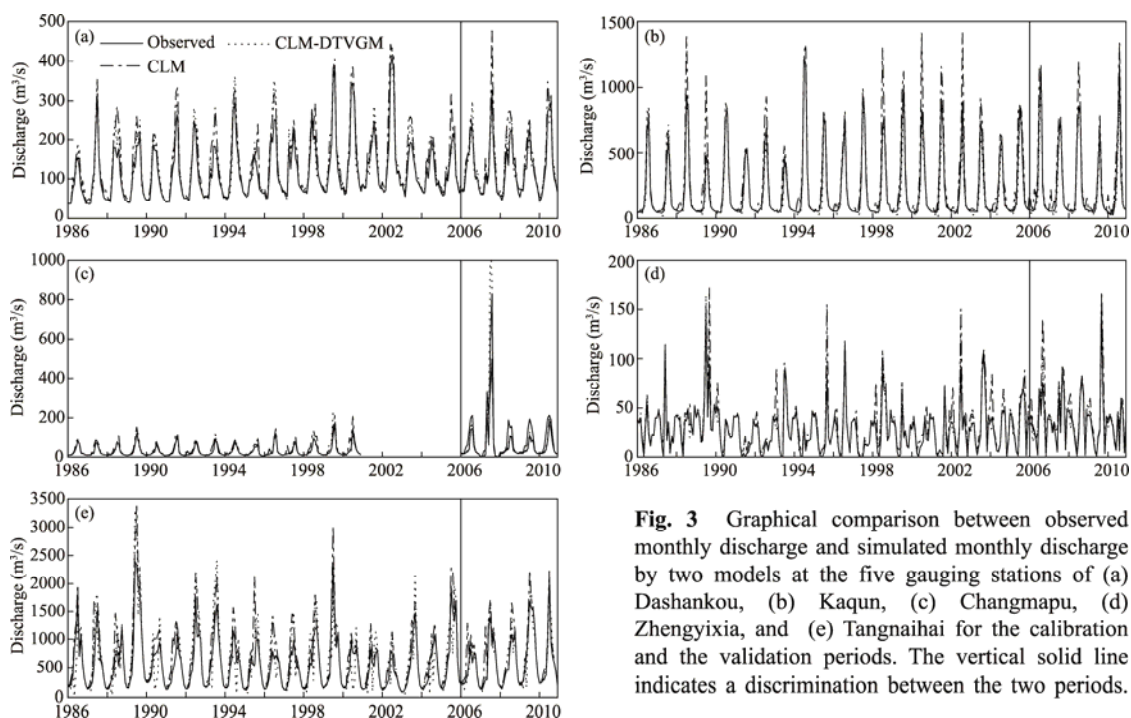
As for CLM-DTVGM, the  $R^2$  value of the model fitting for observed and simulated discharge was within 0.83–0.97 for the calibration period, and within 0.72–0.98 for the validation period. The average  $R^2$  value at the five gauging stations was 0.91 for the calibration period and 0.88 for the validation period. The NSE for the calibration period varied from 0.80 at Tangnaihai Station to 0.96 at Kaqun Station, and it varied from 0.72 at Changmapu Station to 0.96 at Kaqun station for the validation period. The average NSE at the five gauging stations was 0.86 for the calibration period and 0.83 for the validation period. In addition, the B values was –0.11 for the calibration period at Kaqun Station, which means that the simulated discharge was 11% less than the observed discharge. Also, for the validation period, the B values was –0.13, meaning that the simulated discharge was 13% less than the observed. However, the percent bias was greater than 0 at the other four gauging stations. For the calibration period, the B values changed from 0.05 at Dashankou Station to 0.22 at Zhengyixia Station, and the B values in simulated and observed values vary within 5%–22%. For the validation period, the B values varied from 0.04 at Dashankou Station to 0.28 at Changmapu station, and the differences in simulated and observed values varied within 4%–28%. The average B values at the five gauging stations was 0.07 for the calibration period and 0.10 for the validation period, which indicates that the simulated discharge may be slightly larger than the observed discharge.

Figure 3 provides a graphical comparison between observed monthly flows and simulated monthly flows by two models at five stations for the calibration and validation periods. It shows that the discharge simulated by CLM 3.5 at the five stations were tended to overestimate the peak flows, compared with the discharge projected by the CLM-DTVGM.

According to Fig. 3 and Table 1, the CLM-DTVGM can better track the observed monthly discharge than the CLM 3.5 for both the calibration and validation periods, although some of the future peak flows are either lower or higher than the observed data.

### 3.2 Spatial distribution of simulated future average annual runoff

The spatial distribution of simulated average annual runoff over a 90-year future timescale is illustrated in Fig. 4. In general, the average annual runoff was spatially, unevenly distributed, and varied across the study area from low (approximately 10 mm) to high (approximately 470 mm),



**Fig. 3** Graphical comparison between observed monthly discharge and simulated monthly discharge by two models at the five gauging stations of (a) Dashankou, (b) Kaqun, (c) Changmapu, (d) Zhengyixia, and (e) Tangnaihai for the calibration and the validation periods. The vertical solid line indicates a discrimination between the two periods.

influenced by regional atmospheric conditions and topography. The average annual runoff increased by approximately 0.52% under the RCP 4.5 scenario, and by approximately 1.23% under the RCP 8.5 scenario. For the period 2010–2100, the average annual runoff in the Tianshan Mountains was relatively high for all the three RCP scenarios (Figs. 4a–c). However, the average annual runoff in the Turpan-Hami Basin and downstream of the Hexi Corridor was relatively low, due to scarce local precipitation and a widespread Gobi with a sparse river network.

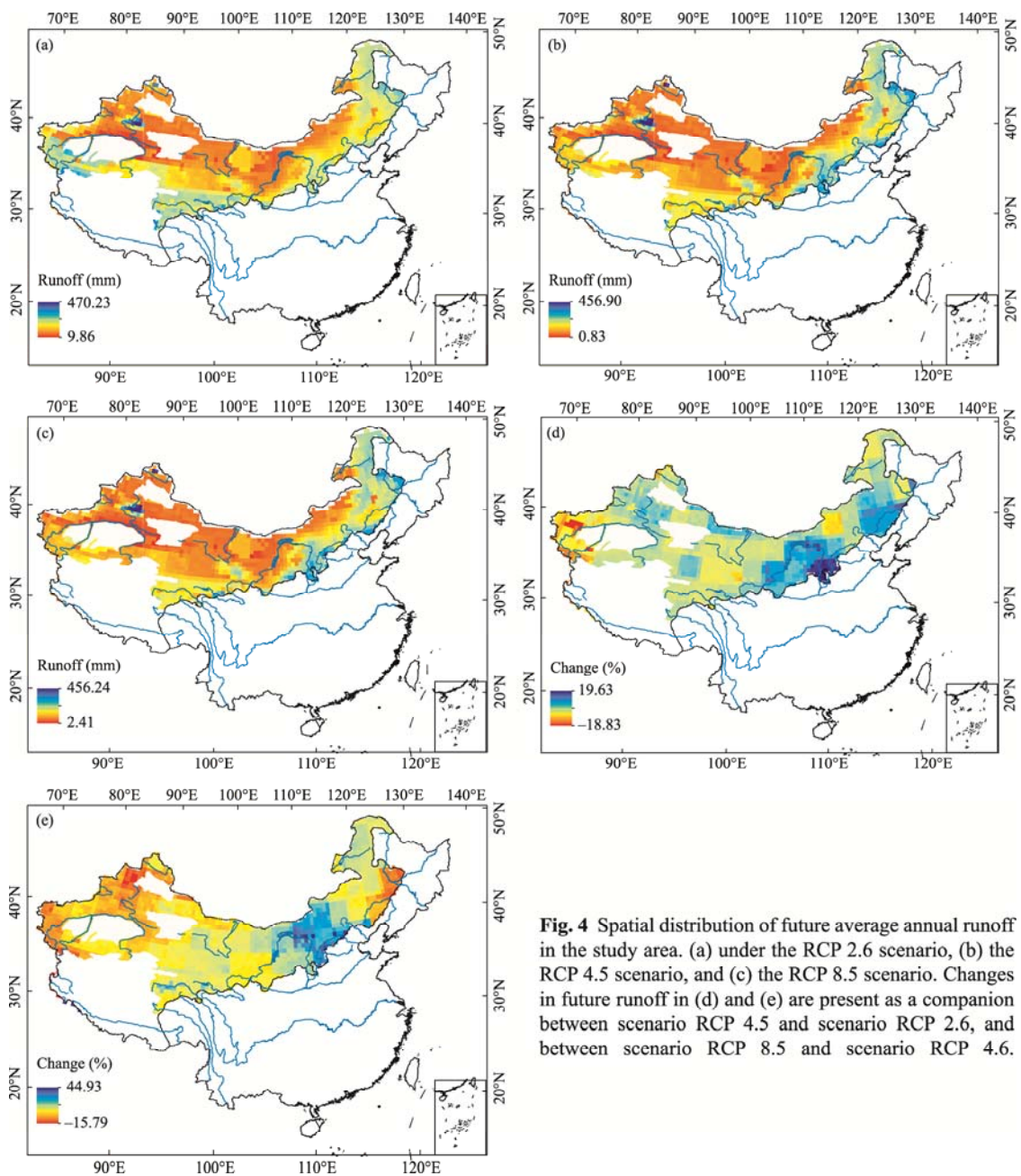
Furthermore, Figs. 4d and e also show changes in future average annual runoff with different RCP scenarios for the period 2010–2100. The eastern and western parts of the study area experienced different changes in future average annual runoff. Compared with RCP 2.6, the changes in future average annual runoff under the RCP 4.5 scenario ranged from  $-18.8\%$  in the upper reaches of the Tarim River to  $19.63\%$  in the middle Yellow River region. Additionally, the future average annual runoff increased in the middle Tianshan Mountains, the middle Yellow River region and the upper reaches of the Xiliaohe and Songhua rivers. However, the changes in future average annual runoff under the RCP 8.5 scenario ranged from approximately  $-15.79\%$  at the western margin of the Junggar Basin to  $44.93\%$  in Karakorum Mountains, which has many glaciers. Furthermore, the future average annual runoff in most parts of the study area increased, except in Xinjiang and the upper parts of the Xiliaohe and Songhua rivers.

### 3.3 EEMD analysis for future average annual runoff

The EEMD analysis found five IMFs extracted from the future areal average annual runoff in the study area which range from high frequency to low frequency, and indicate the scale of annual to biennial or decadal time spans as shown in Fig. 5.

Taking EEMD decomposition of the average annual runoff under the RCP 2.6 scenario as an example (Fig. 5a), the original future annual runoff signal was displayed at the top. The IMFs were presented in a descending order. IMF1 was displayed following the original signal and can be seen as the one with the highest-frequency fluctuation. The subsequent IMFs contained fluctuations with increasingly lower frequencies. The bottom shows the instantaneous mean, which in this study was interpreted as a trend.

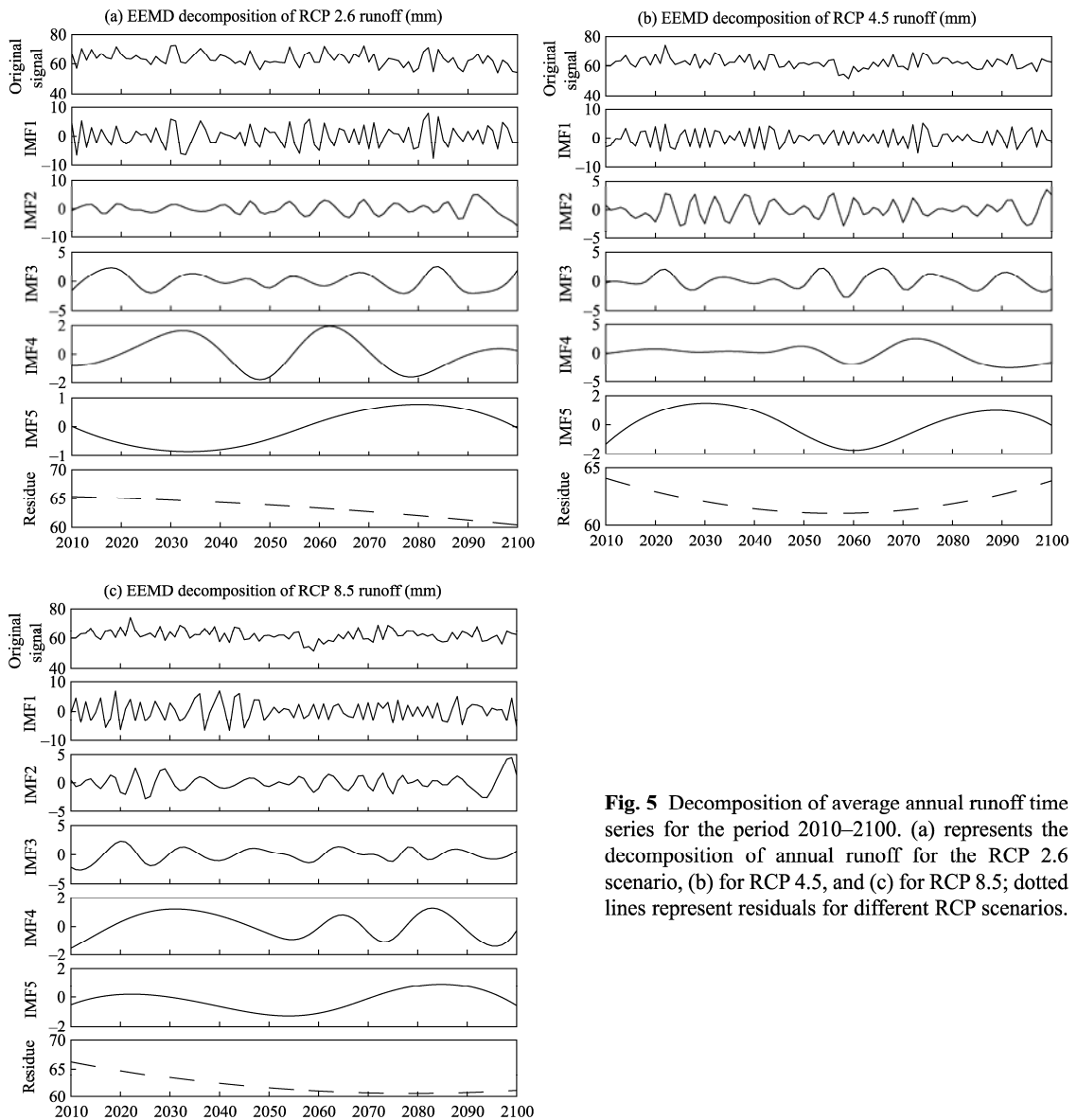




**Fig. 4** Spatial distribution of future average annual runoff in the study area. (a) under the RCP 2.6 scenario, (b) the RCP 4.5 scenario, and (c) the RCP 8.5 scenario. Changes in future runoff in (d) and (e) are present as a companion between scenario RCP 4.5 and scenario RCP 2.6, and between scenario RCP 8.5 and scenario RCP 4.6.

The extracted IMF components showed nonlinear variations of original annual runoff signals, revealing a temporal distribution on multiple time scales. They also elaborated the fluctuations of the major periods in the time-frequency domain. For each scenario, the statistical results of five IMF components obtained from the EEMD are reported in Table 2.

According to Fig. 5 and Table 2, the original runoff variations and the major periods of different IMFs varied considerably. The major periods of IMF1 was 2–3 years for all the three RCP scenarios. The period was relatively stable in the time-frequency domain, which may be associated with the quasi-biennial oscillation and quasi-triennial oscillation (Naujokat, 1986; Tian et al., 2012). The major periods of IMF2 was 5–7 years, and it may be associated with an atmospheric oscillation closely connected to the El Niño Southern Oscillation (ENSO). IMF3 exhibited runoff variations on a scale of 10–13 years, which was similar to the 11-year cycle of sunspots (Katsavrias et al., 2012).



**Fig. 5** Decomposition of average annual runoff time series for the period 2010–2100. (a) represents the decomposition of annual runoff for the RCP 2.6 scenario, (b) for RCP 4.5, and (c) for RCP 8.5; dotted lines represent residuals for different RCP scenarios.

**Table 2** Central frequencies and average cycles of different IMFs

IMF	Center frequency (/a)			Average cycle (a)		
	RCP 2.6	RCP 4.5	RCP 8.5	RCP 2.6	RCP 4.5	RCP 8.5
IMF1	0.330	0.396	0.363	3.0	2.5	2.8
IMF2	0.143	0.209	0.187	7.0	4.8	5.4
IMF3	0.077	0.099	0.088	13.0	10.1	11.4
IMF4	0.027	0.028	0.036	37.5	35.5	28.1
IMF5	0.012	0.019	0.021	83.7	53.7	47.2

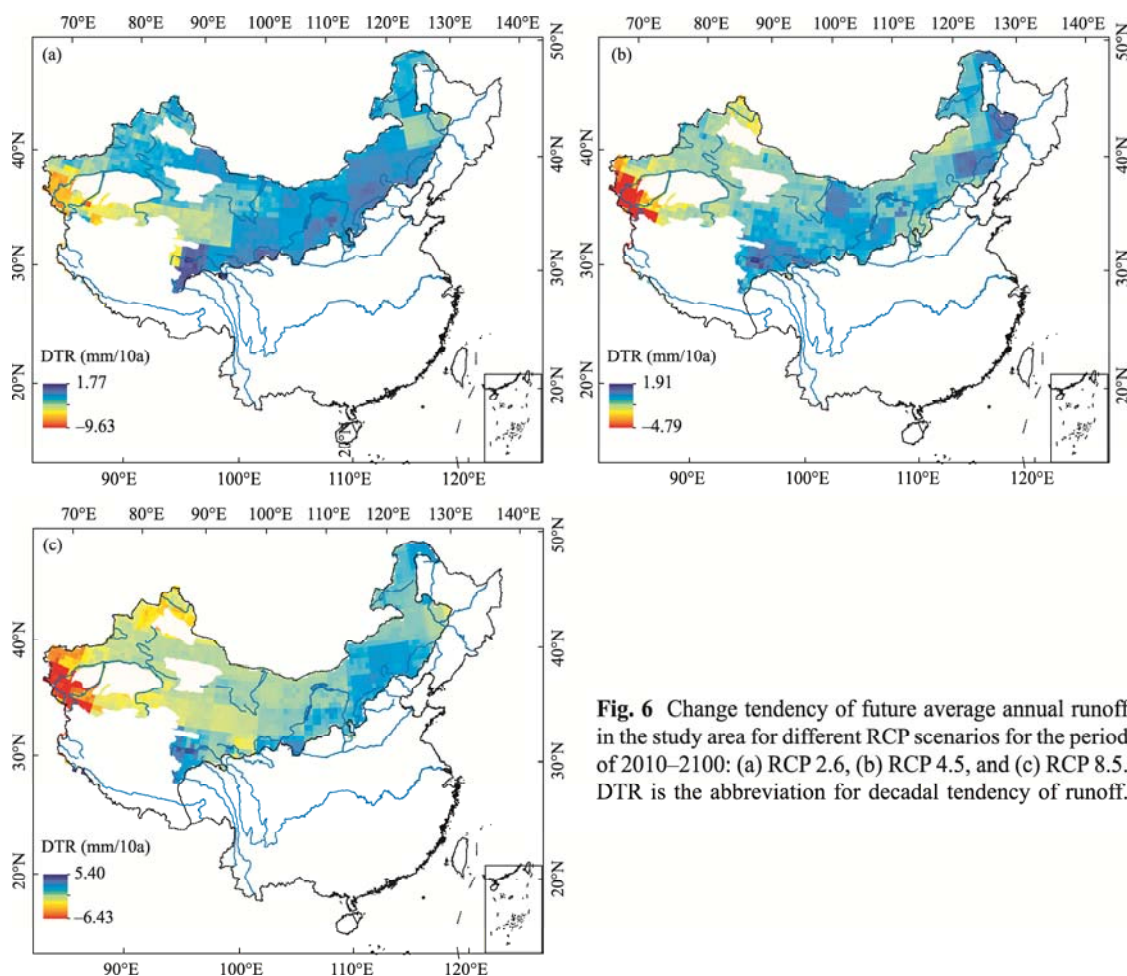
Note: IMF, the intrinsic mode function; RCP, the Representative Concentration Pathways.

IMF4 represented multi-decadal variations in runoff. IMF5 came as the last residue after all of the IMFs were extracted, which indicated an overall trend for the coming 90 years. However, the multi-decadal variations and overall trend for different RCP scenarios varied. The major period of

IMF4 was approximately 35 years for the RCP 2.6 and RCP 4.5 scenarios, but it was approximately 30 years for the RCP 8.5 scenario. Additionally, the annual runoff had a slowly decreasing trend for the RCP 2.6 and RCP 8.5 scenarios, whereas it showed a non-monotonic trend for the RCP 4.5 scenario, with a slow increase after the 2050s.

### 3.4 Decadal variability of future average annual runoff

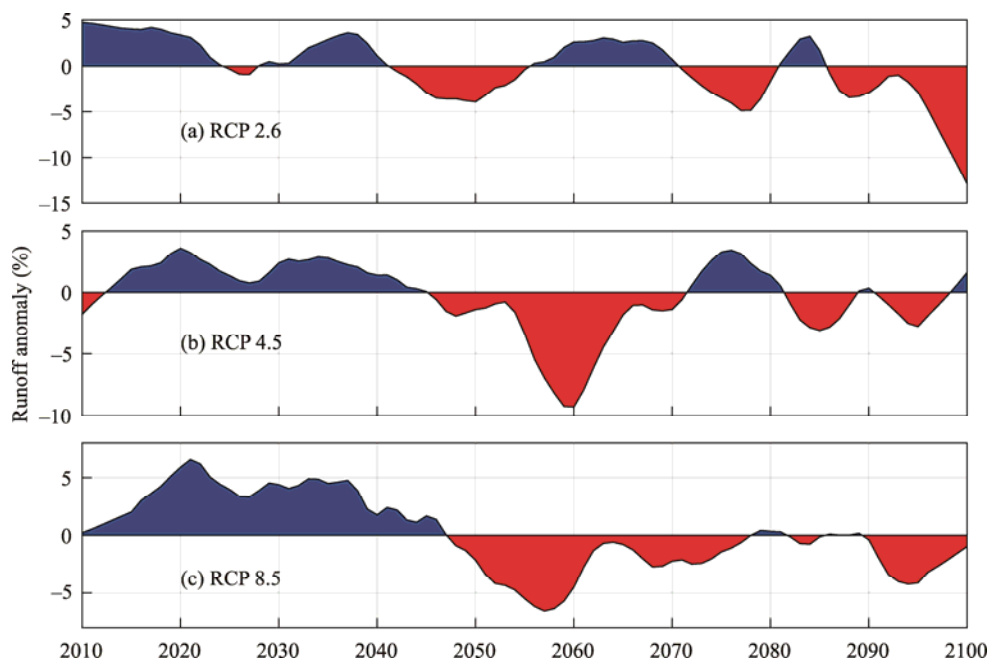
To investigate the inter-annual variation of future average annual runoff, we calculated the change tendency for the period of 2010–2100, as shown in Fig. 6. The change rate of future average annual runoff was  $-0.45$  mm/10a for the RCP 2.6 scenario,  $-0.39$  mm/10a for RCP 4.5, and  $-0.14$  mm/10a for RCP 8.5. Moreover, Fig. 6 indicates that as a result of climate change, the runoff change rate during the coming 90 years varied spatially among the different RCP scenarios. According to the statistics in the GIS system, the area with a decreasing trend occupied 60.5% of the study area for the RCP 2.6 scenario, and the percentages were 48.6% and 58.6% for the RCP 4.5 and RCP 8.5 scenarios, respectively. In light of the spatial distribution of the change rate, most of Xinjiang had a decreasing trend for all the three RCP scenarios, and the eastern part of the study area had an increasing trend during the period 2010–2100.



**Fig. 6** Change tendency of future average annual runoff in the study area for different RCP scenarios for the period of 2010–2100: (a) RCP 2.6, (b) RCP 4.5, and (c) RCP 8.5. DTR is the abbreviation for decadal tendency of runoff.

An understanding of the decadal or multi-decadal variability of runoff is important to water resources management under climate change. In this study, the normalized annual anomalies were smoothed to obtain the decadal components of the future average annual runoff series. The method is a robust version of “loess” that assigns lower weights to outliers in the regression, and also assigns zero weight to data outside six mean absolute deviations (Cleveland, 1979). The

span in this method is a percentage of the total number of data points, which in this study was 0.11, representing approximately 10 years. Decadal fluctuations were detected in future average annual runoff, as shown in Fig. 7.



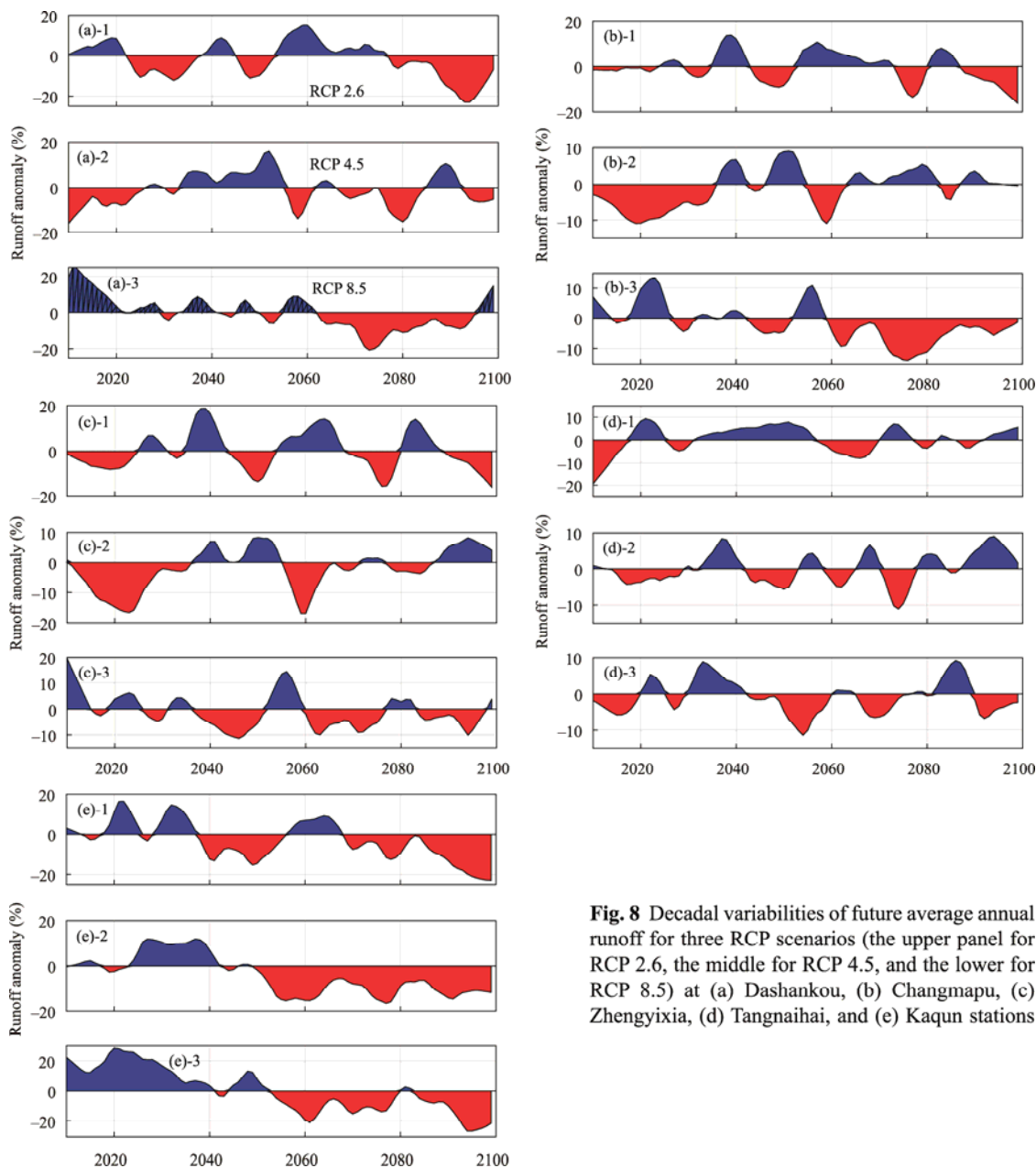
**Fig. 7** Decadal variabilities of future average annual runoff for three RCP scenarios: (a) RCP 2.6, (b) RCP 4.5, and (c) RCP 8.5

The results show that the future average annual runoff clearly exhibited decadal variability, which indicates that the average annual runoff had a clear division between dry and wet periods. For all the three RCP scenarios, the average annual runoff tended to alternate between dry and wet divisions. However, the alternating cycles and transitions differed. There were four alternating cycles for the coming 90 years, and the transitions between wet and dry initiated in the years 2040, 2055, 2070 and 2080 for the RCP 2.6 scenario. As a comparison, there were three alternating cycles for the RCP 4.5 scenario and two for the RCP 8.5 scenario.

Furthermore, the span of dry periods varied among the three RCP scenarios. The longest dry period was approximately 15 years during 2040–2055 for the RCP 2.6 scenario, and 25 years during 2045–2070 for the RCP 4.5 scenario. However, for the RCP 8.5 scenario, there is a long-spanning dry period starting from 2045. Under this scenario, the water resources situation of the study area will be then extremely severe.

### 3.5 Decadal variability of future average annual runoff in the test catchments

The decadal variabilities of the future average annual runoff in five test catchments were also analyzed, using the same method as mentioned above. The results were illustrated in Fig. 8. These results illustrate that the future average annual runoff in the five test catchments clearly exhibited decadal variabilities, indicating clear wet-dry cycles. Furthermore, the alternating cycles and transitions differed with RCPs and basins. Taking Dashankou Station as an example, three main dry periods were present for the RCP 2.6 scenario for the coming 90 years, which were during 2022–2040, 2050–2060 and 2075–2099, respectively. Two main dry periods occurred during 2010–2030 and 2045–2085 for the RCP 4.5 scenario. However, only one main dry period was found during 2065–2095 for the RCP 8.5 scenario. In addition, the situation of water resources at Kaqun Station was severe for the coming 90 years, especially for the RCP 4.5 and RCP 8.5 scenarios.



**Fig. 8** Decadal variabilities of future average annual runoff for three RCP scenarios (the upper panel for RCP 2.6, the middle for RCP 4.5, and the lower for RCP 8.5) at (a) Dashankou, (b) Changmapu, (c) Zhengyixia, (d) Tangnaihai, and (e) Kaqun stations

#### 4 Conclusion

The coupled model (CLM-DTVGM) can better track the observed monthly discharge than CLM 3.5 in the study area. The future annual runoff had a slow decreasing trend for the RCP 2.6 and RCP 8.5 scenarios during the period 2010–2100, whereas it had a non-monotonic trend for the RCP 4.5 scenario, with a slow increase after the 2050s. The future annual runoff exhibited decadal variations, indicating that the annual runoff had a clear division of dry and wet periods. The longest dry period was 15 years during 2040–2055 for the RCP 2.6 scenario, and it was 25 years during 2045–2070 for the RCP 4.5 scenario. However, the RCP 8.5 scenario predicts a long dry period starting from 2045. The same phenomenon was shown at the five test catchments. Under these circumstances, the water resources situation of the study area will be extremely severe. Therefore, future work focusing on risk assessment and management of water resources is imperative in this study area.

## Acknowledgements

This work was supported by the National Basic Research Program of China (2012CB956204). We acknowledge the modeling groups for providing the data for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project for collecting and archiving the model output and organizing the data analysis. Reviewer comments greatly enriched the manuscript.

## References

- Beven K J. 2011. *Rainfall-Runoff Modelling: The Primer*. Hoboken: John Wiley & Sons.
- Cammalleri C, Micalle F, Vogt J. 2015. On the value of combining different modelled soil moisture products for European drought monitoring. *Journal of Hydrology*, 525: 547–558.
- Choi H I, Liang X Z. 2010. Improved terrestrial hydrologic representation in mesoscale land surface models. *Journal of Hydrometeorology*, 11(3): 797–809.
- Cleveland W S. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368): 829–836.
- Collins W D, Bitz C M, Blackmon M L, et al. 2006. The community climate system model version 3 (CCSM3). *Journal of Climate*, 19(11): 2122–2143.
- Dai Y J, Zeng X B, Dickinson R E. 2001. *The Common Land Model (CLM): Technical Documentation and User's Guide*. Georgia: Georgia Institute of Technology.
- Gudmundsson L, Tallaksen L M, Stahl K, et al. 2012. Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe. *Journal of Hydrometeorology*, 13(2): 604–620.
- Gupta H V, Sorooshian S, Yapo P O. 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2): 135–143.
- Huang N E, Shen Z, Long S R, et al. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: *Proceedings of Mathematical, Physical and Engineering Sciences*. London: The Royal Society, 903–995.
- Huang N E, Wu Z H. 2008. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, 46(2), doi: 10.1029/2007RG000228.
- Kallache M, Rust H W, Kropp J. 2005. Trend assessment: applications for hydrology and climate research. *Nonlinear Processes in Geophysics*, 12(2): 201–210.
- Katsavrias C, Preka-Papadema P, Moussas X. 2012. Wavelet analysis on solar wind parameters and geomagnetic indices. *Solar Physics*, 280(2): 623–640.
- Lai X, Wen J, Cen S X, et al. 2014. Numerical simulation and evaluation study of soil moisture over China by using CLM 4.0 model. *Chinese Journal of Atmospheric Sciences*, 38(3): 499–512. (in Chinese)
- Li M X, Ma Z G, Niu G Y. 2011. Modeling spatial and temporal variations in soil moisture in China. *Chinese Science Bulletin*, 56(17): 1809–1820.
- Moss R H, Edmonds J A, Hibbard K A, et al. 2010. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282): 747–756.
- Nash J E, Sutcliffe J V. 1970. River flow forecasting through conceptual models part I-A discussion of principles. *Journal of Hydrology*, 10(3): 282–290.
- Naujokat B. 1986. An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *Journal of the Atmospheric Sciences*, 43(17): 1873–1877.
- Niu G Y, Yang Z L, Dickinson R E, et al. 2005. A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *Journal of Geophysical Research: Atmospheres*, 110(D21): D07103.
- Oleson K W, Dai Y G, Bonan G, et al. 2004. Technical description of the community land model (CLM). In: *NCAR Technical Note NCAR/TN-461+STR*. National Center for Atmospheric Research. Colorado, USA.
- Oleson K W, Niu G Y, Yang Z L, et al. 2008. Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research: Biogeosciences*, 113(G1): G01021.
- Qian S, Fu Y, Pan F F. 2010. Climate change tendency and grassland vegetation response during the growth season in Three-River Source Region. *Science China Earth Sciences*, 53(10): 1506–1512.
- Shangguan W, Dai Y J, Liu B Y, et al. 2013. A China data set of soil properties for land surface modeling. *Journal of Advances in Modeling Earth Systems*, 5(20): 212–224.



- Sheffield J, Goteti G, Wood E F. 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(13): 3088–3111.
- Song X M, Zhan C S, Xia J. 2012. Integration of a statistical emulator approach with the SCE-UA method for parameter optimization of a hydrological model. *Chinese Science Bulletin*, 57(26): 3397–3403.
- Tank A M G K, Zwiers F W, Zhang X B. 2009. Guidelines on: Analysis of extremes in a changing climate in support of informed decisions for adaptation. Switzerland: World Meteorological Organization.
- Thompson S A. 1999. *Hydrology for Water Management*. Rotterdam: Balkema.
- Tian Q H, Zhou X J, Gou X H, et al. 2012. Analysis of reconstructed annual precipitation from tree-rings for the past 500 years in the middle Qilian Mountain. *Science China Earth Sciences*, 55(5): 770–778.
- Wang A H, Lettenmaier D P, Sheffield J. 2011. Soil moisture drought in China, 1950–2006. *Journal of Climate*, 24(13): 3257–3271.
- Wu Z H, Huang N E, Long S R, et al. 2007. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proceedings of the National Academy of Sciences of the United States of America*, 104(38): 14889–14894.
- Wu Z H, Huang N E. 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1): 1–41.
- Xia J. 2002. *Hydrological Nonlinear Theories and Approaches*. Wuhan: Wuhan University Press. (in Chinese)
- Xia J, Wang G S, Tan G, et al. 2005. Development of distributed time-variant gain model for nonlinear hydrological systems. *Science in China Series D: Earth Sciences*, 48(6): 713–723.
- Xu W T, Wu B F, Yan C Z, et al. 2005. China land cover 2000 using SPOT VGT S10 data. *Journal of Remote Sensing*, 9(2): 204–214. (in Chinese)
- Yang J P, Ding Y J, Chen R S, et al. 2002. Spatial change of dry and wet climate boundary in China in the recent 50 years. *Journal of Glaciology and Geocryology*, 24(6): 731–736. (in Chinese)