DESERT

Desert
Online at http://desert.ut.ac.ir

Desert 20-2 (2015) 101-115

Climate change scenarios generated by using GCM outputs and statistical downscaling in an arid region

Zh. Liu[a*], Z. Xu[b*]

[a] *Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 100101, Beijing, China*
[b] *Key Laboratory of Water and Sediment Sciences, Ministry of Education; College of Water Sciences, Beijing Normal University, Beijing 100875, China*

Received: 25 November 2015; Received in revised form: Accepted by Editor; Accepted: 8 January 2016

Abstract

Two statistical downscaling models, the non-homogeneous hidden Markov model (NHMM) and the Statistical Down–Scaling Model (SDSM) were used to generate future scenarios of both mean and extremes in the Tarim River basin, which were based on nine combined scenarios including three general circulation models (GCMs) (CSIRO30, ECHAM5, and GFDL21) predictor sets and three special report on emission scenarios (SRES) (SRES A1B, SRES A2, and SRES B1). Local climate change scenarios generated from statistical downscaling models was also compared with that projected by raw GCMs outputs. The results showed that the magnitude of changes for annual precipitation projected by raw GCMs outputs was greater than that generated by using statistical downscaling model. The difference between changes of annual maximum air temperature projected by statistical downscaling model and raw GCMs outputs was not as significant as that for annual precipitation. In total, the magnitude of these increasing trends projected by both statistical downscaling models and raw GCMs outputs was the greatest under SRES A2 scenario and the smallest under B1 scenario, with A1B scenario in–between. Generally, the magnitude of these increasing trends in the period of 2081 to 2100 was greater than that in the period of 2046 to 2065. The magnitude of standard deviation changes for daily precipitation projected by raw GCMs outputs was greater than that generated by statistical downscaling model under most of combined scenarios in both periods.

Keywords: Climate change; Statistical downscaling; Non-homogeneous hidden Markov; Probability density function; Tarim River

1. Introduction

General circulation models (GCMs) are the primary tool to simulate the present climate and project future climate (Christensen *et al*., 2007). However, GCM simulations of local climate on spatial scales smaller than grid cells are often poor, especially when the study area shows complex topography (Schubert, 1998; McAvaney *et al*., 2001). Therefore, large–scale GCM scenarios should not be used directly for impact studies (Schubert, 1997). Generating information below the grid scale of GCMs, which is referred to as downscaling, is needed in assessing the impact of climate change. There are two main approaches for downscaling, dynamical and statistical (Fowler *et al*., 2007; Christensen *et al*., 2007). Fowler et al. (2007) reviewed downscaling techniques and found out that dynamical downscaling methods provide little advantage over statistical techniques, at least for the present day climates. Murphy (1999) also pointed out that while dynamical and statistical downscaling methods generate similar reproduction of current climate (Wilby *et al*., 2000), they can however differ significantly in the projection of future climate conditions. By

* Corresponding author. Tel.: +86 10 58801136, Fax: +86 10 58801136.
E-mail address: zongxuexu@vip.sina.com

considering the advantages of being computationally inexpensive, being able to access finer scales than dynamical methods and relatively easily applied to different GCMs, parameters, and regions (Cubasch *et al*., 1996; Timbal *et al*., 2003; Wilby *et al*., 2004; Wood *et al*., 2004), two statistical downscaling (SD) methods were selected in this study.

SD methods use an empirical statistical technique to establish relationships between observed large–scale and regional/local climate (predictors and predicands) for a baseline period (Karl *et al*., 1990; Busuioc *et al*., 2001; Christensen *et al*., 2007). These relationships are then applied to downscale future climate scenarios using GCMs output. The predictors should be realistically modeled by the GCM and fully represent the climate change signal. This implies that both circulation–based and humidity predictors should be included in the downscaling model (Giorgi *et al*., 2001). SD methods can be basically classified into three types (Wilby *et al*., 2004): regression models (transfer functions), weather generators, and weather classification. In general, SD methods, that combined these techniques, are possibly the most appropriate (Christensen *et al*., 2007). The Statistical Down–Scaling Model (SDSM), which was recommended by the Canadian Climate Impacts and Scenarios (CCIS) project, includes both deterministic transfer functions and stochastic components (Wilby and Wigley, 1997; Wilby *et al*., 2002, 2003). The non-homogeneous hidden Markov Model (NHMM), which was developed by Hughes and Guttorp (1994), models multi–site patterns of daily precipitation occurrence conditioned on a finite number of 'hidden' (that is, unobserved) weather states. Charles *et al*. (1999) extended NHMM by using conditional multiple linear regressions to simulate daily precipitation amounts. Therefore, the new version of NHMM contains weather classification, regression, and stochastic components. Recently, these two downscaling methods have been widely used in statistical downscaling studies for both climate variables and air quality variables (Dibike and Coulibaly, 2005; Diaz–Nieto and Wilby, 2005; Wetterhall *et al*., 2006; Gachon and Dibike, 2007; Wise, 2009; Hughes *et al*., 1999; Bates *et al*., 1998; Robertson *et al*., 2004; Charles *et al*., 1999). All these studies have proved that SDSM and NHMM are promising methods to downscale climate change scenarios.

Tarim River Basin (TRB), which is the largest inland river basin in China (Figure 1), is famous for its rich natural resources and fragile eco–environment (Xu *et al*., 2010). The aridity index (ratio of annual potential evaporation to precipitation) is one of the greatest regions in the world, with annual precipitation of less than 150 mm, but with annual potential evaporation of 1000 to 1600 mm. Liu *et al*. (2010) found out that NHMM showed a better model performance in downscaling daily precipitation than SDSM over the TRB. Therefore, NHMM and SDSM were selected to generate future climate change scenarios for precipitation and air temperature, respectively. In assessing each model performance, indices of discrete mean or percentile values were easy to make a misleading, because these statistics only explained model performance on limited values. Therefore, distributions of variables, which could capture the entire statistical characteristics of variables, were also used for model assessment. The focus in most of the previous works for downscaling has been on reproducing the mean behavior of climate variables and little has been done to downscale variables with emphasis on reproducing extreme values. However, it is the extremes of these variables that have severe consequences and there is a growing need for reproducing and being able to reasonably project future scenarios of extreme values for impact studies. This study is, therefore, aimed at applying two different statistical downscaling methods in the generation of future scenarios for both mean and extremes values.

The main objective of this study is to construct regional daily precipitation, maximum and minimum air temperature scenarios projected by raw GCMs simulations and statistical downscaling models in the TRB, based on nine combined scenarios including three GCMs (CSIRO30, ECHAM5, and GFDL21) predictor sets and three Special Report on Emission Scenarios (SRES) (SRESA1B, SRESA2, and SRESB1). Besides this fact, regional climate change scenarios generated between statistical downscaling models and raw GCMs outputs were also compared with each other.

2. Dataset Description

2.1. Observed stations data

There were 25 observed stations data used in this study, which included daily time series of precipitation, maximum and minimum air temperatures (T_{max} and T_{min}) with continuous data series of 1961 to 2000. These stations were from

National Meteorological Administration of China, which applies data quality control (such as, stations correction and equipments correction) before releasing these data. The geographical locations and spatial distribution of these stations are as shown in Figure 1.

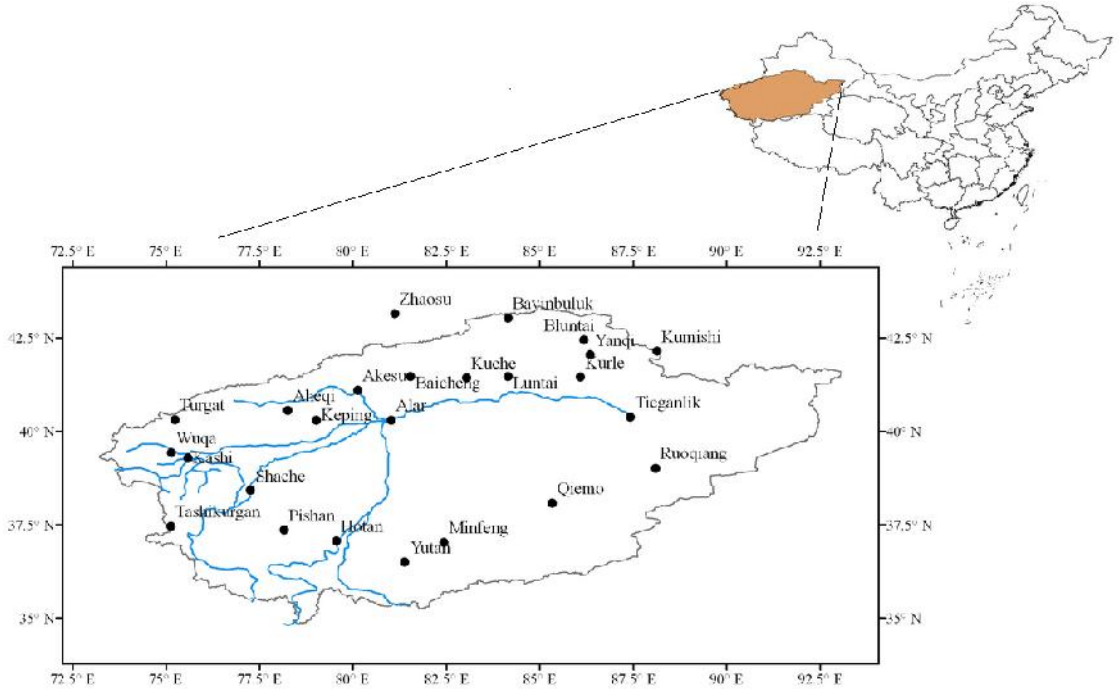

Fig. 1. Location of the Tarim River basin and observed stations used in this study

2.2. NCEP reanalysis data

The National Centers for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis datasets (Kalnay *et al*., 1996; Kistler *et al*., 2001) for the period of 1961 to 2000 containing the suite of daily atmospheric variables were selected as observed large scale predictors. Both pressure and surface variables were selected as candidate predictors (Table 1), which include five pressure variables with three pressures (400, 500, and 600 hPa) and five surface variables. Three pressures were selected, because data for the pressures more than 600 hPa were not available from GCMs outputs (see Section 2.3) in the study area. The raw reanalysis datasets, which have several different horizontal resolutions for different variables, were interpolated into 2.5×2.5º resolution grids for all candidate predictors in this study. There were twelve grids covering the TRB (Figure 1). All reanalysis data has been standardized before downscaling. Standardized predictors are routinely employed in order that the same models could be applied to climate scenario generation using standardized GCM output.

2.3. GCMs outputs

Model output from three different GCM datasets, which include CSIRO30, ECHAM5, and GFDL21, were used to assess TRB precipitation and air temperature changes in the 21st century. All GCMs runs used in this study were run1. Three emission scenarios of the Intergovernmental Panel on Climate Change (IPCC) SRES, SRES A1B, A2, and B1 simulations from each of the GCMs were selected. Data series include both present control period (1960 to 2000) and future period (2046 to 2065 and 2081 to 2100).

The SRES A1B describes a future world with rapid economic growth, global population that peaks in mid–century and declines thereafter, and rapid introduction of new and more efficient technologies, which is based on a balance of all energy sources. The SRES A2 describes a very heterogeneous world, which results in continuously increasing population, and its economic development is primarily regionally oriented and technological change more fragmented and smaller than other scenarios. In the SRES B1 scenarios, a convergent world with the same global population,

that peaks in mid–century and declines thereafter, as in the SRES A1B scenario, but with rapid change in economic structures with reductions in material intensity and introduction of clean and resource–efficient technologies. More details on SRES emission scenarios could be found in the IPCC AR4 (IPCC, 2007).

Table 1. Candidate predictors

ID	Full name	classification
air	air temperature	Pressure variables (three pressures 400hPa, 500hPa and 600hPa)
hgt	geopotential height	
shum	specific humidity	
uwnd	east–west velocity	
vwnd	north–south velocity	
slp	air pressure at sea level	Surface variables

3. Methodology Description

3.1. Model settings and evaluation criteria

Long observed data series (1961 to 2000) were split into periods of 1981 to 2000 and 1961 to 1980, and were used for model calibration and validation, respectively. The period of 1981 to 2000 was selected for model calibration because both observed and NCEP data have more quality data than previous periods (Kalnay *et al*., 1996; Kistler *et al*., 2001). The NHMM was fit on a seasonal basis with the wet season (April to September) and the dry season (October to March), respectively, in which precipitation takes up 84.3 and 15.7% of annual precipitation. The SDSM was based on monthly type, in which model parameters are derived for each month. Models calibrated and validated using observed predictors derived from reanalysis data were then driven by predictors derived from large-scale predictor variables simulated by GCMs in the control period representing the present climate (1981 to 2000) and future periods (2046 to 2065 and 2081 to 2100). Finally, plausible future changes of daily precipitation and air temperature were estimated as the difference of downscaled values between those using predictors derived from GCMs outputs corresponding to the future period under the IPCC SRES A1B, A2, and B1 emission scenarios and the corresponding values run for the control period.

Several model performance criteria were selected to evaluate the skill of the downscaling models to reproduce observed precipitation and extreme air temperature. First, the model bias was evaluated by the relative error (RE) which reflected the difference between modeled and observed mean values. Extreme events being of interest in climate impact assessments also need to be reasonably well modeled. Therefore, percentile values including 5 and 95% values (referenced as 5 and 95th), the minimum and maximum values of each variable, were also considered. As for precipitation, three characteristics included wet–day precipitation amount (wpa), wet–spell length (wsl), and dry–spell length (dsl) were analyzed.

The normalized root mean square error (NRMSE), defined as root mean square error divided by the corresponding standard deviation of the observed field (Randall et al., 2007), is an index considering both mean value and standard deviation for time series:

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_{mi} - X_{oi})^2}}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_{oi} - \overline{X_o})^2}} \quad (1)$$

Empirical orthogonal function (EOF) analysis was used to compare the space-time variability of the observed and modeled data (Harvey and Wigley, 2003). One advantage of using EOFs is the ability to identify and quantify the spatial structures of correlated variability (Mu *et al*., 2004). The first two leading modes of each EOF, accounting for majority of the total variance, were compared.

Finally, two skill scores, BS and S_{score}, based on probability density functions (PDFs) were used to measure how well each model can capture the PDFs of each variable, which was assessed with daily data series and computed as follows:

$$BS = \frac{1}{n}\sum_{i=1}^{n}(P_{mi} - P_{oi})^2 \quad (2)$$

$$S_{score} = \sum_{i=1}^{n} Minimum(P_{mi}, P_{oi}) \quad (3)$$

where P_{mi} and P_{oi} are the modeled and observed ith probability values of each bins, n is the number of bins. Observed and modeled data were binned around centers determined by the range of each data series. Therefore, the bin sizes were different for each data series. However, the bins' number is

fixed as 100 for all data series. All daily values of precipitation below 0.1 mm/day were omitted because rates below this amount are not recorded in the observations. The BS is a mean squared error measure for probability forecasts (Brier, 1950) and S_{score} calculates the cumulative minimum value between observed and modeled distributions for each binned value, thereby can measure the overlap area between two PDFs (Perkins *et al*., 2007).

All objective functions were evaluated for both calibration and validation periods.

3.2. Predictors for downscaling methods

Selection of predictor variables should be given high consideration in statistical downscaling methods. Circulation variables, which are well simulated by GCMs, are usually selected as candidate predictors for downscaling (Cavazos and Hewitson, 2005). However, it is increasingly acknowledged that circulation variables alone are not sufficient, as they fail to capture key precipitation mechanisms based on thermodynamics and moisture content. Therefore, humidity has increasingly been used to downscale precipitation (Karl *et al*., 1990; Wilby and Wigley, 1997; Murphy, 2000; Beckmann and Buishand, 2002), particularly as it may be an important predictor under a changed climate. It is essential to have prior knowledge of climate model limitations when selecting candidate predictors before using downscaling tools. It is suggested that the optimal predictors must be strongly correlated with the predictand, being physically sensible and can capture multiyear variability (Wilby and Wigley, 2000; Wilby *et al*., 2004; Gachon and Dibike, 2007). In this study, 16 NCEP reanalysis large–scale variables (see section 2.2.2) that had potential physical relationships with precipitation and air temperature and were well simulated by GCMs, and were selected as candidate observed predictors. These predictors include not only circulation variables (that is, geopotential and wind component), but also air temperature and moisture variables (specific humidity). Exploratory and partial analyses were used to determine which candidate predictors had the strongest statistical relationship with each variable over the region. The choice of predictor domain was based on the method developed by Liu *et al*. (2010). Multi-grid predictor domains were selected for each observed station in the application of SDSM. Potential grids include the grid nearest to the observed station and the eight grids surrounding it. The predictor domains were selected from these candidate potential grids based on mean sea level pressure gradients. For each station, the centre grid in which a station was located and its neighbor grids which have a mean sea level pressure gradient towards this centre grid were selected as the grid domain for the station in SDSM.

3.3. SDSM method

As mentioned above, SDSM is a hybrid between stochastic weather generator and multilinear regression method. That is because regional circulation patterns and atmospheric moisture variables are used to condition local precipitation at each station. The regression component represents the deterministic part of the model, while the stochastic component is a random element of the model (Prudhomme and Davies, 2009). The multilinear regression of the model is used to derive a statistical relationship between predictors and precipitation, which includes some transform functions in order to obtain secondary data series of precipitation and predictors that have stronger correlations than the original data series. It allows the prediction of local weather conditions from large–scale predictor variables simulated by GCMs under baseline and changed climatic scenarios. Precipitation is then modeled through the stochastic weather generator conditioned on predictors. Stochastic component was added to replicate a variance closer to the observed variability, because the chaotic nature of local weather conditions cannot be fully explained by the variations in predictor variables. Therefore, SDSM enables to generate multiple simulations with slightly different time series attributes, but the same overall statistical properties. Full technical details of SDSM could be found in Wilby *et al*. (2002).

Generally, application of SDSM contains four steps: (1) selection of predictors; (2) model calibration; (3) weather generator; and (4) generation of future climate variables. Model calibration is based on multiple linear regression equations, given daily precipitation and large scale atmospheric variables. Then, daily precipitation is simulated by weather generator in the model, which is based on NCEP re–analysis predictors. Assessment of the model is based on comparison between the simulated and observed series. Finally, the calibrated SDSM was applied to generate future climate variables.

3.4. NHMM method

The NHMM can be defined by a state transition probability matrix and a precipitation occurrence probability distribution (Charles *et al*., 1999). It defines stochastic conditional relationships between multi–site daily precipitation occurrence patterns and a discrete set of weather states, which are referred to as hidden states, because they are not directly observable. Then, the transition probabilities between these hidden states, conditioning on a set of predictors and not fixed as in the homogeneous HMM (so called non-homogeneous HMM), are defined by a first–order Markov chain. Finally, multi–site daily precipitation amounts are simulated by conditional multiple linear regression.

There are two assumptions in the NHMM. The first assumption is that precipitation occurrence is only conditioned on a small number of given hidden states. By conditioning precipitation occurrence on hidden states, rather than atmospheric circulation patterns, the NHMM can capture much of the spatial and temporal variability of daily multi–site precipitation occurrence records (Hughes *et al*., 1999; Charles *et al*., 1999). Other assumption describes that the hidden state on a given day only depends on the previous day's state, which corresponds to the Markov property.

The application of NHMM contains five steps: (1) choice of predictors; (2) selection for hidden state number; (3) generation of precipitation occurrence; (4) simulation for precipitation amount; and (5) model validation. Predictors are selected in section 3.1. The most appropriate hidden state number is estimated based on maximum likelihood, which is evaluated by the Bayes Information Criterion (BIC). Once hidden state number has been estimated, the most likely daily sequence of hidden states, along with dry states (zero precipitation) can be determined by the Viterbi algorithm (Forney, 1978), a dynamic programming scheme. Then, daily precipitation amount in wet days at observed stations is modeled for each state based on the Gamma distribution. In all cases, the NHMM was run 10 times from random seeds. Full details on NHMM could be found in Charles *et al*. (1999).

4. Results Analysis

4.1. Predictor selection

Table 2 presents the predictors selected for T_{max} and T_{min} in SDSM, and for precipitation in NHMM, respectively. The shum400, shum500, shum600, vwnd600, hgt400, slp, and air600 are selected as predictors for the NHMM, based on explained variances. These predictors are consistent with the dynamic mechanism of precipitation in the TRB, which is mainly controlled by westerly water vapour transport (Dai *et al*, 2007; Qian and Qin, 2008). Predictor sets are varied for each individual station for the SDSM. Generally, the selected predictor number for each station is five and six; air600, shum600, hgt400 and slp are the most selected predictors, and air500 and vwnd600 are followed, while other variables were not selected.

4.2. Statistical downscaling models and GCMs performance on simulating climate variables in the control period

4.2.1 Statistical downscaling models performance based on residual functions

Model biases of air temperature and precipitation for both mean and percentile values are as shown in Figure 2. Model biases at most of the stations were smaller than 1°C for T_{max} and T_{min} during both calibration and validation periods. Only one station showed model bias greater than 1°C for the simulated mean values of T_{max}. Mean value of model biases at all stations was about 0.5°C for the simulated T_{max} and T_{min} during both calibration and validation periods Figure 2 shows that NHMM tends to underestimate three statistics of precipitation, including dsl, wsl, and wpa. Underestimation of dsl and wsl means that the simulated wet or dry spell lengths are shorter than observed ones. However, there is no significant difference between simulated and observed precipitation, with relative errors of three statistics of precipitation smaller than 15%. In general, SDSM and NHMM showed the abilities to reproduce observed air temperature and precipitation, respectively. Figure 2 also shows that there was little difference for models performance during calibration and validation periods. This means that SDSM and NHMM showed abilities on the stability in simulation. When NCEP predictors were replaced by GCM predictors to downscale T_{min} and T_{max} in calibration period, model performance (bias range for mean and percentile values) based on GCMs predictors were similar to that based on NCEP predictors, especially model performance based on ECHAM5 and GFDL21 predictors were almost same as that based on NCEP

predictors. There was no significant difference in downscaling precipitation among three GCMs predictors, while CSIRO30 showed a little better model performance than other GCMs. Model performance also did not change much when NCEP predictors were replaced by GCM predictors in downscaling precipitation, with little changes on the biases of three statistics for precipitation. Therefore, it was concluded that the relationships between predictors and predicands (T_{max}, T_{min} and precipitation) in SDSM and NHMM, which were calibrated based on NCEP predictors, could also be used for these three GCMs predictors.

Table 2. Explained variances of each predictor and SDSM and NHMM selected predictors

Variables	T_{max}		T_{min}		Prec	Variables	T_{max}		T_{min}		Prec
	Explained variance	NM	Explained variance	NM			Explained variance	NM	Explained variance	NM	
air400	0.500		0.797			shum600	0.363	25	0.692	25	★
air500	0.496	2	0.778			uwnd400	0.032		0.083		
air600	0.513	23	0.825	25	★	uwnd500	0.036		0.064		
hgt400	0.303	25	0.435	25	★	uwnd600	0.025		0.055	2	
hgt500	0.124		0.077			vwnd400	0.015		0.030		
hgt600	0.051		0.015			vwnd500	0.016		0.017		
shum400	0.273		0.583		★	vwnd600	0.047	2	0.120	2	★
shum500	0.279		0.565		★	slp	0.440	25	0.650	25	★

Note: "NM" is the number of stations selected as predictor by SDSM; ★ indicates selected by NHMM. "T_{max}", "T_{min}" and "Prec" describe maximum air temperature, minimum air temperature, and precipitation, respectively.

The SDSM could capture distribution characteristics of T_{min}. Although it lost some abilities in simulating extreme large values for T_{max}; SDSM was able to capture trends of these values. Figure 3 shows the distribution of wpa including extreme values could be captured by NHMM. Model performance on simulating the distribution of dsl was better than that for wsl in the TRB. The NHMM could capture distribution of dsl for values smaller than 95 percentile, while it tends to underestimate those values greater than 95 percentile. The ability of NHMM for capturing distribution of wsl was not as good as that for wpa and dsl. It was only able to capture distribution of wsl for values smaller than median value and tend to underestimate greater values than this value.

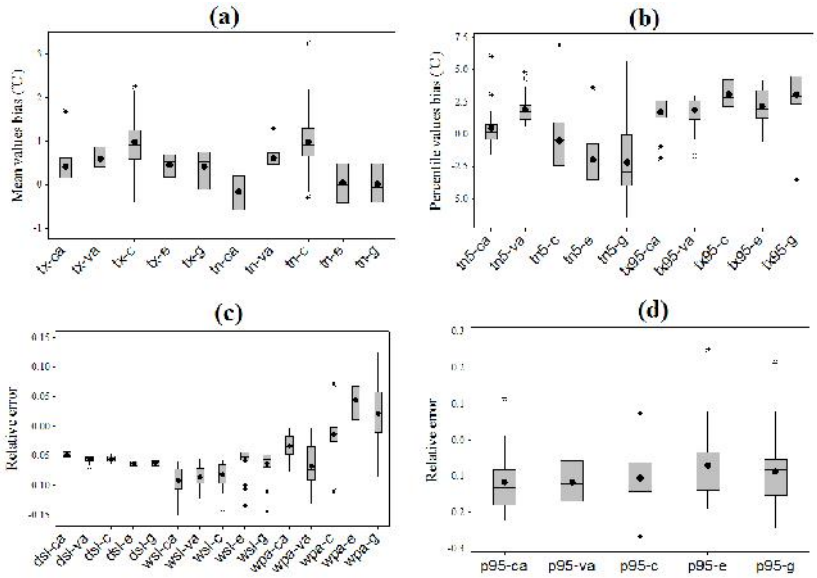

Fig. 2. Model bias for precipitation and temperature ((a), (b) describe bias of mean and percentile values for temperature, respectively; (c), (d) are the same as that but for precipitation. "–ca" and "–va" mean calibration and validation respectively; "–c", "–e", "–g" mean predictors of CSIRO30, ECHAM5 and GFDL21 respectively; "tx", "tn" mean maximum and minimum air temperatures respectively; "tn5", "tx95", "p95" mean 5th value of minimum air temperature, 95th value of maximum air temperature and wpa, respectively. The following are the same.)

4.2.2. Statistical downscaling models performance based on NRMSE and EOF

The spatial and temporal variability of monthly temperature was characterized by EOF analysis. The first two leading modes of each EOF for all three variables are shown in Table 3. The first mode of EOF for observed that T_{max} and T_{min} accounts for 75.6 and 62.2% of the total variance in calibration period. It is well simulated by SDSM, with relative error of these variances to only about 10%. These characteristics are also captured by the model in validation period. The first two EOFs for precipitation explain 40.5 and 11.2% in calibration period, respectively, of the total variance. NHMM is able to simulate this spatial and temporal variability with explained variance values of 49 and 17.4%, respectively. This suggests that the physical processes dominating temperature and precipitation variability are captured by two statistical downscaling models. However, modeled precipitation, T_{max} and T_{min} generally have larger explained variance for the first EOF than observed ones in both calibration and validation periods. This implies that the model is tends to simulate a more uniform spatial distribution.

The NRMSE measures the goodness-fit between observed and modeled time series, where a smaller value indicates a better fit between observed and simulated values. NRMSE values for maximum air temperature, minimum air temperature, and precipitation are as shown in Figure 4. In generally, the NRMSE values for T_{max} and T_{min} ranges from 0.2 to 0.5 in calibration period. There is little difference for model performance on T_{max} and T_{min} in both calibration and validation periods. The model performance on simulating precipitation was not so well as that on T_{max} and T_{min}, with NRMSE values larger than 0.6 for most of the stations. When NCEP predictors were replaced by GCMs predictors, it could be found from the figure that there is little difference between GCMs and NCEP predictors when simulating precipitation, which means that the NHMM has model stabilities on capturing precipitation time series. It also shows some consistency of model performance between GCMs and NCEP predictors when simulating T_{max} and T_{min}, but the difference is a little larger than that for precipitation. It is concluded that there is a better model stability for NHMM than SDSM when simulating precipitation and air temperature, respectively.

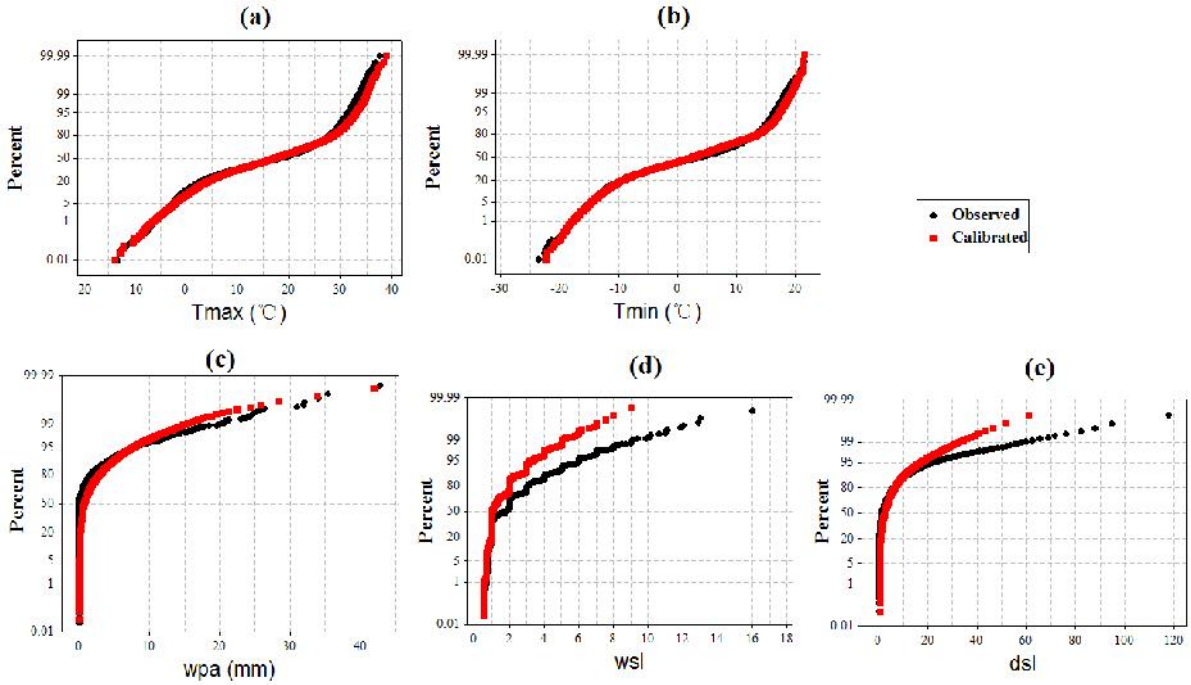

Fig. 3. Cumulative distribution functions for observed and simulated time series of precipitation and temperature

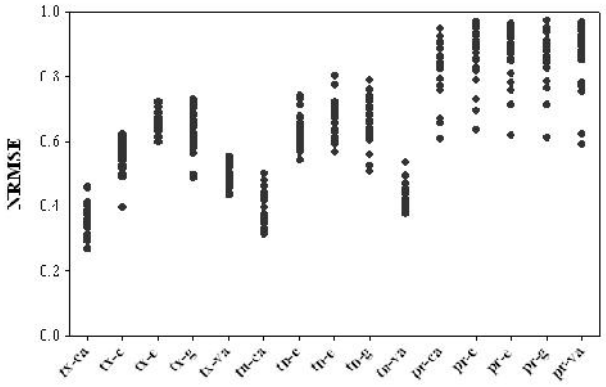

Fig. 4. NRMSE values for maximum air temperature, minimum air temperature and precipitation ("pr" means precipitation.)

Table 3. The first two leading modes of each EOF for precipitation and temperature

		Precipitation		Minimum air temperature		Maximum air temperature	
		EOF1	EOF2	EOF1	EOF2	EOF1	EOF2
Calibration	Observed	0.405	0.112	0.622	0.109	0.756	0.096
	Simulated	0.490	0.174	0.719	0.129	0.823	0.127
	CSIRO30	0.443	0.115	0.811	0.075	0.734	0.111
	ECHAM5	0.396	0.152	0.762	0.144	0.819	0.082
	GFDL21	0.430	0.144	0.751	0.125	0.730	0.165
Validation	Observed	0.282	0.143	0.593	0.123	0.735	0.100
	Simulated	0.346	0.129	0.724	0.103	0.822	0.129

Note: "EOF1" and "EOF2" mean the first and second modes of EOF respectively.

4.2.3. Statistical downscaling models performance based on PDFs

Results of two PDFs based skill scores S_{score} and BS for air temperature and precipitation are as shown in Figure 5. As shown in figure, NHMM showed better skills on simulating dsl than that on wsl in the TRB, with greater S_{score} values and less BS values for dsl than that for wsl. This might be as a result of the fact that values of dsl are much higher than wsl in the arid region, and the statistical downscaling model shows better ability on greater values of time series. That is consistent with the results from Liu *et al*. (2010). Besides, NHMM also had skills in modeling wpa with S_{score} being greater than 0.9 in both calibration and validation periods. SDSM also showed good skills on downscaling T_{max} and T_{min} over the TRB, with S_{score} values at most stations greater than 0.8 in both calibration and validation periods. It was similar to the results of residual functions and correlation analysis that, there was little different for model performance of NHMM and SDSM between calibration and validation periods, which showed that both models had a better model stability in downscaling precipitation and air temperature over the TRB. When NCEP predictors were replaced by GCM predictors in downscaling precipitation, T_{max} and T_{min} were replaced in calibration period, there were no significant differences for model performance with little changes on the values of S_{score} and BS, which indicated that the relationships between predictors and predicands calibrated by NCEP predictors could also be used for these three GCMs predictors.

4.3. Comparison of climate change scenarios generated by raw GCMs outputs and statistical downscaling

Performance of NHMM and SDSM showed that the two models had abilities in downscaling daily precipitation, T_{max} and T_{min} in the TRB. It also showed that the relationship between predictors and predicands calibrated by NCEP predictors could be used for three GCMs predictors to generate regional climate scenarios. Therefore, models calibrated and validated by NCEP predictors were driven by GCMs predictors in the control period (1981 to 2000) and future periods

(2046 to 2065 and 2081 to 2100) to project local precipitation, T_{max} and T_{min} scenarios, which were estimated by the difference of downscaled variables between those two periods under the IPCC SRES A1B, A2, and B1 emission scenarios. Besides this fact, climate scenarios based on raw GCMs outputs were also generated by the difference of GCMs variables (precipitation, T_{max} and T_{min}) between the control period and future periods. Finally, climate change scenarios generated by raw GCMs outputs and statistical downscaling were compared with each other. Although, the local variables from GCMs raw outputs were not well simulated as that from statistical downscaling models;, climate scenarios based on raw GCMs outputs were also analyzed in this study. Owing to the fact that there are many uncertainties in the projection of future climate change scenarios, could be as optional scenarios for related studies.

Figure 7 described the changes for mean values of monthly, seasonal and annual precipitation, T_{max} and T_{min} projected by raw GCMs outputs and statistical downscaling models. Changes of annual precipitation generated by the statistical downscaling model were not obvious, with values smaller than 15% under all combined scenarios (Figure 7a and b). The magnitude of changes for annual precipitation in the period of 2081 to 2100 was smaller than that in the period of 2046 to 2065. In total, the magnitude of changes for annual precipitation projected by raw GCMs outputs was greater than that generated by using the statistical downscaling model. In both periods, CSIRO30 tends to project increasing trends for annual precipitation, while ECHAM5 and GFDL21 projected decreasing trends under three combined scenarios. As for monthly precipitation, it showed much greater changes than annual precipitation based on the statistical downscaling model, with values ranging from -80 to 60% approximately. Based on statistical downscaling results, it should be pointed out that monthly precipitation showed increasing trends only in March and July under most of combined scenarios for the period of 2046 to 2065. It exhibited increasing trends in April, May, October and November, while decreasing trends in July and August under all combined scenarios for the period of 2081 to 2100. Changes of monthly precipitation in period of 2081 to 2100 were also smaller than that in the period of 2046 to 2065. The most significant difference for changes of monthly precipitation generated between the statistical downscaling model and raw GCMs outputs appeared from July to October, in which changes of monthly precipitation projected by raw GCMs outputs were much greater than that generated by using the statistical downscaling model. In both periods, the most significant decreasing trends for monthly precipitation appeared from July to October, in which precipitation projected by ECHAM5 and GFDL21 decreased by 20 to 80%.

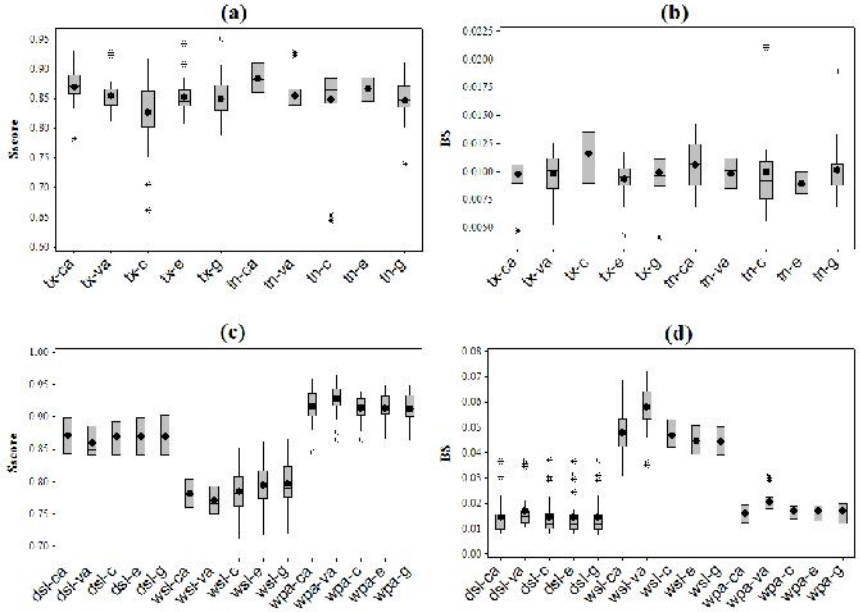

Fig. 5. Boxplots for precipitation, maximum and minimum air temperatures based on skill scores ((a), (b) describe air temperature; (c), (d) describe precipitation.)

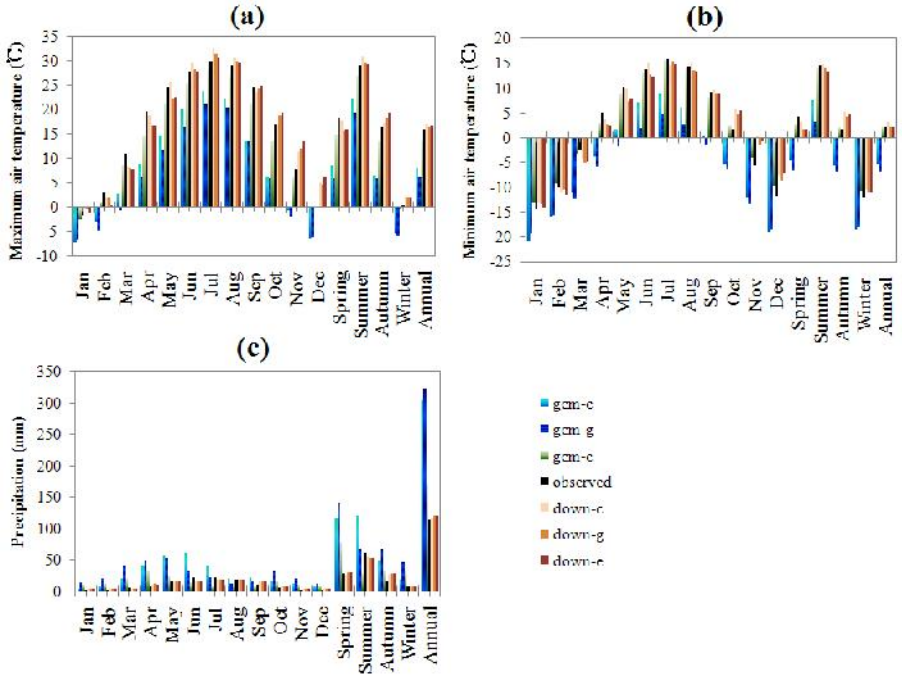

Fig. 6. Mean values of climate variables simulated by GCMs and statistical downscaling models

Figure 7c and d shows that T_{max} exhibited increasing trends under all combined scenarios. Annual T_{max} in the period of 2046 to 2065 generated by the statistical downscaling model increased by 1.3 to 4.0°C under different combined scenarios. Increasing trend in the period of 2081 to 2100 was greater than that in the period of 2046 to 2065, with the greatest change greater than 7.0°C. The difference between changes of annual T_{max} projected by the statistical downscaling model and raw GCMs outputs was not as obvious as that for annual precipitation. Generally, for both the statistical downscaling model and raw GCMs outputs, the magnitude of increasing trend was the greatest under SRES A2 scenario and was the smallest under B1 scenario, with A1B scenario in–between. In both periods, the increasing trend of T_{max} projected by CSIRO30 was smaller than that by ECHAM5 and GFDL21. The difference between changes of monthly and annual T_{max} was not as obvious as that for precipitation. The magnitude of increasing trend in the period of 2081 to 2100 was the greater than that in the period of 2046 to 2065. Similar to annual T_{max}, the magnitude of increasing trends projected by both the statistical downscaling model and raw GCMs outputs was the greatest under SRES A2 scenario and the smallest under B1 scenario, with A1B scenario in–between. The magnitude of increasing trend projected by using CSIRO30 was approximately from 0.5 to 2.1°C in the period of 2046 to 2065 and from 1.1 to 3.3°C in the period of 2081 to 2100, which were smaller than that projected by ECHAM5 and GFDL21.

Figure 7e and f shows that T_{min} projected by raw GCMs outputs exhibited increasing trends under all combined scenarios, while that generated by the statistical downscaling model also exhibited the same trends under most of combined scenarios, except CSIRO30 related scenarios and GFDL21–B1 combined scenarios. Changes of annual T_{min} generated by the statistical downscaling model ranged from –0.4 to 3.3°C in the period of 2046 to 2065 and from –0.9 to 5.9°C in the period of 2081 to 2100. The magnitude of increasing trends for annual T_{min} projected by raw GCMs outputs was greater than 1°C in both periods. The difference between changes of monthly and annual T_{min} was not as obvious as that for precipitation. It was worthy to point out that winter T_{min} showed increasing trends under all combined scenarios for both periods.

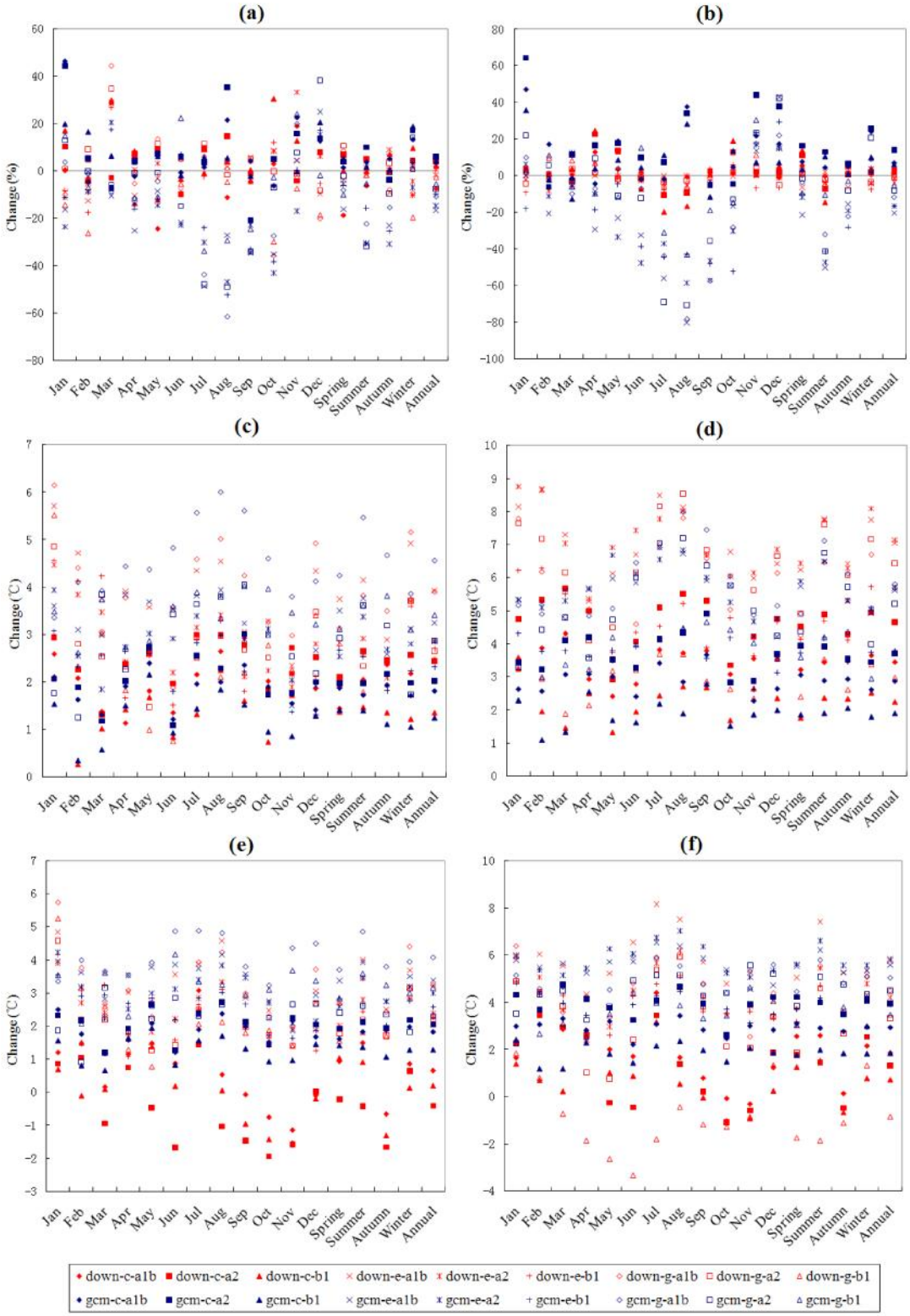

Fig. 7. Changes of mean values for precipitation, maximum and minimum air temperatures projected by raw GCMs outputs and statistical downscaling models ((a), (c), (e) describe changes of precipitation, maximum and minimum air temperatures in the period of 2046–2065 respectively; (b), (d), (f) were changes of precipitation, maximum and minimum air temperatures in the period of 2081–2100 respectively.)

Figure 8 described changes for standard deviation of daily precipitation, T_{max} and T_{min} time series projected by raw GCMs outputs and statistical downscaling models. In both periods of 2046 to 2065 and 2081 to 2100, standard deviation of daily T_{max} projected by raw GCMs outputs showed increasing trends under all combined scenarios, while that generated by the statistical downscaling model exhibited decreasing trends under most of combined scenarios. All of these changes were smaller than 10%. The magnitude of increasing trends under CSIRO30–A2 and EHAM5–A2 combined scenarios was greater than the other combined scenarios in both periods. In both periods, standard deviation of daily T_{min} projected by raw GCMs outputs and the statistical downscaling model exhibited increasing trends under some combined scenarios, while it showed opposite trends under other combined scenarios. The magnitude of these trends generated by the statistical downscaling model was greater than that projected by raw GCMs outputs. In the period of 2081 to 2100, it was worthy to point out that standard deviation of daily T_{min} generated by the statistical downscaling model exhibited increasing trends under ECHAM5 related scenarios, while it showed decreasing trends under CSIRO30 and GFDL21 related scenarios. In both periods, the magnitude of changes projected by raw GCMs outputs was greater than that generated by the statistical downscaling model under most of the combined scenarios. As for standard deviation changes for daily precipitation, the magnitude of that projected by raw GCMs outputs was greater than that generated by the statistical downscaling model under most of combined scenarios in both periods, with values smaller than 10%.

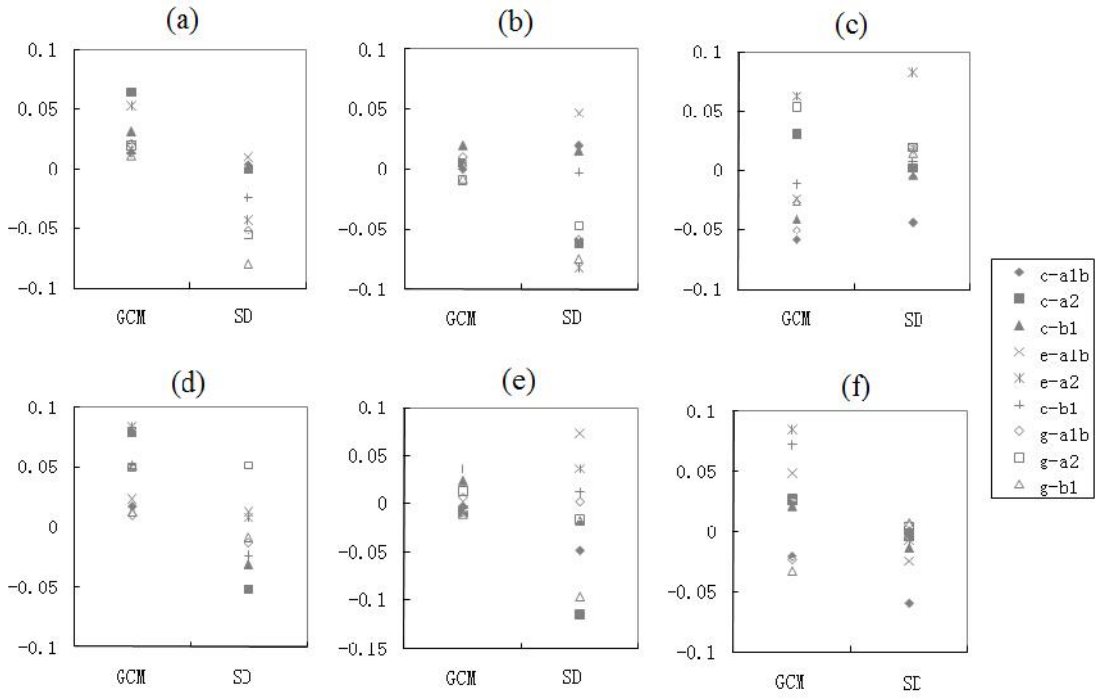

Fig. 8. Changes of standard deviation for precipitation, maximum and minimum air temperatures projected by raw GCMs outputs and statistical downscaling models ((a), (b), (c) describe changes of maximum and minimum air temperature, and precipitation in the period of 2046–2065 respectively; (d), (e), (f) were changes of maximum and minimum air temperature, and precipitation in the period of 2081–2100 respectively.)

5. Conclusion and Discussions

Precipitation, T_{max} and T_{min} scenarios generated by raw GCMs outputs and statistical downscaling models are compared with each other in an arid region, TRB.

Two statistical downscaling models, NHMM and SDSM, showed abilities to reproduce observed precipitation, T_{max} and T_{min} with NCEP reanalysis predictors in the TRB. Besides that, when NCEP predictors were replaced by GCMs predictors, model performance changed a little, including

model biases, spatial and inter–annual correlation coefficients, PDFs based skill scores S_{score} and BS. It shows that the relationships between predictors and predicands calibrated by NCEP predictors could also be used for GCMs predictors to generate local climate change scenarios.

Based on the nine combined scenarios, including three GCMs and three emission scenarios, local annual precipitation projected by the statistical downscaling model did not change obviously in periods of 2046 to 2065 and 2081 to 2100, with values smaller than 15%. Annual precipitation showed increasing trends under CSIRO30 related scenarios, while it exhibited decreasing trends under ECHAM5 and GFDL21 related scenarios. The magnitude of changes for annual precipitation in the period of 2081 to 2100 was smaller than that in the period of 2046 to 2065. The most significant difference for changes of monthly precipitation generated between the statistical downscaling model and raw GCMs outputs appeared from July to October, in which changes of monthly precipitation projected by raw GCMs outputs were much greater than that generated by the statistical downscaling model. T_{max} showed increasing trends under all combined scenarios and T_{min} also exhibited increasing trends under most of the combined scenarios. In total, the magnitude of these increasing trends projected by both statistical downscaling model and raw GCMs outputs was the greatest under SRES A2 scenario and the smallest under B1 scenario, with A1B scenario in–between. The magnitude of these increasing trends in the period of 2081 to 2100 was greater than that in the period of 2046 to 2065. The difference between changes of annual T_{max} projected by statistical downscaling model and raw GCMs outputs was not as obvious as that for annual precipitation. Standard deviation of daily T_{max} projected by raw GCMs outputs showed increasing trends under all combined scenarios, while that generated by the statistical downscaling model exhibited decreasing trends under most of the combined scenarios. The magnitude of standard deviation changes for daily precipitation projected by raw GCMs outputs was greater than that generated by the statistical downscaling model under most of the combined scenarios in both periods.

Acknowledgements

This study is financed by the “Hundred Talents Program”, of the Chinese Academy of Science. Thanks also should be given to the Chinese Scholarship Council-CSIRO Joint Supervision of Chinese PhD student project, which makes it possible for the first author to study in Australia for one year. The kind assistance from the colleagues in CSIRO, Australia is also greatly appreciated.

References

Bates, B.C., S.P. Charles, J.P. Hughes, 1998. Stochastic downscaling of numerical climate model simulations. Eviron. Modell. Softw. 13, 325–331.

Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. Mon. Wea. Rev. 78, 1–3.

Busuioc, A., D. Chen, C. Hellström, 2001. Performance of statistical downscaling models in GCM validation and regional climate change estimates, Application for Swedish precipitation. Int. J. Climatol. 21, 557–578.

Charles, S.P., B.C. Bates, J.P. Hughes, 1999. A spatiotemporal model for downscaling precipitation occurrence and amounts. J. Geophys. Res. 104(D24), 31657–31669.

Christensen, J.H., B. Hewitson, A. Busuioc, A. Chen, X. Gao, I. Held, R. Jones, R.K. Kolli, W.T. Kwon, R. Laprise, V. Magaña Rueda, L. Mearns, C.G. Menéndez, J. Räisänen, A. Rinke, A. Sarr, P. Whetton, 2007. Regional Climate Projections. In: Climate Change 2007, The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Cubasch, U., H. von Storch, J. Waszkewitz, E. Zorita, 1996. Estimates of climate change in Southern Europe derived from dynamical climate model output. Climate Res. 7, 129–149.

Dai, X.G., Li, W.H., Ma, Z.G., P. Wang, 2007. Water–vapor source shift of Xinjiang region during the recent twenty years. Prog. Nat. Sci. 17(5), 569–575.

Diaz–Nieto, J., R.L. Wilby, 2005. A comparison of statistical downscaling and climate change factor methods, impacts on low flows in the river Thames, United Kingdom. Climatic Change 69, 245–268.

Dibike, Y.B., P. Coulibaly, 2005. Hydrologic impact of climate change in the Saguenay watershed, comparison of downscaling methods and hydrologic models. J. Hydrol. 307, 145–163.

Dibike, Y.B., P. Coulibaly, 2006. Temporal neural networks for downscaling climate variability and extremes. Neural Networks 19, 135–144.

Forney, G.D. Jr., 1978. The Viterbi algorithm. P. IEEE 61, 268–278.

Gachon, P., Y. Dibike, 2007. Temperature change signals in northern Canada, convergence of statistical downscaling results using two driving GCMs. Int. J. Climatol. 27, 1623–1641.

Giorgi, F., Coauthors, 2001. Regional climate information Evaluation and projections. In: Climate Change 2001, The Scientific Basis. J. T. Houghton et al. (Eds.), Cambridge University Press, 583–638.

Harvey, L.D.D., T.M.L. Wigley, 2003. Charactering and comparing control–run variability of eight coupled AOGCMs and of observations Part 1: temperature.

Climate dyn. 21, 619–646.
Huges, J.P., P. Guttopp, 1994. Incorporating spatial dependence and atmospheric data in a model of precipitation. J. Appl. Meteorol. 33(12), 1503–1515.
Hughes, J.P., P. Guttorp, S.P. Charles, 1999. A non–homogeneous hidden Markov model for precipitation occurrence. Appl. Stat.-J. Roy. St. C. 48(1), 15–30.
IPCC, 2007, Summary for Policymakers. In: Climate Change 2007, The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. QIn: M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. pp 18.
Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. GandIn: M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, D. Joseph, 1996. The NCEP/NCAR 40–Year Reanalysis Project. B. Am. Meteorol. Soc. 77(3), 437–471.
Karl, T.R., W.C. Wang, M.E. Schlesinger, R.W. Knight, D. Portman, 1990. A method of relating general circulation model simulated climate to observed local climate. Part I, Seasonal statistics. J. Climate 3, 1053–1079.
Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, V. Kanamitsu, M. Kousky, H. van den Dool, R. Jenne, M. Fiorino, 2001. The NCEP/NCAR 50–year reanalysis. B. Am. Meteorol. Soc. 82, 247–267.
Liu, Z.F., Z.X. Xu, S.P. Charles, G.B. Fu, L. Liu, 2010. Evaluation of two statistical downscaling models for daily precipitation over an arid basin in China. Int. J. Climatol. (in press)
McAvaney, B., C. Covey, S. Joussaume, V. Kattsov, A. Kitoh, W. Ogana, A. Pitman, A. Weaver, R. Wood, Z.C. Zhao, 2001. Climate change 2001, the scientific basis, chap. 8, model evaluation. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC. University Press, Cambridge
Mu, Q.Z., C.S. Jackson, P.L. Stoffa, 2004. A multivariate empirical–orthogonal–function–based measure of climate model performance. J. geophys. Res. 109, D15101.
Murphy, J.M., 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. J. Climate 12, 2256–2284.
Perkins, S.E., A.J. Pitman, N.J. Holbrook, J. Mcaneney, 2007. Evaluation of the AR4 climate models' simulated daily maximum air temperature, minimum air temperature, and precipitation over Australia using probability density functions. J. Climate 20, 4356–4376.
Prudhomme, C., H. Davies, 2009. Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1, baseline climate. Climate Change 93, 177–195. DOI, 10.1007/s10584–008–9464–3.
Qian, W.H., A. Qin, 2008. Precipitation division and climate shift in China from 1960 to 2000. Theor. Appl. Climatol. 93, 1–17.
Robertson, A.W., S. Kershner, P. Smyth, 2004. Downscaling of Daily Rainfall Occurrence over Northeast Brazil Using a Hidden Markov Model. J. Climate 17, 4407–4424.
Schubert, S., A.H. Sellers, 1997. A statistical model to downscale local daily temperature extremes from synoptic–scale atmospheric circulation patterns in the Australian region. Clim. Dynam. 13, 223–234.
Schubert, S., 1998. Downscaling local extreme temperature changes in south–eastern Australia from the CSIRO Mark2 GCM. Int. J. Climatol. 18, 1419–1438.
Timbal, B., A. Dufour, B. McAvaney, 2003. An estimate of future climate change for western France using a statistical downscaling technique. Clim. Dynam. 20, 807–823. DOI 10.1007/s00382–002–0298–9.
Wetterhall, F., A. Bárdossy, D.L. Chen, S. HalldIn: C.Y. Xu, 2006. Daily precipitation–downscaling techniques in three Chinese regions. Water Resour. Res. 42, W11423. DOI,10.1029/2005WR004573.
Wilby, R.L., S.P. Charles, E. Zorita, B. Timbal, P. Whetton, L.O. Mearns, 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change (IPCC), prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA). <http,//ipccddc.cru.uea.ac.uk/guidelines/StatDown_ Guide.pdf>.
Wilby, R.L., O.J. Tomlinson, C.W. Dawson, 2003. Multi–site simulation of precipitation by conditional resampling. Climate Res. 23, 183–194.
Wilby, R.L., P.G. Whitehead, A.J. Wade, D. Butterfield, R.J. Davis, G. Watts, 2006. Integrated modelling of climate change impacts on water resources and quality in a lowland catchment_River Kennet, UK. J. Hydrol. 330, 204–220.
Wilby, R.L., T.M.L. Wigley, 2000. Precipitation predictors for downscaling, Observed and general circulation model relationships. Int. J. Climatol. 20, 641–661.
Wilby, R.L., T.M.L. Wigley, 1997. Downscaling general circulation model output, A review of methods and limitations. Prog. Phys. Geog. 214, 530–548.
Wilby, R.L., C.W. Dawson, E.M. Barrow, 2002. SDSM—A decision support tool for the assessment of regional climate change impacts. Eviron. Modell. Softw. 17, 147–159.
Wilby, R.L., L.E. Hay, W.J.Jr. Gutowski, R.W. Arritt, E.S. Takle, Z. Pan, G.H. Leavesley, M.P. Clark, 2000. Hydrological Responses to Dynamically and Statistically Downscaled Climate Model Output. Geophys. Res. Lett. 27, 1199–1202.
Wilks, D.S., R.L. Wilby, 1999. The weather generation game, a review of stochastic weather models. Prog. Phys. Geog. 23, 329–357.
Wise, E.K., 2009. Climate–based sensitivity of air quality to climate change scenarios for the southwestern United States. Int. J. Climatol. 29, 87–97.
Wood, A.W., L.R. Leung, V. Sridhar, D.P. Lettenmaier, 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62, 189–216.
Xu, Z.X., Z.F. Liu, G.B. Fu, Y.N. Chen, 2010. Hydro-Climate Trends of the Tarim River Basin for the Last 50 Years. J. Arid Environ. 74, 256–267.