Dear Dr. Bettina Schaefli,

First of all, we would like to thank you for your decision regarding the revision of our manuscript. We would like to thank the anomalous reviewer and Dr. Markus Muerth for their instructive and helpful comments. These comments will be very valuable to improve the manuscript.

We have addressed all the comments and incorporated most of them into the revised manuscript. The point to point reply is listed in *"Reply to Reviewers and Editor.pdf*" and the corresponding changes are shown in "*Revised manuscript.pdf*".

If you have any questions, please don't hesitate to contact us.

Best regards,

Jing Yang

13.04.2015

Reviewer #1

1. Since the raw RCM simulation is greatly biases, it is necessary to give some explanation on the data reliability.

Our reply: *Firstly, It has shown that GCM or RCM outputs are generally biased (Ahmed et al., 2013;Teutschbein and Seibert, 2012;Mehrotra and Sharma, 2012), which demonstrates the need for bias correction before their use in regional impact studies. Secondly, though the biases in the raw RCM simulation are large, the RCM outputs show reasonable simulation of temperature and precipitation over most parts of China especially when compared with its driving GCM BCC_CSM1.1, which is validated by Gao et al (2013) using the observational dataset (CN05.1).*

We have incorporated the explanation in the revised manuscript.

Lines 121 ~ 132:

GCM or RCM outputs are generally biased (Ahmed et al., 2013; Teutschbein and Seibert, 2012; Mehrotra and Sharma, 2012), and there is a need to correct these outputs before used for regional impact studies. The RCM outputs used in this study are based on the work done by Gao et al. (2013) where the RCM model (RegCM, Giorgi and Mearns, 1999) was driven by a global climate model BCC_CSM1.1 (Beijing Climate Center Climate System Model; Wu et al., 2013; Xin et al., 2013) at a horizontal resolution of 50 km over China.

The RCM outputs were validated with the observational dataset (CN05.1) over China for the period from 1961 to 2005. The RCM outputs show reasonable simulation of temperature and precipitation in most parts of China except some regions where our study area is located (for more details refer to Gao et al., 2013).

2. P12666 Line 16: what do you mean by "bias correction methods were conducted on a monthly basis", since the inputs required for SWAT is normally daily climate data.

Our reply: *The time step of climate variables is daily. We altered the sentence into "bias correction methods were conducted on a daily basis" in the revised manuscript.*

Lines 189 ~ 190:

All these bias correction methods were conducted on a daily basis from 1975 to 2005.

3. Table 6 could be improved if you provide the MAE (mean absolute error) or RMSE value, so the readers could quickly acquire the relative errors that are still existed in the corrected meteorological data and can compare with other studies easily.

Our reply: *That is a good point. We added the MAE values in Table 6 and Table 3. The equation of MAE has also been added in the revised manuscript (Eq. 18).*

In addition, the corresponding analysis has been replaced.

Equation 18:

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_i^{obs} - Y_i^{sim}\right|}{n} \quad (18)$$

Lines 431 ~ 438:

For precipitation, the performance of the raw RCM simulated precipitation is very poor with NS = -6.78, P_{BIAS} = 293.28% and MAE = 65.40 mm for monthly data, and the improvements of correction are obvious. The "P_{BIAS}"s of the corrected precipitation are within $\pm 7\ \%$ and "NS"s approach 0.64. It is worth noting that LS and LOCI methods perform better than PT and QM methods in terms of time series performances. For temperature, although the raw RCM simulation obtains an acceptable NS value (0.84), it overestimates the observation with P_{BIAS} = 15.78% and MAE = 4.31 ℃.

4. P12663 line 24: The "precipitation falls as rain from May to September", therefore, the hydrological regime is different among seasons. It is advisable to alter Figure 5 and Figure 6 to demonstrate the differences of performances.

Our reply: *In order to investigate the performances of bias correction methods for different hydrological seasons, we divided the streamflow into two different periods according to the hydrograph (Fig. 3): wet period is from April to September and dry period is from October to March of next year. Streamflow statistics for each simulation scenarios are shown in Fig. S1 and Fig. S2.*

In Fig. S1, except for magnitudes, the results are similar for both wet and dry period. Therefore, there is no need to demonstrate the streamflow distribution in wet period and dry period separately, as the aim of this study is comparing the performances of bias correction methods. In Fig. S2, the exceedance probability curves can represent streamflow data for each frequency, therefore, there is no need to display separately. The similar performance of bias correction methods for the wet and dry periods in term of simulated streamflow confirms that evaluation of bias *correction is robust and can provide useful information for both wet and dry climate.*

We incorporated the following discussion in the revised manuscript.

Lines 454 ~ 460:

To investigate the performances of bias correction methods for different hydrological seasons, we divided the streamflow into two different periods according to the hydrograph (Fig. 3): wet period is from April to September and dry period is from October to March of next year. It indicates that the performances of bias correction methods are, except for magnitudes, similar for both wet and dry period (not shown), which demonstrates that the evaluation is robust and can provide useful information for both dry and wet seasons.

5. Present some discussion on the differences of bias correction method applied in the arid area and humid area.

Our reply: *We agree and added some discussions on the performances of bias correction methods based on previous studies.*

Lines 372 ~ 378:

These results are consistent with previous studies (Themeßl et al., 2011, 2012; Wilcke et al., 2013; Graham et al., 2007), but are different from the research by Piani et al. (2010) who found that performance of DM method is unexpectedly well for the humid Europe region. This discrepancy can be partly attributed to the precipitation regime for different regions since better fit of the assumed distribution leads to better performance of DM.

Lines 515 ~518:

Results show slightly better performance of PT and QM methods than LOCI and DM in predicting extreme flood and low flow, which is consistent with previous studies in North America and Europe (e.g., Chen et al., 2013a; Teutschbein and Seibert, 2012).

Technical corrections:

6. Some expressions should stay consistent throughout the paper, e.g., P12667 line18 Capital the first letter “Transformation”. Also, some items are confusing, e.g., RCM simulations, RCM outputs, climate variables from the RCMs, RCM output. I think they all indicate the RCM simulated climate variables, why not use one expression?

Our reply: *We corrected them in the revised version. And we carefully checked the expressions, which is not included here since many corrections have been made.*

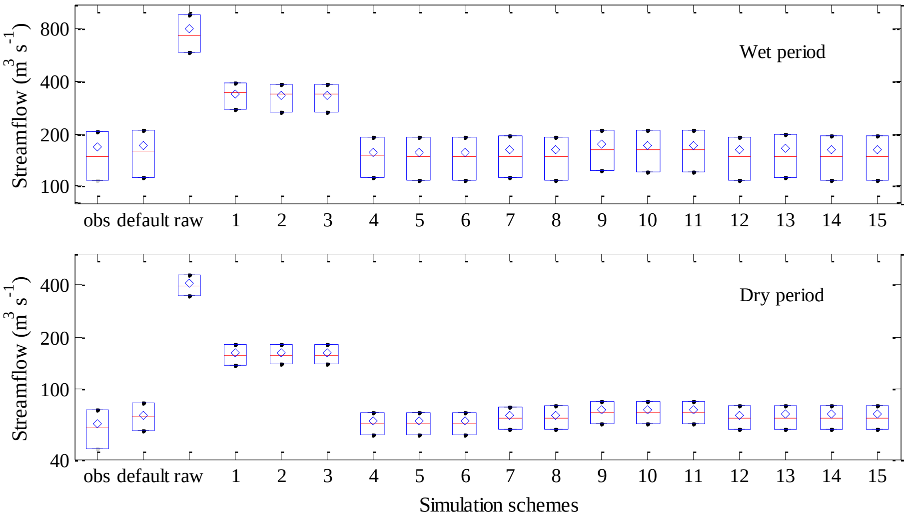

Fig. S1 Same as Fig. 6 but for the wet period and dry period.

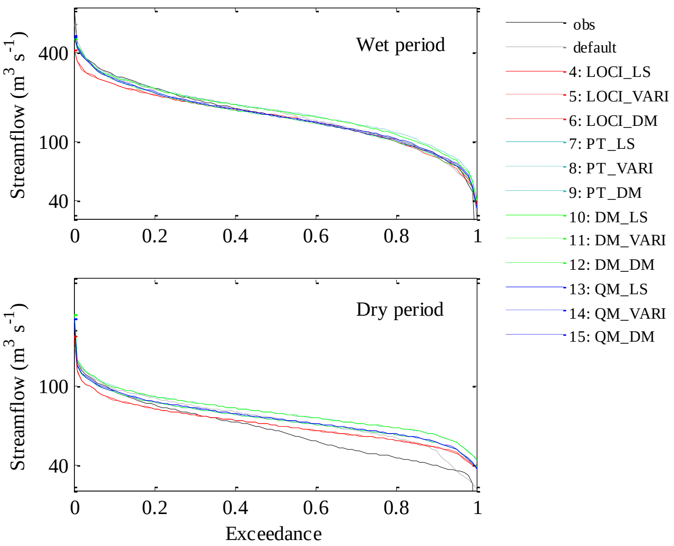

Fig. S2 Same as Fig. 7 but for the wet period and dry period.

Reviewer #2 (Dr. Markus Muerth)

The manuscript of Fang et al. on 'Comparing bias correction methods in downscaling meteorological variables' is generally well structured and well written besides some minor typos. Yet, there are three major issues with this paper that have to be clarified in my point of view before it can be accepted.

1. The first problem I have is that I don't understand which data is compared and which is used for modeling. Meteorological variables are corrected based on data of one climate station. As we can see, precipitation is overestimated by the RCM-GCM model chain. Yet, for the whole catchment, this is not true looking at the maps of Gao et al 2013. In some parts, rainfall seems to be well simulated by the RCM. So how does the station relate to the whole basin? Do you compare RCM simulated data for the whole basin with one station or just the RCM box at the station? So, what does "raw" precipitation mean? I can hardly imagine that a correction factor found for one station evens out the heterogeneity of biases in the basin. The authors should be clearer about the relevance of the bias correction in the light of this, especially as the accompanying paper is not available yet.

Our reply: *1) We added a flow chart of comparison procedure (Figure 2) and its description (Lines 135 ~ 143) to better understand the comparison work. Data comparison (as shown as dash lined boxes in Figure 2) includes the comparison of "corrected meteorological data at the station scale" and observed meteorological data, and the comparison of simulated streamflow with corrected meteorological data, simulated streamflow with observed meteorological data, and observed streamflow data. The data used for modeling include "corrected meteorological data at the station scale" and observed meteorological data.*

2) Meteorological data were compared with observed data at two stations (i.e., Bayanbulak station and Baluntai station). As indicated in Lines 352 ~ 354, we only show the results for Bayanbulak as the results for Baluntai Station are very similar to those for Bayanbulak.

3) In Gao et al. (2013), precipitation was well simulated in most parts of China except some regions where our study area is located. As shown in Figure 2 of Gao et al. (2013), the precipitation is overestimated for our study area (located in the *northwest China, 82°58′ ~ 86°05′E, 42°14′ ~ 43°21′N).*

4) As indicated in the newly added Figure 2, station based meteorological data were upscaled to areal meteorological data with precipitation and temperature lapse rates before hydrological modeling.

5) As indicated in the newly added Figure 2, RCM meteorological data were compared with the observations at the station scale through downscaling.

6) The "raw" precipitation means the RCM simulated grid precipitation without any correction.

7) The accompany paper has been accepted and is online now (http://link.springer.com/article/10.1007/s12665-015-4244-7), and we updated the citation accordingly.

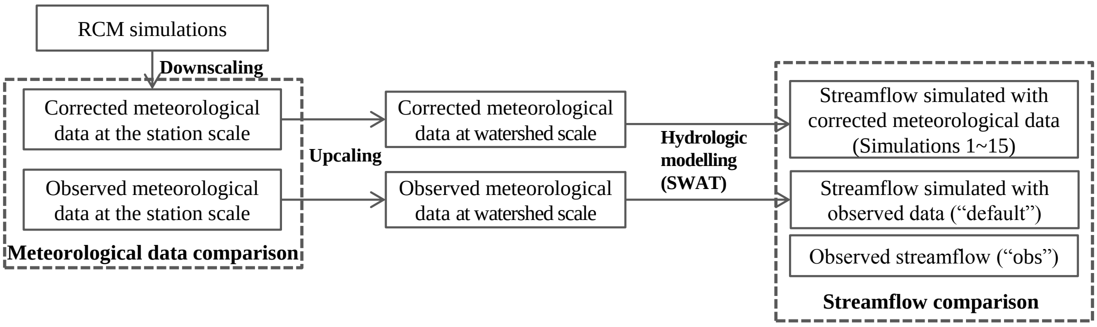

Figure 2 Flow chart of comparison procedure.

Lines 135 ~ 143:

Figure 2 shows the flow chart of the comparison procedure. First, grid based RCM simulation was downscaled to station scale using bias correction methods, and then the corrected meteorological data were compared to the observation at these two stations and to each other ("Meteorological data comparison" in Fig. 2). These station based meteorological data were then upscaled to watershed scale with the precipitation and temperature lapse rates before they were used to drive the hydrological model (SWAT). Finally, the simulated streamflow driven by the corrected and observed meteorological data were compared to observed streamflow and to each other ("Streamflow comparison" in Fig. 2).

Lines 352 ~ 354:

2) Similar results were obtained for minimum temperature and maximum temperature, and for Bayanbulak and Baluntai. Therefore we only listed and discussed results for Bayanbulak, and maximum temperature.

2. The second issue is the weak performance of the LS method for precipitation on

runoff. How large are drizzle values? Are they so low, you could easily cut them off because these (daily) values would never affect hydrological processes? Why don’t you connect them to a humidity threshold? Does SWAT not compute evapotranspiration if there is very little rainfall? Your graphs look like there nearly no evapotranspiration when using LS, meaning runoff seems to equal rainfall

Our reply: *1) In this study, after LOCI method was applied to correct the wet day probability (or drizzle effect), drizzle values are precipitation values less than the corresponding threshold for month m ($P_{thres,\,m}$ in Eq. 7) and vary from 0.01 to 5.6 mm. As they are not always low, this cutoff will influences simulated hydrological processes; therefore, the LOCI method can often be used to improve the streamflow simulation by correcting overestimated wet day probability. This method has been used by many previous studies, e.g., Chen et al. (2013), Berg et al. (2012), and Maraun et al. (2010).*

2) We didn’t use humidity because the simulated streamflow is not sensitive to humidity (see Lines 342 ~ 343) according to sensitivity analysis in section 4.1.

3) For SWAT, it does compute evapotranspiration on a daily basis no matter whether there is a rainfall. In this study, the potential evapotranspiration is computed using Penman-Monteith method (Monteith, 1965) and it is based on temperature, radiation, humidity, etc.

4) For simulations with precipitation corrected by LS, the annual mean precipitation at Bayanbulak Station is about 267 mm and annual mean precipitation of the watershed is 664 mm (i.e. after upscaling from station scale to watershed scale through precipitation lapse rate), runoff is 310 mm, and evapotranspiration is 278 mm which is far from 0. We checked our graphs and do not know which graphs you were referring to. As a result, overestimation of the watershed precipitation is the main reason leading to the poor performance of LS. We addressed the reason in the revised manuscript on Lines 491 ~ 498.

Lines 342 ~ 343:

Relative humidity and wind speed are insensitive in this case.

Lines 491 ~ 498:

It is worth noting that simulations 1 to 3 and simulations 4 to 6, whose precipitation is corrected by LS and LOCI, respectively, vary significantly. The difference between LS and LOCI is that LOCI introduces a threshold for precipitation on wet days to correct the wet day probability while LS doesn’t. That is a simple but *quite pragmatic approach since the raw RCM simulated precipitation usually has too many drizzle days (Teutschbein and Seibert, 2012). Obviously, wet day probability is crucial to streamflow simulation when using elevation bands to account for spatial variation in SWAT (see more details in SWAT manual, http://www.brc.tamus.edu/).*

3. Finally, I think the conclusion is a bit weak. If there is drizzle in the RCM, you of course have to correct wet days. Why did you pick the PT and not the QM method as best for precipitation correction after drizzle correction? Please be clearer on how you think your results apply to your specific catchment or region and to bias correction for hydrology in general.

Our reply: *1) We applied PT instead of QM after drizzle correction because QM can correct the drizzle days and performs well in the literature (e.g., Teutschbein and Seibert, 2012). This makes the comparison between PT and QM meaningful.*

2) We added Lines 539 ~ 554 in the conclusion on how our result could be applied to other studies.

Lines 539 ~ 554:

4) For simulated streamflow, precipitation correction methods have more significant influence than temperature correction methods and their performances on streamflow simulations are consistent with these of corrected precipitation, i.e., PT and QM methods performed equally best in correcting flow duration curve and peak flow while LOCI method performed best in terms of the time-series based indices. Note the LOCI and DM methods should be used with caution when analyzing drought or extreme streamflows. Besides, LS method is not suitable in hydrological impact assessment where there is a large variation in precipitation distribution when few meteorological stations are used since LS fails to correct wet day probability.

Generally, selection of precipitation correction method is more important than the selection of temperature correction method to downscale GCM/RCM simulations and thereafter for streamflow simulations. This might be generally true for other regional studies as GCMs/RCMs normally tend to better represent the temperature field than the precipitation field. However, the selection of precipitation correction method will be case dependent. The comparison procedure listed in Figure 2 could be applied for other cases.

Minor comments and typos:

4. Title - You don't downscale with these methods, so remove that from the title!

Our reply: As discussed in *Chen et al. (2011; Section 3.2 on the third page), Maraun et al. (2010; Section 4 on the seventh page), Schmidli et al. (2007; Section 3.1 on the third page) and Colette et al. (2012; Section 2.2 on the second page), we tend to take these correction methods as downscaling methods. Therefore we think the title is appropriate.*

5. p12660 Line 21 – 'these simulations'

p12661 Line 4 – 'areas and models'

p12661 line 13 - 'the hydrologic system of the arid region is: : :'

p12661 line 24 – 'to study potential climate change'

p12662 line 14 – 'used to drive a hydrological model especially in an arid region where the hydrology is sensitive to climatic changes'

p12664 line 13 – (for more details: : :)'

p12665 line 12 – daily or monthly NS and R2?

p12672 line 16 – 'NS equal to -6.65'

p12672 line 17 – 'used for a hydrological: : :'

p12673 line 14 – 'All methods improve the raw: : :'

p12674 line 15ff – 'The LS method underestimates high precipitation values with probabilities below 0.06: : :' (See also later in this paragraph. You write either a probability of 0.05 or probabilities below 0.06)

p12676 line 24 – 'Teutschbein and Seibert

Our reply: *Thank you. We corrected these typos and carefully checked the revised manuscript. The corrected lines involved are listed below.*

Lines 23 ~24:

…all bias correction methods effectively improved these simulations;

Line 35:

…some results can be applied to other areas and models

Lines 47 ~ 48:

…the hydrologic system of the arid region is particularly vulnerable to climate change …

Lines 58 ~ 59:

… to study potential climate change on water resources (Liu et al., 2010, 2011).

Lines 77 ~ 78:

…they are used to drive a hydrological model especially in an arid region where the hydrology is sensitive to climate changes.

Line 132:

…(for more details refer to Gao et al., 2013).

Lines 159 ~ 160:

…with daily "NS"s (Nash-Sutcliffe coefficients, Nash and Sutcliffe, 1970; see the definition in Eq. 16) and "R^2"s over 0.80…

About the typo mistake (p12672 line 16 – 'NS equal to -6.65'), we think it is correct as NS here is not a plural noun and "equals" agrees with it.

Lines 332 ~ 334:

The streamflow simulated by the re-calibrated model was plotted in Fig. 3, and it systematically overestimates the observation with NS equals to -6.65.

Lines 334 ~ 336:

Therefore, it is necessary to correct the climate variables before they can be used for a hydrological impact study.

Lines 361 ~ 362:

All the bias correction methods improve the raw RCM simulated precipitation, however, there are differences in their corrected statistics.

Lines 396 ~ 409:

The LS method underestimates the high precipitation with probabilities below 0.06 and overestimates the low precipitation with probabilities between 0.06 ~ 0.32. The overestimation of precipitation with probabilities between 0.32 ~ 0.73 indicates LS method has a very limited ability in reproducing dry day precipitation (below 0.1 mm). Similar to LS method, the LOCI method also overestimates the low precipitation with probabilities between 0.08 ~ 0.32 and underestimates the high intensities with probabilities below 0.08, which is in line with previous arguments by Berg et al. (2012). However, unlike LS method, LOCI method performs well on the estimation of the dry days with precipitation below 0.1 mm. The PT, DM and QM methods well adjust precipitation exceedance except that DM method slightly overestimates the precipitation with probabilities between 0.12 ~ 0.28. For temperature, the raw temperature overestimates low temperature with probabilities above 0.65 and underestimates high temperature with probabilities below 0.65.

Line 476:

Teutschbein and Seibert (2012)

References

Berg, P., Feldmann, H., and Panitz, H. J.: Bias correction of high resolution regional climate model data, J Hydrol, 448, 80-92, 10.1016/j.jhydrol.2012.04.026, 2012.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins, J Hydrol, 479, 200-214, 10.1016/j.jhydrol.2012.11.062, 2013.

Chen, J., Brissette, F. P., Poulin, A., and Leconte, R.: Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed, Water Resour Res, 47, W12509, 10.1029/2011WR010602, 2011.

Colette, A., Vautard, R., and Vrac, M.: Regional climate downscaling with prior statistical correction of the global climate forcing, Geophys Res Lett, 39, 10.1029/2012gl052258, 2012.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev Geophys, 48, RG3003, 10.1029/2009RG000314, 2010.

Schmidli, J., Goodess, C., Frei, C., Haylock, M., Hundecha, Y., Ribalaygua, J., and Schmith, T.: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps, Journal of Geophysical Research: Atmospheres (1984–2012), 112, 2007.

Monteith JL (1965) Evaporation and environment. Symp Soc Exp Biol 19:205–234

Comparing bias correction methods in downscaling meteorological variables for hydrologic impact study in an arid area in China

G.H. Fang [1, 2, 3], J. Yang [1*,4], Y.N. Chen [1], C. Zammit [4]

[1] State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Xinjiang, China

[2] University of Chinese Academy of Sciences, Beijing, China

[3] Department of Geography, Ghent University, Ghent, Belgium

[4] National Institute of Water and Atmospheric Research, Christchurch, New Zealand

Corresponding author:

Jing Yang*

State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Xinjiang, 830011, China

818 South Beijing Road, Urumqi, Xinjiang, 830011, China

Tel: +86-991-7823171

Email: yangjing@ms.xjb.ac.cn

Comparing bias correction methods in downscaling meteorological variables for hydrologic impact study in an arid area in China

Abstract:

Water resources are essential to the ecosystem and social economy in the desert and oasis of the arid Tarim River Basin, Northwest China, and expected to be vulnerable to climate change. It has been demonstrated that Regional Climate Models (RCM) ~~have been proved to~~ provide more reliable results for regional impact study of climate change (e.g., on water resources) than General Circulation Models (GCM). However, due to their considerable bias it is still necessary to apply bias correction before they are used for water resources research ~~due to often considerable biases~~. In this paper, after a sensitivity analysis on input meteorological variables based on Sobol' method, we compared five precipitation correction methods and three temperature correction methods in downscaling ~~to the outputs of a~~ RCM simulations applied over the ~~model with its application to the~~ Kaidu River Basin, one of the headwaters of the Tarim River Basin. Precipitation correction methods applied include Linear Scaling (LS), LOCal Intensity scaling (LOCI), Power Transformation (PT), Distribution Mapping (DM) and Quantile Mapping (QM); while ~~and~~ temperature correction methods are ~~include~~ LS, VARIance scaling (VARI) and DM. The~~se~~ corrected precipitation and temperature were compared to the observed meteorological data, prior to be used as ~~climate~~meteorological inputs of a distributed hydrologic model to ~~determine~~study ~~and then~~ their impacts on streamflow ~~were also compared by driving a~~

~~distributed hydrologic model~~. The results show: 1) Streamflows are sensitive to ~~P~~precipitation, temperature, solar radiation but not to ~~are sensitivity to streamflow while~~ relative humidity and wind speed~~ are not~~; 2) Raw RCM simulations are heavily biased from observed meteorological data~~, –~~, and its use~~d~~ ~~–~~for streamflow simulations ~~which~~ results in large biases ~~with~~from ~~observationed~~observed streamflow~~in the simulated streamflows~~, and all bias correction methods effectively improved theses simulations; 3) For precipitation, PT and QM methods performed equally best in correcting the frequency-based indices (e.g., standard deviation, percentile values) while LOCI method performed best in terms of the time~~-~~-series based indices (e.g., Nash-Sutcliffe coefficient, R^2); 4) For temperature, all ~~bias~~ correction methods performed equally well in correcting raw temperature;~~.~~ 5) For simulated streamflow, precipitation correction methods have more significant influence than temperature correction methods and the performances of streamflow simulations are consistent with ~~these~~ those of corrected precipitation, i.e., PT and QM methods performed equally best in correcting flow duration curve and peak flow while LOCI method performed best in terms of the time~~-~~-series based indices. The case study is for an arid area in China based on a specific RCM and hydrologic model, but the methodology and some results can be applied to other areas and~~area and other~~ models.

1. Introduction

In recent decades, the ecological situation of the Tarim River Basin in China has seriously degraded especially in the lower reaches of the Tarim River due to water scarcity. In the meantime, climate change is significant in this region with an ~~consistent~~ increase in temperature at a rate of 0.33 ~ 0.39 ℃/decade and a slight increase in precipitation (Li et al., 2012) over the past 5 decades. Under the context of regional climate change, water resources in this region are expected to be more unstable and ecosystems are likely to suffer from severe water stress because the hydrologic system of the arid region is particularly vulnerable to climate change ~~in the arid region~~ (Arnell et al., 1992; Shen and Chen, 2010; ~~Sun et al., 2013;~~ Wang et al., 2013). The impact of climate change on hydrologic system has already been observed and it is expected that the hydrological system will continue to change in the future (Liu et al., 2010, 2011; Chen et al., 2010). Therefore, projecting reliable climate change and its impact on hydrology are important to study the ecology in the Tarim River Basin.

Only recently efforts have been made to evaluate and project the impact of climate change on hydrology in the Tarim River Basin. These studies include research on the relationships of meteorological variables and streamflow based on the historical measurements (e.g. Chen et al., 2013c; Xu et al., 2013), and use of the GCM outputs ~~of General Circulation Models (GCMs)~~ to drive a hydrologic model to study ~~the future~~ potential climate change on water resources (Liu et al., 2010, ~~; Liu et al.,~~ 2011). Study of~~on~~ historical climate - hydrology relationships has limited applications on future water resource management, especially under the context of global climate change

~~background~~. ~~And t~~Though GCMs have been widely used to study impacts of future climate change on hydrological systems and water resources, they are impeded by their inability to provide reliable information at the hydrological scales (Maraun et al., 2010; Giorgi, 1990). In particular, ~~in~~ for mountainous regions, fine scale information such as the altitude-dependent precipitation and temperature information, which is critical for hydrologic modeling, is not represented in GCMs (Seager and Vecchi, 2010). ~~Although there are options to downscale GCM outputs to the regional scale,~~ Therefore, recent studies tend to use the higher-resolution Regional Climate Models (RCMs) to preserve the physical coherence between atmospheric and land surface variables (Bergstrom et al., 2001; Anderson et al., 2011). As such, when evaluating the impact of climate change on water resources ~~in~~ on a watershed scale, the use of RCMs instead of GCMs is preferable since RCMs have been proved to provide more reliable results for impact study of climate change on regional water resources than GCM models (Buytaert et al., 2010; Elguindi et al., 2011). However, the raw RCM simulations may be still biased especially in the mountainous regions (Murphy, 1999; Fowler et al., 2007), which makes the use of RCM outputs as ~~the~~ direct input for hydrological model challenging~~,~~. As a result~~thus~~ it is of significance to properly correct the RCM simulated meteorological variables before they are used to drive a~~the~~ hydrological model especially in ~~the~~ an arid region~~s~~ where the hydrology is sensitive to climate changes.

Several bias correction methods have been developed to downscale meteorological variables from the RCMs, ranging from the simple scaling approach to sophisticated distribution mapping (Teutschbein and Seibert, 2012). And their applicability in the arid Tarim River Basin has not been investigated, thereby, evaluating and finding the appropriate bias correction method is necessary to evaluate the impact of climate change on~~to~~ water resources.

This study evaluates performances of five precipitation bias correction methods and three temperature bias correction methods in ~~correcting~~ downscaling RCM simulations and applied to the Kaidu River Basin, one of the most important headwaters of the Tarim River. These bias correction methods include most frequently used bias correction methods. We compare their performances in ~~terms of~~downscaling precipitation and temperature and evaluate their impact on streamflow through hydrological modeling.

The paper ~~remaining~~ is constructed as follows: Section 2 introduces the study area and data; Section 3 describes the bias correction methods for precipitation and temperature along with the hydrological model, sensitivity analysis method and result analysis strategy; and then Section 4 presents~~gives~~ results and discussion, followed by conclusions in Section 5.

2 Study area and data

2.1 Study area and observed data

The Kaidu River Basin, with a drainage area of 18,634 km^2 above the Dashankou hydrological station, is located on the south slope of the Tianshan Mountains in

Northwest China (Fig. 1). Its altitude ranges from 1,342~~0~~ m to 4,796 m above sea level (a.s.l.) with an average elevation of 2,995 m, and its climate is ~~featured by~~ temperate continental ~~climate~~ with alpine climate characteristics. As one of the headwaters of the Tarim River, it provides water resources for agricultural activity and ecological environment of the oasis in the lower reaches. This oasis, with a population of over 1.15 million, is stressed by lack of water and water resources are the main factor constricting the development (Chen et al., 2013b). Therefore, projecting the impact of future climate change on water resources is urgent to the sustainable development of this region.

Daily observed meteorological data, including precipitation, maximum/minimum temperature, wind speed and relative humidity of two meteorological stations (Bayanbulak and Baluntai, stars in Fig. 1), are from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/). The mean annual maximum and minimum temperature at the Bayanbulak meteorological station are 3.1 ℃ and -10.6 ℃ and mean annual precipitation is 267 mm, and generally precipitation falls as rain from May to September and as snow from October to April of the next year.

The observed streamflow data at the Dashankou hydrologic station (the triangle in Fig. 1) are from Xinjiang Tarim River Basin Management Bureau. The average daily flow is around 110 $m^3 s^{-1}$ (equivalent to 185 mm runoff per year), ranging from 15 $m^3 s^{-1}$ to 973 $m^3 s^{-1}$.

2.2 Simulated meteorological variables from the ~~regional climate model~~RCM

GCM or RCM outputs are generally biased (Ahmed et al., 2013; Teutschbein and Seibert, 2012; Mehrotra and Sharma, 2012), and there is a need to correct these outputs before used for regional impact studies. The RCM outputs used in this study are based on the work done by Gao et al. (2013), where the RCM outputs used in this study are based on the work done by Gao et al. (2013). In Gao et al. (2013), the RCM model (RegCM, Giorgi and Mearns, 1999) was driven by a global climate model BCC_CSM1.1 (Beijing Climate Center Climate System Model; Wu et al., 2013; Xin et al., 2013) at a horizontal resolution of 50 km over China.

The RCM outputs were validated with the observational dataset (CN05.1) over China for the period from 1961 to 2005. The RCM outputs show reasonable simulation of temperature and precipitation ~~over~~ in most parts of China except some regions where our study area is located~~especially when compared with its driving GCM BCC_CSM1.1~~ (for more details refer to Gao et al., 2013). ~~In this paper, meteorological outputs of the RCM model used include maximum/minimum temperature, precipitation, wind speed, solar radiation and humidity.~~

3 Methodology

Figure 2 shows the flow chart of the comparison procedure. First, grid based RCM simulation was downscaled to station scale using bias correction methods, and then the corrected meteorological data were compared to the observation at these two stations and to each other ("Meteorological data comparison" in Fig. 2). These station based meteorological data were then upscaled to watershed scale with the precipitation and temperature lapse rates before they were used to drive the hydrological model (SWAT). Finally, the simulated streamflow driven by the corrected and observed meteorological data were compared to observed streamflow and to each other ("Streamflow comparison" in Fig. 2).

3.1 Hydrologic model and sensitivity analysis~~of input meteorological variables~~

SWAT (Soil and Water Assessment Tool; Arnold et al., 1998) is a distributed and time continuous watershed hydrologic model. The climatic input (driving force) consists of daily precipitation, maximum/minimum temperature, solar radiation, wind speed and relative humidity. ~~, and SWAT uses elevation bands~~ T~~t~~o account for orographic effects on precipitation and temperature, elevation bands were used. Within each elevation band, the precipitation and temperature are estimated based on their lapse rates and elevation. For more details, refer to SWAT manuals (http://www.brc.tamus.edu/). ~~It~~ SWAT has been being widely used for comprehensive modeling of the impact of management practices and climate change on the hydrologic cycle and water resources at a watershed scale (e.g., Arnold et al., 2000; Arnold and Fohrer, 2005; Setegn et al., 2011).

In this study, SWAT model was firstly set up with available DEM, landuse, soil, and observed climate data, and then model parameters were calibrated with the observed streamflow data at the Dashankou S~~s~~tation. The simulation results show: 1) model application shows excellent performances for both calibration period (1986 ~ 1989) and validation period (1990 ~ 2001) with daily "NS"s (Nash-Sutcliffe coefficients, Nash and Sutcliffe, 1970; see the definition in Eq. 16) and "R^2"s over 0.80, which is highly acceptable; 2) model parameters are reasonable and spatial patterns of precipitation and temperature are in agreement with other studies in the region (see more details in Fang et al., ~~under submission~~2015). Figure ~~2~~ 3 shows a comparison of mean hydrographs of the observed ("obs") and simulated flows ("default"). This calibrated model hence provides a basis for evaluation of the impact of different correction methods on streamflow.

To study the relative importance of the five meteorological variables, the Sobol' sensitivity analysis method (Sobol', 2001) was applied. The Sobol' method is based on the decomposition of the variance V of objective function:

$$\mathrm{V} = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \cdots + V_{1,2,\cdots,n} \tag{1}$$

where

$$\mathrm{V_i} = V(\mu(Y|X_i))$$

$$\mathrm{V_{ij}} = V\left(\mu\big(Y\big|X_i, X_j\big)\right) - \mathrm{V_i} - \mathrm{V_j}$$

and so on. Herein, $V(.)$ denotes the variance operator, V is the total variance, and $\mathrm{V_i}$ and $\mathrm{V_{ij}}$ are main variance of X_i (the i^{th} factor of X) and partial variance of X_i and X_j. Here factors X are the changes applied to these five meteorological variables, respectively (see Table 1 for a list of these factors). In practice, normalized indices are often used as sensitivity measures:

$$S_i = \frac{\mathrm{V_i}}{\mathrm{V}}, 1 \le i \le n \tag{2}$$

$$S_{ij} = \frac{\mathrm{V_{ij}}}{\mathrm{V}}, 1 \leq i < j \leq n \quad (3)$$

$$S_{Ti} = S_i + \sum_j S_{ij} + \sum_j \sum_k S_{ijk} + \cdots + S_{1,2,\ldots,n}\,,\ 1 \leq i \leq n \quad (4)$$

Where S_i, S_{ij} and S_{Ti} are the main effect of X_i, first order interaction between X_i and X_j, and total effect of X_i. S_{Ti} ranges from 0 to 1 and denotes the importance of the factor to model output. The larger S_{Ti}, the more important this factor is. The difference between S_{Ti} and S_i denotes the significance of the interaction of this factor with other factors. As a result, the larger this difference, the more significant the interaction is.

3.2 Bias correction methods

In this study, five bias correction methods were used for precipitation, and three for temperature. These methods are listed in Table 2. All these bias correction methods were conducted on a daily basis from 1975 to 2005.

3.2.1 Linear Scaling (LS) of precipitation and temperature

LS method aims to perfectly match the monthly mean of corrected values with that of observed ones (Lenderink et al., 2007). It operates with monthly correction values based on the differences between observed and raw data (raw RCM simulated data in this case). Precipitation is typically corrected with a multiplier and temperature with an additive term on a monthly basis:

$$P_{cor,m,d} = P_{raw,m,d} \times \frac{\mu(P_{obs,m})}{\mu(P_{raw,m})} \quad (5)$$

$$T_{cor,m,d} = T_{raw,m,d} + \mu(T_{obs,m}) - \mu(T_{raw,m}) \quad (6)$$

where $P_{cor,m,d}$ and $T_{cor,m,d}$ are corrected precipitation and temperature on the d^{th} day of m^{th} month and $P_{raw,m,d}$ and $T_{raw,m,d}$ are the raw precipitation and temperature on the d^{th}

day of m^{th} month. $\mu(.)$ represents the expectation operator (e.g., $\mu(\cancel{T}P_{obs,m})$ represents the mean value of observed precipitation at given month m).

3.2.2 LOCal Intensity scaling (LOCI) of precipitation

LOCI method (Schmidli et al., 2006) corrects the wet-day frequencies and intensities and can effectively improve the raw data which have too many drizzle days (~~defined as~~ days with little precipitation). It normally involves two steps: firstly, a wet-day threshold for the m^{th} month $P_{thres,m}$ is determined from the raw precipitation series to ensure that the threshold exceedance matches the wet-day frequency of the observation; secondly, a scaling factor $s_m = \frac{\mu(P_{obs,m,d}|P_{obs,m,d}>0)}{\mu(P_{raw,m,d}|P_{raw,m,d}>P_{thres,m})}$ is calculated and used to ensure that the mean of the corrected precipitation is equal to that of the observed precipitation:

$$P_{cor,m,d} = \begin{cases} 0, & if\ P_{raw,m,d} < P_{thres,m} \\ P_{raw,m,d} \times S_m, & otherwise \end{cases} \quad (7)$$

3.2.3 Power Transformation (PT) of precipitation

While the LS and LOCI account for the bias in the mean precipitation, it does not correct biases in the variance. PT method uses an exponential form to further adjust the standard deviation of precipitation series. Since PT has the limitation in correcting the wet day probability (Teutschbein and Seibert, 2012), which was also confirmed in our study (not shown), LOCI method is applied to correct precipitation prior to the correction by PT method.

Therefore, to implement this PT method, firstly, we estimate b_m that minimizes:

$$f(b_m) = \frac{\sigma(P_{obs,m})}{\mu(P_{obs,m})} - \frac{\sigma(P_{LOCI,m}^{b_m})}{\mu(P_{LOCI,m}^{b_m})} \quad (8)$$

where b_m is the exponent for the m^{th} month, $\sigma(.)$ represents the standard deviation operator, and $P_{LOCI,m}$ is the LOCI-corrected precipitation in the m^{th} month. If b_m is larger than one, it indicates that the LOCI-corrected precipitation underestimates its coefficient of variance in month m.

After finding the optimal b_m, the parameter $s_m = \frac{\mu(P_{obs,m})}{\mu\left(P_{LOCI,m}^{b_m}\right)}$ is then determined such that the mean of the corrected values corresponds to the observed mean. The corrected precipitation series are obtained based on the LOCI corrected precipitation $P_{cor,m,d}$:

$$P_{cor,m,d} = s_m \times P_{LOCI,m,d}^{b_m} \quad (9)$$

3.2.4 ~~Variance~~ VARIance scaling (VARI) of temperature

The PT method is an effective method to correct both the mean and ~~the~~ variance of precipitation, but it cannot be used to correct temperature time series, as temperature is known to be approximately normally distributed (Terink et al., 2010). VARI method was developed to correct both the mean and variance of normally distributed variable such as temperature (Teutschbein and Seibert, 2012; Terink et al., 2010). Temperature is normally corrected using VARI method with Eq. (10).

$$T_{cor,m,d} = \left[T_{raw,m,d} - \mu\left(T_{raw,m}\right)\right] \times \frac{\sigma\left(T_{obs,m}\right)}{\sigma\left(T_{raw,m}\right)} + \mu\left(T_{obs,m}\right) \quad (10)$$

3.2.5 Distribution ~~mapping~~ Mapping (DM) of precipitation and temperature

DM method is to match the distribution function of raw data to that of observation.

It is used to adjust mean, standard deviation and quantiles. Furthermore, it preserves the extremes (Themeßl et al., 2012). However, it also has its limitation due to the assumption that both the observed and raw meteorological variables follow the same proposed distribution, which may introduce potential new biases.

For precipitation, the Gamma distribution (Thom, 1958) with shape parameter α and scale parameter β is often used for precipitation distribution and has been proven to be effective (e.g., Block et al., 2009; Piani et al., 2010):

$$f_r(x|\alpha,\beta) = x^{\alpha-1} \times \frac{1}{\beta^{\alpha} \times \Gamma(\alpha)} \times e^{\frac{-x}{\beta}}; x \geq 0, \alpha, \beta > 0 \tag{11}$$

where $\Gamma(.)$ is the Gamma function. Since the raw RCM-simulated precipitation contains a large number of drizzle days, which may substantially distort the raw precipitation distribution, the correction is done on LOCI corrected precipitation $P_{LOCI,m,d}$:

$$P_{cor,m,d} = F_r^{-1}\left(F_r\left(P_{LOCI,m,d}\middle|\alpha_{LOCI,m}, \beta_{LOCI,m}\right)\middle|\alpha_{obs,m}, \beta_{obs,m}\right) \tag{12}$$

Where $F_r(.)$ and $F_r^{-1}(.)$ are Gamma CDF (cumulative distribution function) and its inverse. $\alpha_{LOCI,m}$ and $\beta_{LOCI,m}$ are the fitted Gamma parameter for the LOCI corrected precipitation in a given month m, and $\alpha_{obs,m}$ and $\beta_{obs,m}$ are these for observation.

For temperature, the Gaussian distribution (or normal distribution) with mean μ and standard deviation σ is usually assumed to fit temperature best (Teutschbein and Seibert, 2012):

$$f_N(x|\mu,\sigma) = \frac{1}{\sigma \times \sqrt{2\pi}} \times e^{\frac{-(x-\mu)^2}{2\sigma^2}}; x \in \mathbf{R} \tag{13}$$

And then similarly the corrected temperature can be expressed as:

$$T_{cor,m,d} = F_N^{-1}\left(F_N\left(T_{raw,m,d}\middle|\mu_{raw,m}, \sigma_{raw,m}\right)\middle|\mu_{obs,m}, \sigma_{obs,m}\right) \quad (14)$$

where $F_N(.)$ and $F_N^{-1}(.)$ are Gaussian CDF and its inverse, $\mu_{raw,m}$ and $\mu_{obs,m}$ are the fitted and observed means for the raw and observed precipitation series at a given month m, and $\sigma_{raw,m}$ and $\sigma_{obs,m}$ are the corresponding standard deviations, respectively.

3.2.6 Quantile Mapping (QM) of precipitation

QM method is a non-parametric bias correction method and is generally applicable for all possible distributions of precipitation without any assumption on precipitation distribution. This approach originates from the empirical transformation (Themeßl et al., 2012) and was successfully implemented in the bias correction of RCM simulated precipitation (Sun et al., 2011; Themeßl et al., 2012; Chen et al., 2013a; Wilcke et al., 2013). It can effectively correct bias in the mean, standard deviation and wet day frequency as well as quantiles.

For precipitation, the adjustment of precipitation using QM can be expressed in terms of the empirical CDF (*ecdf*) and its inverse (*ecdf*$^{-1}$):

$$P_{cor,m,d} = ecdf_{obs,m}^{-1}\left(ecdf_{raw,m}\left(P_{raw,m,d}\right)\right) \quad (15)$$

3.3 Performance evaluation

The performance evaluation of these correction methods is based on their abilities to reproduce precipitation, temperature, and streamflow simulated with a hydrological model (SWAT) driven by bias corrected RCM simulations~~, specifically~~. When evaluating ability to reproduce streamflow, streamflow is firstly simulated by running the hydrological model driven by 15 different combinations of corrected precipitation, max/min temperature with different correction methods (these hydrologic simulations are then referred to as simulations 1 to 15, which are listed in Table 3) together with hydrologic simulations driven by observed meteorological data (“default”) and raw RCM simulation (“raw”). These 15 simulations were then compared with observed streamflows and “default” and “raw”.

The performance evaluation of precipitation, temperature and streamflow ~~with different correction methods are:~~are as follows.

1) For corrected precipitation, frequency-based indices and time series performances are compared with observed precipitation data. The frequency-based indices include mean, median, standard deviation, 99~~0~~th percentile, probability of wet days, and intensity of wet day while time-~~-~~series based metrics include Nash-Sutcliffe coefficient(NS), Percent bias (P_{BIAS}), R^2 and Mean Absolute Error (MAE) defined as follows:

$$NS = 1 - \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)^2}{\sum_{i=1}^{n}\left(Y_i^{obs} - Y^{mean}\right)^2} \tag{16}$$

$$P_{BIAS} = \frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right)}{\sum_{i=1}^{n}\left(Y_i^{obs}\right)} \tag{17}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_i^{obs} - Y_i^{sim}\right|}{n} \tag{18}$$

Where Y_i^{obs} and Y_i^{sim} are the i^{th} observed and simulated variables, Y^{mean} is the mean of observed variables, and n is the total number of observations.

NS indicates how well the simulation matches the observation and it ranges between -∞ and 1.~~0~~, with NS =1 meaning a perfect fit. The higher this value, the more reliable the model is in comparison to the mean. P_{BIAS} measures the average tendency of the simulated data to their observed counterparts. Positive values indicate an overestimation of observation, while negative values indicate an underestimation. The optimal value of P_{BIAS} is 0.0, with low-magnitude values indicating accurate model simulations. MAE demonstrates the average model prediction error with less sensitivity to large errors.

2) For corrected temperature, frequency-based indices and time series performances are compared with observed temperature data. The frequency-based indices include mean, median, standard deviation, and 10th, 90th percentile~~s~~ while time-series based metrics include NS, P_{BIAS}, R^2 and MAE.

3) For simulated streamflow driven by corrected RCM simulations, the frequency-based indices are visualized using boxplot, exceedance probability curve~~,~~ ~~and exceedance probabilities of 7-day peak flow and low flow~~. Time-~~-~~series based metrics include NS, P_{BIAS}, R^2 and MAE.

4 Results and discussion

4.1. Initial streamflow simulation driven with raw RCM simulations and sensitivity analysis

To illustrate the necessity of bias correction in climate change impact on hydrology, we re-calibrated SWAT using the raw RCM simulations while keeping all

SWAT parameters in their reasonable ranges. The assumption is that if the re-calibrated hydrological model driven by the raw RCM simulations performs well and model parameters are reasonable, then there is no need for bias correction. The streamflow simulated by the re-calibrated model was plotted in Fig. ~~2~~3, and it systematically overestimates the observation ~~a lot~~ with NS equals to -6.65. Therefore, it is necessary to correct the meteorological variables before they can be used for a hydrological impact study.

~~And then t~~The Sobol' method was applied to study which meteorological variables should be corrected for hydrological modeling. Table 1 lists the sensitivity results for these five meteorological variables. As ~~it~~ can be seen, precipitation is the most sensitive factor (the main effect S_i is 44.0% and total effect S_{Ti} is 74.0%), followed by temperature (S_i = 15.0% and S_{Ti} = 36.9%) and solar radiation (S_i = 7.7% and S_{Ti} = 22.6%), and the interactions between these factors are large. ~~The r~~Relative humidity and wind speed are insensitive in this case. This means precipitation, temperature and solar radiation need to be bias corrected before applied to hydrologic models, while relative humidity and wind speed over the region do not need any correction.

4.2 Evaluation of corrected precipitation and temperature

The bias correction was done on RCM simulated precipitation, ~~minimum temperature,~~ max~~imum~~/min temperature, and solar radiation (for solar radiation, LS and VARI methods were used) for two meteorological stations Bayanbulak and Baluntai. Results show: 1) for solar radiation, there is no significant difference for different correction methods. There the results are not shown. 2) Similar results were obtained for minimum temperature and maximum temperature, and for Bayanbulak and Baluntai. Therefore we only listed and discussed results for Bayanbulak, and maximum temperature.

Table 4 lists the frequency-based statistics of observed ("obs"), raw RCM-simulated ("raw") and corrected (denoted by the corresponding correction method) precipitation data at the Bayanbulak Station. This station has a ~~low~~ daily mean precipitation ~~(daily mean~~ of 0.73 mm or annual mean of 266 mm~~)~~ and precipitation falls in 32% days in a year with a mean intensity of 2.3 mm. Compared to the observation, the raw RCM simulation~~s~~ deviates significantly from observation, with overestimation of all the statistics. All the bias correction methods improve~~s~~ the raw RCM simulated precipitation, however, there are differences ~~between~~ in their corrected statistics. LS method has a good estimation of the mean while it shows a large bias in other measures, e.g., it largely overestimated the probability of wet days (e.g., up to 41% overestimation) and underestimated the standard deviation (up to 0.9~~1~~ mm underestimation). LOCI method provides a good estimation in the mean, median, wet-day probability and wet-day intensity; however, there is a slight underestimation in the standard deviation and therefore 99~~0~~th percentile. Compared to LS and LOCI, PT method performs well in all these metrics. ~~In spite of slight better estimation of standard deviation, probability of wet days and intensity of wet day,~~ DM method has a~~n~~ slight overestimation of the mean and an underestimation of standard deviation. This means that precipitation does not follow the assumed Gamma distribution. On the contrary, QM method doesn’t have this assumption and it provides an excellent estimation of these statistics. These results are consistent with previous studies (Themeßl et al., 2011, 2012; Wilcke et al., 2013; Graham et al., 2007), but are different from the research by Piani et al. (2010) who found that performance of DM method is unexpectedly well for the humid Europe region. This discrepancy can be partly attributed to the precipitation regime for different regions since better fit of the assumed distribution lead to better performance of DM.

Table 5 lists the frequency-based statistics of observed (“obs”), raw RCM simulated (“raw”) and ~~bias-~~corrected (denoted by the corresponding method) maximum temperature data at the Bayanbulak Station. The mean and standard deviation of “obs” are 3.1 and 14.5 ℃, with the 90^{th} percentile being 19.2 ℃. Analysis of the “raw”~~ RCM simulations~~ indicates deviation from “obs”~~ervation~~, with an overestimation of the mean, and underestimations of the median, standard deviation, and 90^{th} percentile. All three ~~bias-~~correction methods correct~~ed~~ biases in ~~RCM simulated temperature~~the “raw” and improve~~d~~ estimations of the statistics. LS has a correct estimation of mean but~~ a~~ slight underestimations of median and standard deviation, while VARI and DM have good estimations~~a good match with observations for~~ of all the frequency-based statistics. These results ~~are in accordance with~~ confirm the study by Teutschbein and Seibert (2012), i.e., LS method doesn’t adjust the standard deviation and the ~~10th/90th~~ percentiles while VARI and DM methods do.

Figure ~~3~~ 4 shows the exceedance probability curves of the observed and corrected precipitation and temperature. For precipitation, the raw RCM simulations are heavily biased (as also shown by statistics in Table 4). All correction methods effectively, but in different extent, correct biases in raw precipitation. The LS method underestimates the high precipitation with probabilities~~y~~ below 0.06 and overestimates the low precipitation with ~~probability~~ probabilities between 0.06 ~ 0.32. The overestimation of precipitation with ~~probability~~ probabilities between 0.32 ~ 0.73 indicates LS method has a very limited ability in reproducing dry day precipitation (below 0.1 mm). Similar to LS method, the LOCI method also overestimates the low precipitation with probabilities~~y~~ between 0.08 ~ 0.32 and underestimates the high ~~precipitation~~ intensities with probabilities~~y~~ below 0.08, which is in line with previous arguments by Berg et al. (2012). However, unlike LS method, LOCI method performs well on the estimation of the dry days with precipitation below 0.1 mm. The PT, DM and QM methods well adjust precipitation exceedance except that DM method slightly overestimates the precipitation with probabilities~~y~~ between 0.12 ~ 0.28. For temperature, the raw temperature overestimates low temperature with probabilities~~y~~ above 0.65 and underestimates high temperature with probabilities~~y~~ below 0.65. All temperature correction methods adjust the biases in raw temperature and the corrected temperature has ~~the~~ similar quantile values with the observation. They performed equally well and differences among these correction methods are negligible.

Time-~~-~~series based performances were evaluated and results are ~~listed~~ shown in Fig. ~~4~~ 5 and Table 6. For precipitation, all bias correction methods significantly improve the raw RCM simulations. However, as shown in the right plot of Fig. ~~4~~5, there is a systematic mismatch between observation and corrections which follow the pattern of the raw RCM simulated precipitation, which indicates that all bias correction methods fail to correct the temporal pattern of precipitation. In addition, this mismatch differs between different methods, among which the differences ~~is~~ are smaller for LS and LOCI methods than for PT, DM and QM methods. This resulted in a slightly better squared difference based measures (e.g., NS, R^2) for LS and LOCI than PT, DM and QM methods, as is indicated in Table 6. Similar to precipitation, all correction methods significantly improved the raw RCM simulated temperature. ~~Differences between observation and~~ Biases in raw temperature (e.g., 1.1 ℃ in spring, 1.0 ℃ in summer, 3.3 ℃ in autumn, and up to 7.6 ℃ in winter) were ~~significantly~~ corrected. These three correction methods performed equally well and no significant differences exist ~~between~~ in terms of the average daily temperature graphs.

Table 6 lists ~~performances of~~ the time-series based metrics of corrected ~~ion methods for monthly time series of~~ precipitation and temperature at the Bayanbulak Station. For precipitation, the performance of the raw RCM simulated precipitation is very poor with NS = -6.78, P_{BIAS} = 293.28% and MAE = 65.40 mm for monthly data, and the improvements of correction are obvious. The "P_{BIAS}"s of the corrected precipitation are within ±~~5~~7 % and “NS”s approach 0.64. It is worth noting that LS and LOCI methods perform better than PT and QM methods in terms of time series performances. For temperature, although the raw RCM simulation obtains an acceptable NS value (0.84), it ~~severely~~ overestimates the observation with ~~(~~P_{BIAS} = 15.78% and MAE = 4.31 ℃~~)~~. The “P_{BIAS}”s of the corrected temperatures are within ±5% and “NS”s are over 94% (better than that of the “raw”) for all three correction methods and there is no significant difference between these results, which indicates the corrected monthly temperature series are in good agreement with the observation.

4.3 Evaluation of streamflow simulations

Figure ~~5~~ 6 compares the mean, median, first and third quantiles of daily observed streamflows (“obs”), ~~with~~ simulated streamflows ~~driven by~~ using observed meteorological inputs (“default”), raw RCM simulations (“raw”) and 15 combinations of corrected precipitation and corrected temperature (i.e., simulations 1 to 15). The overestimation of simulated streamflow using raw RCM simulations (i.e., “raw”) is obvious. ~~For s~~Simulations 1 to 3~~, streamflow~~ overestimate streamflow with 100% overestimation of the mean streamflow~~ions are also observed and they substantially overestimate the mean streamflow by over 100%,~~ while simulations 4 to 15 reproduce similar streamflows as the observation or simulation “default”. As the major difference between simulations 1 to 3 and other simulations is that simulations 1 to 3 use the LS-corrected precipitation, ~~this~~ which means precipitation corrected with LS method ~~is~~

~~not suitable~~has great bias in ~~for~~ flow simulation in this study.

To investigate the performances of bias correction methods for different hydrological seasons, we divided the streamflow into two different periods according to the hydrograph (Fig. 3): wet period is from April to September and dry period is from October to March of next year. It indicates that the performances of bias correction methods are, except for magnitudes, similar for both wet and dry period (not shown), which demonstrates that the evaluation is robust and can provide useful information for both dry and wet seasons.

Figure ~~6~~ 7 shows the exceedance probability curves (flow duration curves) of the observed streamflow ("obs"), and streamflows with simulation "default" and simulations 4 to 15. For plotting purpose, simulations "raw" and 1 to 3 are not shown. Generally all simulations are in good agreement with the observation ~~for frequencies~~with probabilities between 0.12 and 0.72, and precipitation correction methods have more significant influence than temperature correction methods. This confirms the previous sensitivity results that precipitation is the most sensitive driving force in~~to~~ streamflow simulation. Similar to performances of bias corrected precipitation, simulations with DM corrected precipitation (i.e. simulations 10 to 12) and ~~deviate the observation the most, followed these with~~ LOCI corrected precipitation (i.e., simulations 4 to 6) deviate the observation the most, ~~and then~~ followed these with PT and QM methods. All simulations encounter the problem to correctly mimic the low flow part (i.e. ~~exceedance~~ probabilities larger than 0.7). This might be a systematic problem of the calibrated hydrologic model (as indicated by simulation "default"), e.g., the objective function of the hydrological modeling is not focused on baseflow. Differences among streamflows driven by different temperature but same precipitation are insignificant~~. This result~~, which is different~~s~~ from the study of Teutschbeien and Seibert (2012). This may be related to the~~ chosen RCM model or~~ watershed characteristic.

The ~~time series~~ performances of ~~simulation "default",~~ simulation "raw"~~ and~~, simulations 1 to 15 at daily and monthly time steps (simulation "default" is taken as reference) are summarized in Table 3.~~ The "default" performs well with NS reaching 0.80 for daily and 0.90 for monthly streamflow and daily MAE within 25 m³/s.~~ The "raw" is heavily biased with NS close to -56~~3~~.3 and P_{BIAS} as large as ~~421~~ 399 % for monthly data. All the 15 simulations improve the statistics ~~of the "raw" scenario~~ significantly. For simulations 1 to 3, whose precipitation series are corrected by LS method, NS ranges from -3.09~~3.10~~ to -~~2.87~~2.85 for monthly streamflow and they substantially overestimate the streamflow with P_{BIAS} over 100~~10~~ %. For simulations 4 to 15, monthly "NS"s are over 0.60, which indicates they can reproduce satisfactory monthly streamflows in this watershed, and simulations with precipitation corrected by LOCI (simulations 4 to 6) have best "NS"s and "P_{BIAS}"s. However, these indices of ~~daily streamflow~~ are lower for daily streamflow (~~the highest~~ "NS"s ~~is~~range from 0.38 to 0.50~~ for simulations 5 and 6~~), and this is related to the mismatch between corrected and observed precipitation time series (see top plot in Fig.~~ure~~ ~~4~~5), which is intrinsic from the RCM model and cannot be improved through these correction methods.

It is worth noting that simulations 1 to 3 and simulations 4 to 6, whose precipitation is corrected by LS and LOCI, respectively, vary significantly. The difference between LS and LOCI is that LOCI introduces a threshold for precipitation on wet days to~~for the wet day precipitation to~~ correct the wet day probability while LS doesn't. That is a simple but quite pragmatic approach since the raw RCM simulated precipitation usually has too many drizzle days (Teutschbein and Seibert, 2012). Obviously, wet day probability is crucial to streamflow simulation when using elevation bands to account for spatial variation in SWAT ~~in this study~~(see more details in SWAT manual, http://www.brc.tamus.edu/).

Figure ~~7~~ 8 shows the ~~simulated~~ monthly mean streamflow and exceedance probability curves of 7-day peak and 7-day low flow. For the monthly mean streamflow, obviously the "raw" is heavily biased with deviations ranging from 282% to 426%. Simulations 1 to 3 also overestimate the observation and the "default" as discussed before, while simulations 4 to 15 reproduced good monthly mean streamflow ~~especially for simulations 4, 5 and 6~~. The annual peak flow and low flow are~~is~~ presented in Fig. 8~~7~~ to investigate the impact of bias correction methods on extreme flows. For the peak flow, the exceedance probabilities of the simulations 4 to 15 are close to the observation while "raw" and simulations 1 to 3 deviate significantly (not shown). It is worth noting that simulations 4, 5 and 6, which perform the best in terms of the "NS"s, ~~slightly~~ underestimate the peak flow by 1% ~ 28%. The reason may be that the LOCI method adjusts all precipitation events in a certain month with a same scaling factor, which leads to the underestimation of the standard deviation ~~(Table 4)~~ and high precipitation intensity (Table 4), and finally results in an underestimation of the peak streamflow. For the low flow, all simulations overestimate the observation, but are in good agreement with the "default", which can be attributed to the systematic deficit in the hydrological model. DM method slightly overestimates both peak flow and low flow. Results show slightly better performance of PT and QM methods than LOCI and DM in predicting extreme flood and low flow, which is consistent with previous studies in North America and Europe (e.g., Chen et al., 2013a; Teutschbein and Seibert, 2012).

~~For the peak flow and low flow, both DM and QM methods perform well and QM method is slightly better than DM method as the latter overestimates both peak flow and low flow. However, there is an essential problem of QM method when comes to correcting future climate since it fails to resolve the "new extreme" (modeled values beyond the observed range) problem (Themeßl et al., 2012) as the corrected precipitation always falls between the maximum and minimum values.~~

5 Conclusions

The~~is~~ work presented in this study compared the abilities of five precipitation ~~bias correction methods~~ and three temperature ~~bias~~ correction methods in ~~correcting~~ downscaling RCM simulations. The downscaled ~~climate~~meteorological ~~information~~data ~~is~~were then used to model ~~and their impact on for~~ hydrologic processes in an arid ~~mountainous~~ region in China. The evaluation of the correction methods ~~is carried out~~ includes their abilities to reproduce precipitation, temperature and ~~simulated~~ streamflow using a hydrological model driven by corrected meteorological variables. Several conclusions can be drawn:

1) Sensitivity analysis shows precipitation is the most sensitive driving force ~~to~~ in streamflow simulation, followed by temperature and solar radiation, while relative humidity andr wind speed are not sensitive.

2) ~~The r~~Raw RCM simulations are heavily biased from observed meteorological data, and this results in biases in the simulated streamflows which cannot be corrected through calibration of the hydrological~~by~~ model~~ calibration~~. However ~~; and~~ all bias correction methods effectively improve precipitation, temperature, and streamflow ~~these~~ simulations.

3) Different precipitation correction methods show a big difference in downscaled precipitations while different temperature correction methods show similar results in downscaled temperatures. For precipitation, the PT and QM methods performed equally best in terms of the frequency-based indices~~, (e.g., mean, standard deviation, percentiles)~~; while LOCI method performed best in terms of the time-~~-~~series based indices~~ (e.g., NS, P_{BIAS} and R^2)~~. ~~For temperature, the raw RCM simulated temperature is highly relevant to the observation but generally biased (R^2 = 0.88 and P_{BIAS} = 15.78% for monthly data). A~~a~~ll correction methods effectively corrected biases in the raw RCM simulated temperature and they performed almost equally well for both frequency-based indices and time series based indices.~~

4) For simulated streamflow, precipitation correction methods have more significant influence than temperature correction methods and their performances ~~of~~ on

streamflow simulations are consistent with these of corrected precipitation, i.e., PT and QM methods performed equally best in correcting flow duration curve and peak flow while LOCI method performed best in terms of the time-~~-~~series based indices~~(e.g., NS = 0.69, $|P_{BIAS}| < 5\%$)~~. Note the LOCI and DM methods should be used with caution when analyzing drought or extreme streamflows. Besides,~~the wet day probability is vital in simulating streamflow in this study and it is recommended the LOCI method be applied to correct precipitation prior to the correction by PT method. This study also stresses~~ LS method is not suitable in hydrological impact assessment where there is a large variation in precipitation distribution when few meteorological stations are used since LS fails to correct wet day probability.~~the need for bias correction when assessing the impact of climate change on hydrology using the RCM simulations.~~

Generally, selection of precipitation correction method is more important than the selection of temperature correction method to downscale GCM/RCM simulations and thereafter for streamflow simulations. This might be generally true for other regional studies as GCMs/RCMs normally tend to better represent the temperature field than the precipitation field. However, the selection of precipitation correction method will be case dependent. The comparison procedure listed in Figure 2 could be applied for other cases. ~~The most appropriate bias correction method for RCM simulations may differ regarding to climate conditions or evaluation indices . As such, it is necessary to find an appropriate bias correction method based on the study purpose.~~

Acknowledgment

The research was supported by the “Thousand Youth Talents” Plan (Xinjiang Project) and the National Natural Science Foundation of China (41471030) and the Foundation of State Key Laboratory of Desert and Oasis Ecology (Y371163)~~Basic Research Program of China (973 Program: 2010CB951003)~~. We ~~wish~~ would like to thank Prof. Xuejie Gao ~~X~~ at National Climate Center (China) for providing the output of Regional Climate Model used in this paper.

References

Ahmed, K. F., Wang, G., Silander, J., Wilson, A. M., Allen, J. M., Horton, R., and Anyah, R.: Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the US northeast, Global Planet Change, 100, 320-332, 2013.

Anderson, W. P., Jr., Storniolo, R. E., and Rice, J. S.: Bank thermal storage as a sink of temperature surges in urbanized streams, J Hydrol, 409, 525-537, 10.1016/j.jhydrol.2011.08.059, 2011.

Arnell, N. W.: Factors controlling the effects of climate change on river flow regimes in a humid temperate environment, J Hydrol, 132, 321-342, http://dx.doi.org/10.1016/0022-1694(92)90184-W, 1992

Arnold, J., Muttiah, R., Srinivasan, R., and Allen, P.: Regional estimation of base flow and groundwater recharge in the Upper Mississippi river basin, J Hydrol, 227, 21-40, 2000.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J.: Large area hydrologic modeling and assessment part I: Model development1, JAWRA Journal of the American Water Resources Association, 34, 73-89, 1998.

Arnold, J. G., and Fohrer, N.: SWAT2000: current capabilities and research opportunities in applied watershed modelling, Hydrol Process, 19, 563-572, 10.1002/hyp.5611, 2005.

Berg, P., Feldmann, H., and Panitz, H. J.: Bias correction of high resolution regional climate model data, J Hydrol, 448, 80-92, 2012.

Bergstrom, S., Carlsson, B., Gardelin, M., Lindstrom, G., Pettersson, A., and Rummukainen, M.: Climate change impacts on runoff in Sweden-assessments by global climate models, dynamical downscaling and hydrological modelling, Climate Res, 16, 101-112, 2001.

Block, P. J., Souza Filho, F. A., Sun, L., and Kwon, H. H.: A Streamflow Forecasting Framework using

Multiple Climate and Hydrological Models1, JAWRA Journal of the American Water Resources Association, 45, 828-843, 2009.

Buytaert, W., Vuille, M., Dewulf, A., Urrutia, R., Karmalkar, A., and Celleri, R.: Uncertainties in climate change projections and regional downscaling in the tropical Andes: implications for water resources management, Hydrol Earth Syst Sc, 14, 1247-1258, 10.5194/hess-14-1247-2010, 2010.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resour Res, 49, 4187-4205, 10.1002/wrcr.20331, 2013a.

Chen, Y., Xu, C., Chen, Y., Li, W., and Liu, J.: Response of glacial-lake outburst floods to climate change in the Yarkant River basin on northern slope of Karakoram Mountains, China, Quaternary International, 226, 75-81, 2010.

Chen, Y., Du, Q., and Chen, Y.: Sustainable water use in the Bosten Lake Basin, Science press, Beijing, 329 pp., 2013b.

Chen, Z., Chen, Y., and Li, B.: Quantifying the effects of climate variability and human activities on runoff for Kaidu River Basin in arid region of northwest China, Theoretical and applied climatology, 111, 537-545, 2013c.

Elguindi, N., Somot, S., Déqué, M., and Ludwig, W.: Climate change evolution of the hydrological balance of the Mediterranean, Black and Caspian Seas: impact of climate model resolution, Clim Dynam, 36, 205-228, 2011.

Fang, G., Yang, J., Chen, Y., Xu, C.: Contribution of meteorological input in calibrating a distributed hydrologic model with the application to a watershed in the Tianshan Mountains, China, Environ Earth Sci, 10.1007/s12665-015-4244-7, ~~submitted Sep, 2014, under submission~~2015.

Fowler, H. J., Ekström, M., Blenkinsop, S., and Smith, A. P.: Estimating change in extreme European precipitation using a multimodel ensemble, J Geophys Res, 112, D18, 2007.

Gao, X., Wang, M., and Giorgi, F.: Climate change over China in the 21st century as simu-lated by BCC_CSM1. 1-RegCM4. 0, Atmos. Oceanic Sci. Lett, 6, 381-386, 2013.

Giorgi, F.: Simulation of regional climate using a limited area model nested in a general circulation model, Journal of Climate 3, 941-963, 1990.

Giorgi, F., and Mearns, L. O.: Introduction to special section: Regional climate modeling revisited, Journal of Geophysical Research: Atmospheres (1984–2012), 104, 6335-6352, 1999.

Graham, L. P., Andréasson, J., and Carlsson, B.: Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods–a case study on the Lule River basin, Climatic Change, 81, 293-307, 2007.

Lenderink, G., Buishand, A., and Deursen, W. v.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, Hydrol Earth Syst Sc, 11, 1145-1159, 2007.

Li, B., Chen, Y., and Shi, X.: Why does the temperature rise faster in the arid region of northwest China?, J Geophys Res, 117, 2012.

Liu, T., Willems, P., Pan, X. L., Bao, A. M., Chen, X., Veroustraete, F., and Dong, Q. H.: Climate change impact on water resource extremes in a headwater region of the Tarim basin in China, Hydrol Earth Syst Sc, 15, 3511-3527, 10.5194/hess-15-3511-2011, 2011.

Liu, Z., Xu, Z., Huang, J., Charles, S. P., and Fu, G.: Impacts of climate change on hydrological processes in the headwater catchment of the Tarim River basin, China, Hydrol Process, 24, 196-208, 10.1002/hyp.7493, 2010.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user, Rev Geophys, 48, RG3003, 2010.

Mehrotra, R., and Sharma, A.: An improved standardization procedure to remove systematic low frequency variability biases in GCM simulations, Water Resour Res, 48, W12601, 10.1029/2012WR012446, 2012.

Murphy, J.: An evaluation of statistical and dynamical techniques for downscaling local climate, Journal of Climate, 12, 2256-2284, 1999.

Nash, J. E., and Sutcliffe, J.: River flow forecasting through conceptual models part I—A discussion of principles, J Hydrol, 10, 282-290, 1970.

Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theoretical and Applied Climatology, 99, 187-192, 2010.

Schmidli, J., Frei, C., and Vidale, P. L.: Downscaling from GC precipitation: A benchmark for dynamical and statistical downscaling methods, International Journal of Climatology, 26, 679-689, 10.1002/joc.1287, 2006.

Seager, R., and Vecchi, G. A.: Greenhouse warming and the 21st century hydroclimate of southwestern North America, Proc. Natl. Acad. Sci. U. S. A., 107, 21277-21282, 10.1073/pnas.0910856107, 2010.

Setegn, S. G., Rayner, D., Melesse, A. M., Dargahi, B., and Srinivasan, R.: Impact of climate change on the hydroclimatology of Lake Tana Basin, Ethiopia, Water Resour Res, 47, W04511, 10.1029/2010WR009248, 2011.

Shen, Y., and Chen, Y.: Global perspective on hydrology, water balance, and water resources management in arid basins, Hydrol Process, 24, 129-135, 2010.

Sobol', I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math Comput Simulat, 55, 271-280, 10.1016/s0378-4754(00)00270-6, 2001.

Sun, F., Roderick, M. L., Lim, W. H., and Farquhar, G. D.: Hydroclimatic projections for the Murray-Darling Basin based on an ensemble derived from Intergovernmental Panel on Climate Change AR4 climate models, Water Resour Res, 47, W00G02, 10.1029/2010wr009829, 2011.

~~Sun, G., Chen, Y., Li, W., Pan, C., Li, J., and Yang, Y.: Spatial distribution of the extreme hydrological events in Xinjiang, north-west of China, Nat Hazards, 67, 483-495, 2013.~~

Terink, W., Hurkmans, R., Torfs, P., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, Hydrology & Earth System Sciences, 14, 687-703, 2010.

Teutschbein, C., and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J Hydrol, 456, 12-29, 10.1016/j.jhydrol.2012.05.052, 2012.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical‐statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530-1544, 2011.

Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Climatic Change, 112, 449-468, 2012.

Thom, H. C.: A note on the gamma distribution, Mon Weather Rev, 86, 117-122, 1958.

Wang, H., Chen, Y., Li, W., and Deng, H.: Runoff responses to climate change in arid region of northwestern China during 1960–2010, Chinese Geographical Science, 23, 286-300, 2013.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-variable error correction of regional climate models, Climatic Change, 120, 871-887, 2013.

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., and Wei, M.: Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, Journal of Geophysical Research: Atmospheres, 118, 4326-4347, 2013.

Xin, X., Wu, T., Li, J., Wang, Z., Li, W., and Wu, F.: How well does BCC_CSM1. 1 reproduce the 20th century climate change over China, Atmos. Oceanic Sci. Lett, 6, 21-26, 2013.

Xu, C., Chen, Y., Chen, Y., Zhao, R., and Ding, H.: Responses of Surface Runoff to Climate Change and Human Activities in the Arid Region of Central Asia: A Case Study in the Tarim River Basin, China, Environ Manage, 51, 926-938, 2013.

Table 1. Sensitivity indices of the five meteorological variables based on the Sobol' method.

Factor[a]	Meaning	Factor Range	Main effect S_i (%)	Total effect S_{Ti} (%)
a__tmp	Additive change to temperature	[-5,5]	15.0	36.9
r__pcp	Relative change to precipitation	[-0.5,0.5]	44.0	74.0
r__hmd	Relative change to humidity	[-0.5,0.5]	0.0	0.0
r__slr	Relative change to solar radiation	[-0.5,0.5]	7.7	22.6
r__wnd	Relative change to wind speed	[-0.5,0.5]	0.3	0.9

[a] Here, '*a*__' or '*r*__' means an addictive or a relative change to the initial parameter values.

Table 2. Bias correction methods for RCM-simulated precipitation and temperature.

Bias correction for precipitation	Bias correction for temperature
Linear Scaling (LS)	Linear Scaling (LS)
LOCal Intensity scaling (LOCI)	VARIance scaling (VARI)
Power Transformation (PT)	Distribution Mapping for temperature using Gaussian distribution (DM)
Distribution Mapping for precipitation using Gamma distribution (DM)	
Quantile Mapping (QM)	

~~**Table 3.** Performances of simulated streamflows driven by observed (default), raw RCM-simulated (raw), and 15 combinations of bias-corrected precipitation and temperature during the period 1986 ~ 2001. For all combinations, solar radiation is corrected with Linear Scaling (LS) method. (Values are given with one decimal except for NS).~~

	~~Bias correction method~~		~~Daily~~				~~Monthly~~			
	~~Precipitation~~	~~Temperature~~	~~NS~~	~~P_{BIAS}~~	~~R^2~~	~~MAE (m^3/s)~~	~~NS~~	~~P_{BIAS}~~	~~R^2~~	~~MAE (m^3/s)~~
~~default~~	~~obs~~	~~obs~~	~~0.80~~	~~4.3~~	~~0.8~~	~~24.2~~	~~0.90~~	~~4.3~~	~~0.9~~	~~16.6~~
~~raw~~	~~raw~~	~~raw~~	~~-44.91~~	~~420.5~~	~~0.4~~	~~487.9~~	~~-53.35~~	~~421.1~~	~~0.6~~	~~487.0~~
~~1~~	~~LS~~	~~LS~~	~~-2.65~~	~~115.6~~	~~0.5~~	~~136.4~~	~~-3.10~~	~~115.8~~	~~0.7~~	~~134.0~~
~~2~~	~~LS~~	~~VARI~~	~~-2.43~~	~~112.7~~	~~0.5~~	~~133.3~~	~~-2.87~~	~~113.0~~	~~0.7~~	~~130.6~~
~~3~~	~~LS~~	~~DM~~	~~-2.43~~	~~112.7~~	~~0.5~~	~~133.3~~	~~-2.87~~	~~113.0~~	~~0.7~~	~~130.6~~
~~4~~	~~LOCI~~	~~LS~~	~~0.49~~	~~-3.7~~	~~0.5~~	~~35.9~~	~~0.69~~	~~-3.7~~	~~0.7~~	~~25.3~~
~~5~~	~~LOCI~~	~~VARI~~	~~0.50~~	~~-4.5~~	~~0.5~~	~~35.6~~	~~0.69~~	~~-4.4~~	~~0.7~~	~~25.4~~
~~6~~	~~LOCI~~	~~DM~~	~~0.50~~	~~-4.5~~	~~0.5~~	~~35.6~~	~~0.69~~	~~-4.4~~	~~0.7~~	~~25.4~~
~~7~~	~~PT~~	~~LS~~	~~0.37~~	~~1.1~~	~~0.4~~	~~40.1~~	~~0.62~~	~~1.1~~	~~0.6~~	~~28.7~~
~~8~~	~~PT~~	~~VARI~~	~~0.38~~	~~0.3~~	~~0.4~~	~~39.8~~	~~0.63~~	~~0.3~~	~~0.6~~	~~28.6~~
~~9~~	~~PT~~	~~DM~~	~~0.38~~	~~8.3~~	~~0.5~~	~~41.2~~	~~0.62~~	~~8.3~~	~~0.7~~	~~30.6~~
~~10~~	~~DM~~	~~LS~~	~~0.40~~	~~7.5~~	~~0.5~~	~~40.7~~	~~0.63~~	~~6.7~~	~~0.6~~	~~30.3~~
~~11~~	~~DM~~	~~VARI~~	~~0.40~~	~~7.5~~	~~0.5~~	~~40.7~~	~~0.63~~	~~5.9~~	~~0.6~~	~~30.3~~
~~12~~	~~DM~~	~~DM~~	~~0.38~~	~~0.3~~	~~0.4~~	~~39.8~~	~~0.63~~	~~5.9~~	~~0.6~~	~~28.6~~
~~13~~	~~QM~~	~~LS~~	~~0.37~~	~~1.8~~	~~0.4~~	~~39.9~~	~~0.63~~	~~1.9~~	~~0.6~~	~~28.6~~
~~14~~	~~QM~~	~~VARI~~	~~0.38~~	~~1.0~~	~~0.4~~	~~39.5~~	~~0.63~~	~~1.0~~	~~0.6~~	~~28.4~~
~~15~~	~~QM~~	~~DM~~	~~0.38~~	~~1.0~~	~~0.4~~	~~39.5~~	~~0.63~~	~~1.0~~	~~0.6~~	~~28.4~~

Table 3. Performances of simulated streamflows driven by raw RCM simulated ("raw") and 15 combinations of bias-corrected precipitation and temperature (donated as numbers from 1 to 15) compared to the simulation driven by observed climate ("default") during the period 1986 ~ 2001. For simulations 1 to 15, solar radiation is corrected with Linear Scaling (LS) method.

	Bias correction method		Daily				Monthly			
	Precipitation	Temperature	NS (-)	P_{BIAS} (%)	R^2 (-)	MAE (m^3/s)	NS (-)	P_{BIAS} (%)	R^2 (-)	MAE (m^3/s)
raw	raw	raw	-47.69	398.9	0.4	547.5	-56.34	399.4	0.6	524.6
1	LS	LS	-2.66	106.2	0.5	150.1	-3.09	106.4	0.7	140.2
2	LS	VARI	-2.43	103.5	0.5	145.4	-2.85	103.7	0.7	135.9
3	LS	DM	-2.43	103.5	0.5	145.4	-2.85	103.7	0.7	135.9
4	LOCI	LS	0.49	-8.0	0.5	56.0	0.70	-7.9	0.7	38.2
5	LOCI	VARI	0.50	-8.6	0.5	55.6	0.70	-8.6	0.7	38.1
6	LOCI	DM	0.50	-8.6	0.5	55.6	0.70	-8.6	0.7	38.1
7	PT	LS	0.38	-3.3	0.4	61.7	0.64	-3.3	0.7	41.4
8	PT	VARI	0.39	-4.1	0.5	61.3	0.65	-4.1	0.7	41.1
9	PT	DM	0.39	-4.1	0.5	61.3	0.65	-4.1	0.7	41.1
10	DM	LS	0.41	3.6	0.5	60.3	0.66	3.6	0.7	40.5
11	DM	VARI	0.42	2.8	0.5	59.5	0.67	2.9	0.7	40.0
12	DM	DM	0.42	2.8	0.5	59.5	0.67	2.9	0.7	40.0
13	QM	LS	0.39	-2.6	0.5	61.3	0.65	-2.6	0.7	40.9
14	QM	VARI	0.40	-3.4	0.5	60.8	0.65	-3.4	0.7	40.7
15	QM	DM	0.40	-3.4	0.5	60.8	0.65	-3.4	0.7	40.7

~~**Table 4.** Frequency-based statistics of daily observed (“obs”), raw RCM-simulated (“raw”) and bias-corrected precipitations at the Bayanbulak Station (values are given with two decimal digits).~~

	~~Mean (mm)~~	~~Median (mm)~~	~~Standard deviation (mm)~~	~~90^{th} percentile (mm)~~	~~Probability of wet days (%)~~	~~Intensity of wet day (mm)~~
~~obs~~	~~0.73~~	~~0.00~~	~~2.44~~	~~1.90~~	~~32~~	~~2.30~~
~~raw~~	~~2.87~~	~~1.44~~	~~4.09~~	~~7.44~~	~~86~~	~~3.34~~
~~LS~~	~~0.73~~	~~0.20~~	~~1.53~~	~~2.10~~	~~73~~	~~1.00~~
~~LOCI~~	~~0.73~~	~~0.00~~	~~1.70~~	~~2.40~~	~~32~~	~~2.29~~
~~PT~~	~~0.73~~	~~0.00~~	~~2.44~~	~~1.80~~	~~32~~	~~2.30~~
~~DM~~	~~0.78~~	~~0.00~~	~~2.30~~	~~2.11~~	~~32~~	~~2.46~~
~~QM~~	~~0.73~~	~~0.00~~	~~2.44~~	~~1.90~~	~~32~~	~~2.31~~

Table 4. Frequency-based statistics of daily observed (“obs”), raw RCM-simulated (“raw”) and bias-corrected precipitations at the Bayanbulak Station ~~(values are given with two~~ one ~~decimal~~ place~~digits).~~

	Mean (mm)	Median (mm)	Standard deviation (mm)	99^{th} percentile (mm)	Probability of wet days (%)	Intensity of wet day (mm)
obs	0.73	0.0	2.4	12.4	32	2.3
raw	2.87	1.4	4.1	19.7	86	3.3
LS	0.73	0.2	1.5	7.6	73	1.0
LOCI	0.73	0.0	1.7	8.1	32	2.3
PT	0.73	0.0	2.4	11.4	32	2.3
DM	0.78	0.0	2.3	11.5	32	2.5
QM	0.73	0.0	2.4	12.4	32	2.3

Table 5. Frequency-based statistics (unit: ℃) of daily observed ("obs"), raw RCM simulated ("raw") and bias corrected maximum temperatures at the Bayanbulak Station ~~(values are given with two decimal places).~~

	Mean	Median	Standard deviation	10th percentile	90th percentile
obs	3.08	7.20	14.50	-18.70	19.20
raw	3.45	3.21	10.88	-10.34	17.90
LS	3.08	6.65	14.14	-17.33	19.40
VARI	3.08	6.85	14.50	-17.76	19.36
DM	3.08	6.85	14.50	-17.76	19.36

Table 6. Time-series based metrics of bias-corrected precipitation and temperature calculated on a monthly scale at the Bayanbulak Station ~~(values are given with two decimal places).~~

		NS (-)	P_{BIAS} (%)	R^2 (-)	MAE (mm or ℃)
Precipitation	raw	-6.78	293.28	0.42	65.40
	LS	0.64	0.06	0.65	9.66
	LOCI	0.61	-0.71	0.64	10.14
	PT	0.42	-0.09	0.53	11.98
	DM	0.46	6.64	0.56	11.78
	QM	0.44	0.03	0.54	11.99
Temperature	raw	0.84	15.78	0.88	4.31
	LS	0.95	3.04	0.95	2.35
	VARI	0.94	4.78	0.94	2.52
	DM	0.94	4.74	0.94	2.52

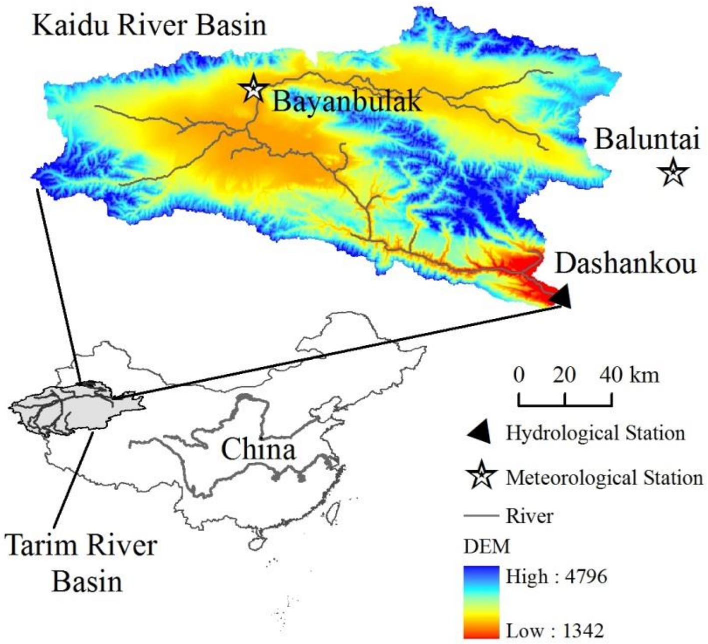

Fig. 1. Location of the study area, two meteorological stations and one hydrological station.

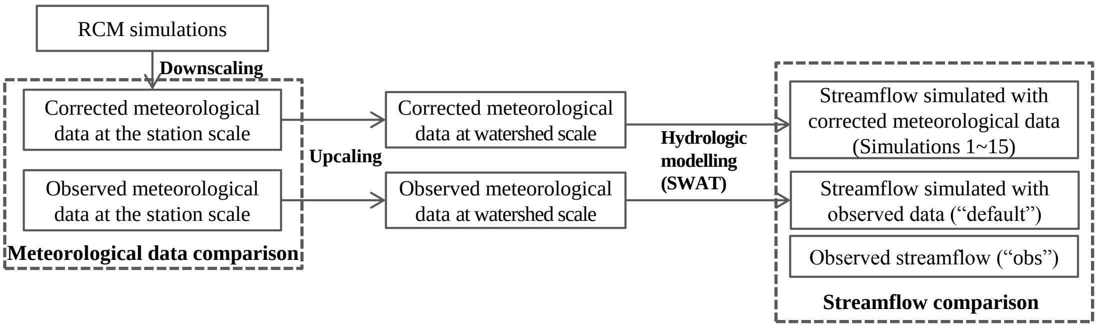

Fig. 2. Flow chart of comparison procedure

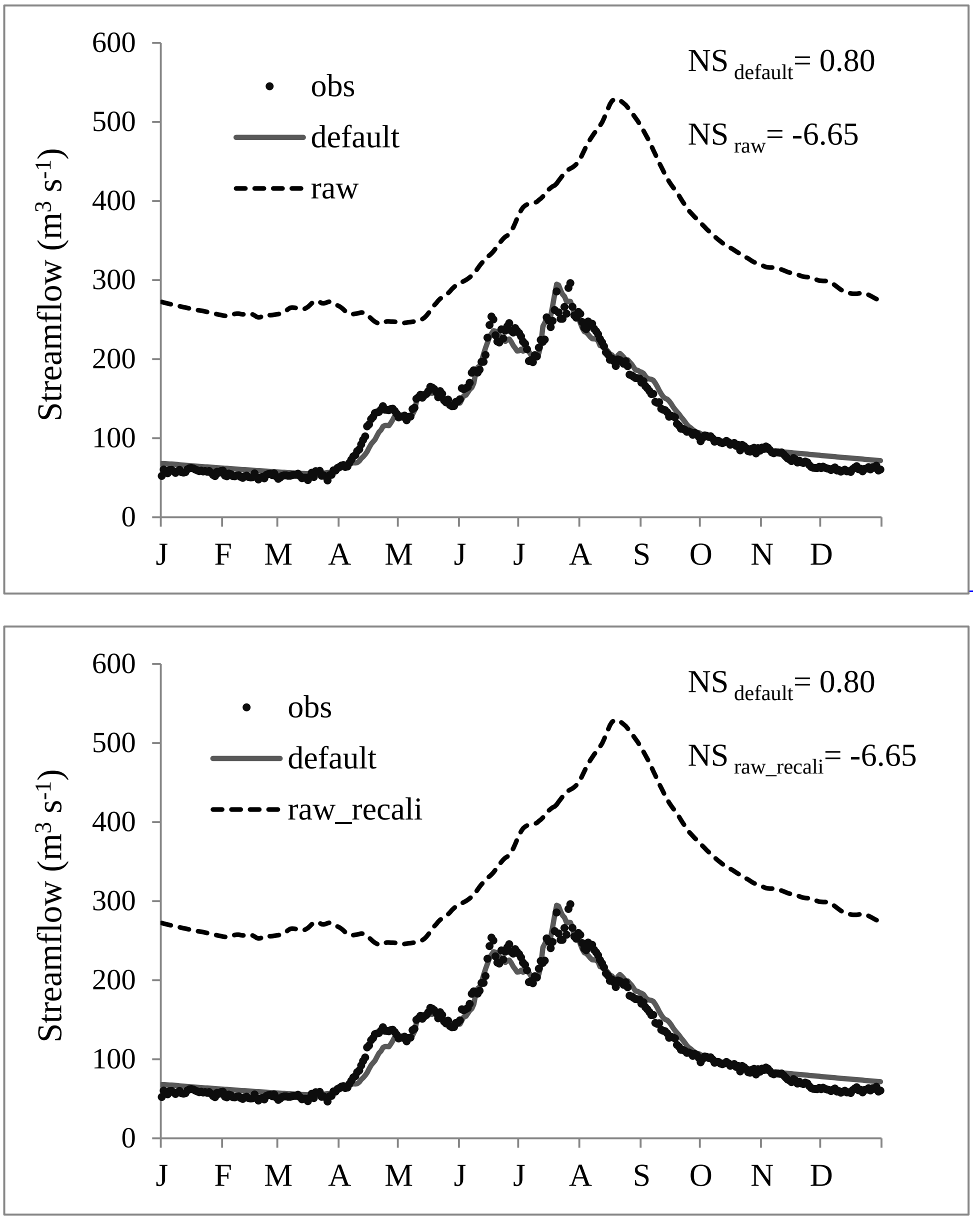

Fig. 3. Mean annual hydrographs of observed streamflow ("obs") and simulated streamflow using observed meteorological data ("default") during the period of 1986 ~ 2001 at the Dashankou Station. The simulated streamflow using raw RCM-simulated meteorological data after re-calibration ("raw_recali") is also plotted. The NS values are for the daily continuous data and not for the mean hydrograph.

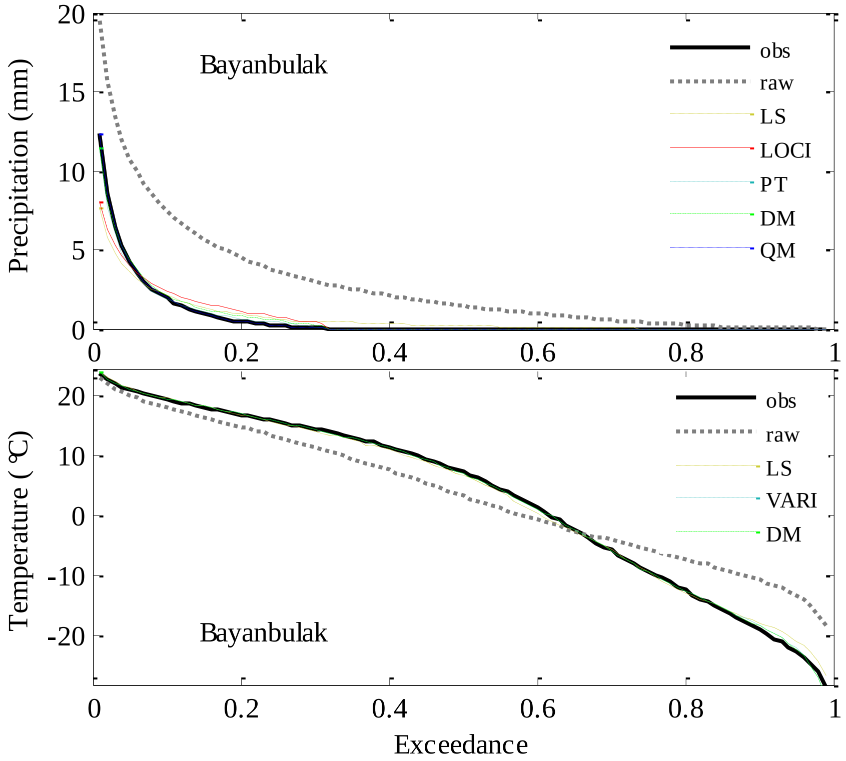

Fig. 4. Exceedance probabilities of the observed (“obs”), raw, and bias-corrected precipitation (top) and temperature (bottom) at the Bayanbulak Station.

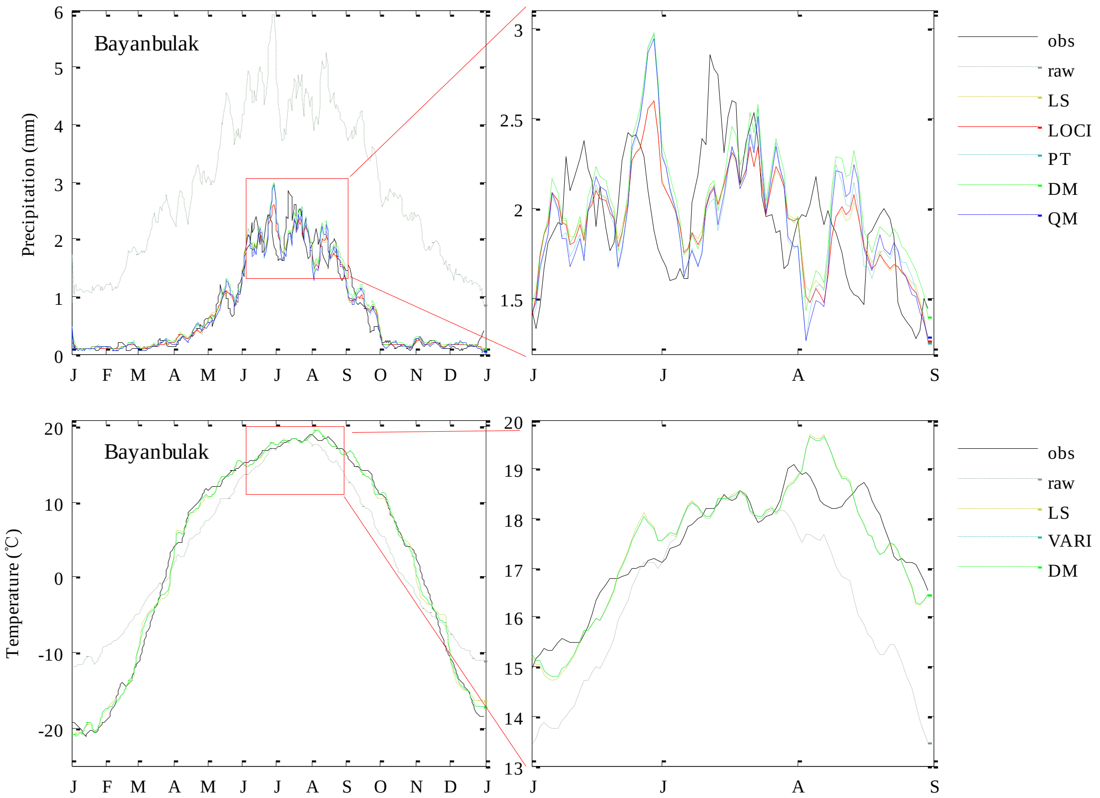

Fig. 5. ~~Average~~ Daily mean precipitation and temperature ~~hydrographs~~ of observed ("obs"), raw RCM simulated ("raw"), and bias corrected values at Bayanbulak Station, which were smoothed with 7-day moving average method. The precipitation and temperature during May to August is amplified to inspect the performance of each correction method.

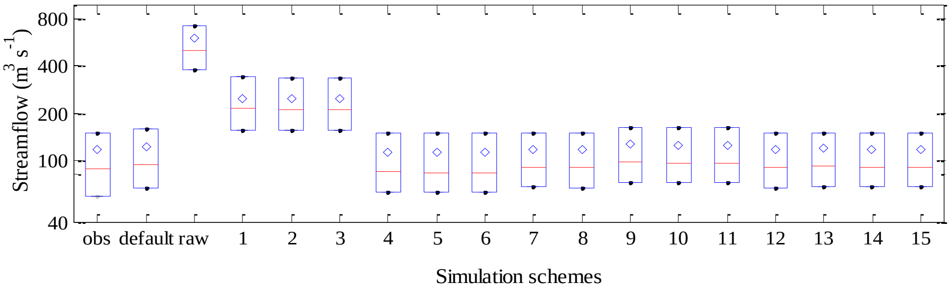

Fig. 6. Box plots of observed (“obs”) and simulated daily streamflows using observed (”default”), raw RCM-~~-~~simulated (“raw”) and corrected meteorological data (~~numbers from 1 to 15; see Table 3 for setup of these~~ setup of ~~15~~ simulations 1to 15 are listed in Table 3). ~~Solid boxes signify values from 1st to 3rd quantile while the median value is shown in the interior of the box, and t~~The mean values are shown with diamonds.

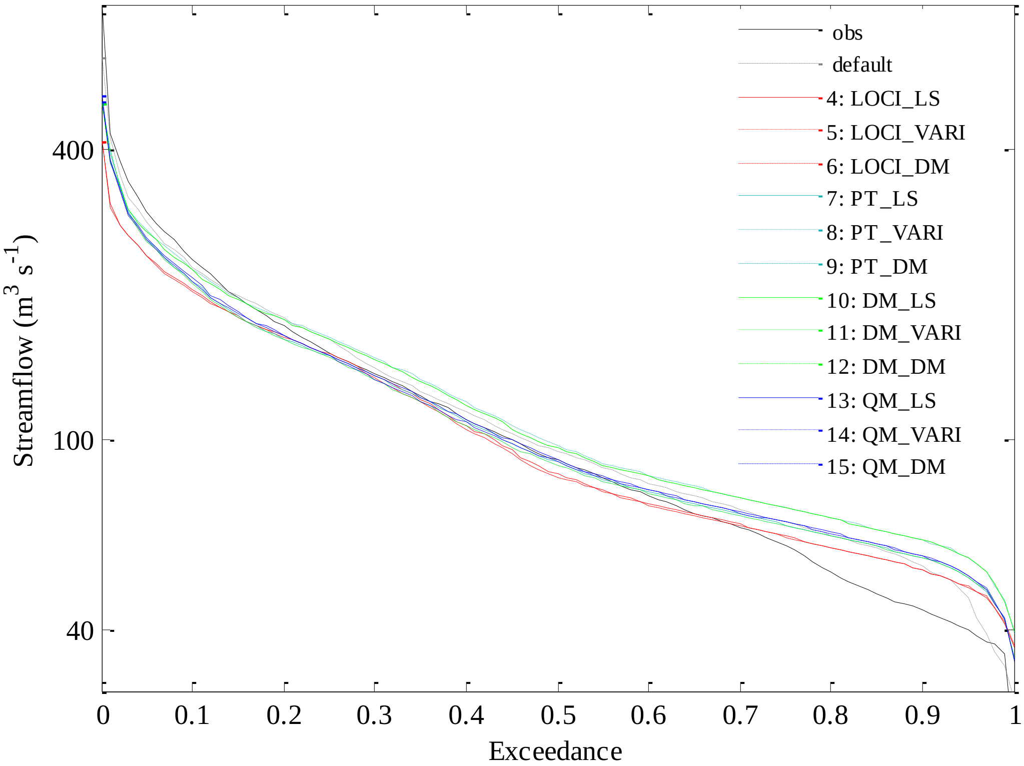

Fig. 7. Exceedance probability curves of observed (“obs”) and simulated streamflow driven by observed (“default”), and bias-corrected meteorological data (numbers from 4 to 15; ~~also~~ see Table 3 for detailed setup of these ~~12~~ simulations). ~~For plotting purpose, simulations “raw” and 1 to 3 are not shown.~~

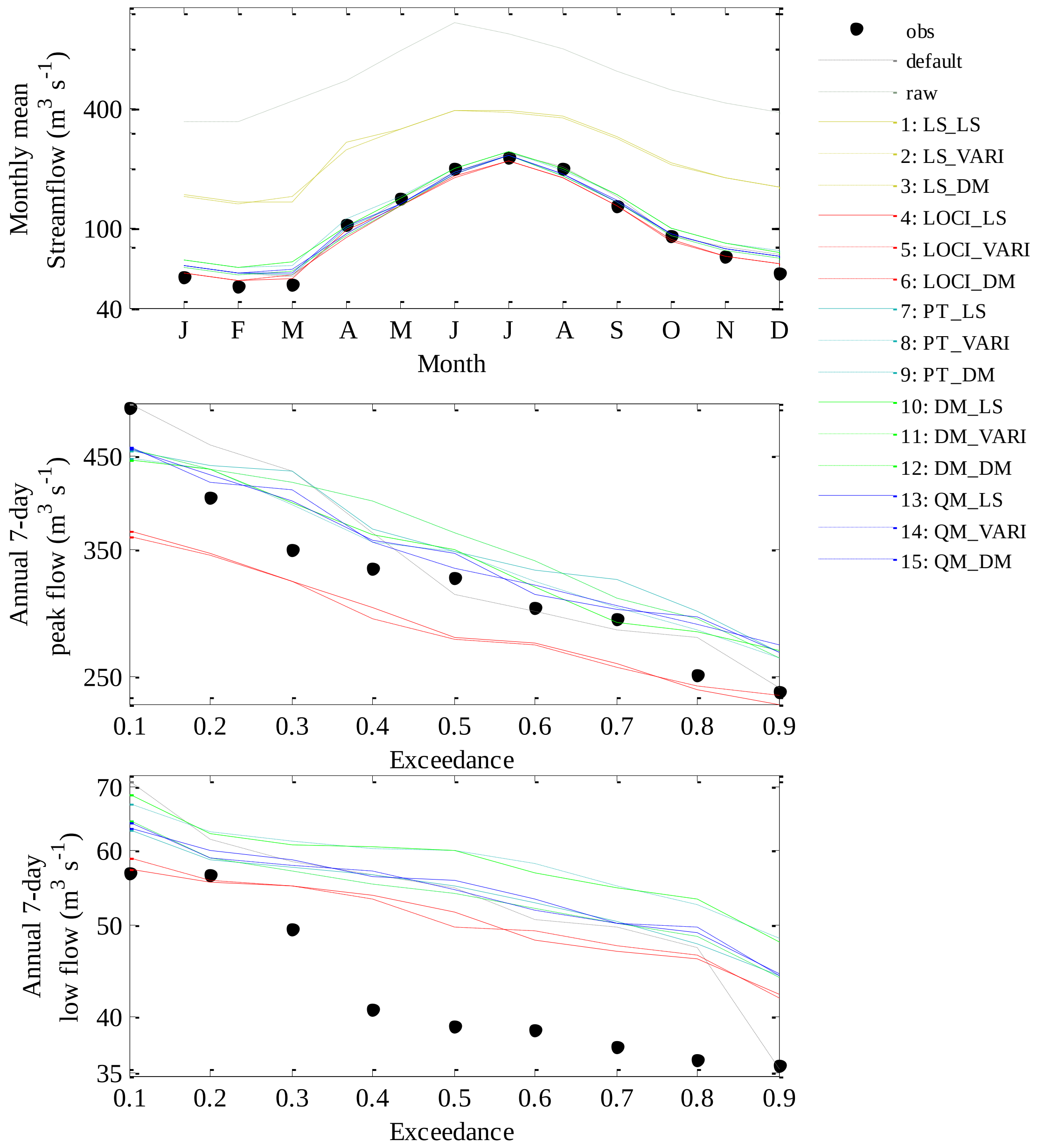

Fig. 8. Monthly mean streamflow (top) and exceedance probability curves of annual 7-day peak flow (middle) and annual 7-day low flow (bottom) during 1986 ~ 2001 in the Kaidu River Basin (obs: observed streamflow; default: simulated with observed meteorological data; raw: simulated with RCM simulated meteorological data; 1~15: simulated with corrected RCM meteorological data listed in Table 3). ~~The observation (“obs”), and simulated streamflows using observed (“default”), raw RCM-simulated (“raw”) and bias-corrected (numbers from 1 to 15; also see Table 3 for detail~~ed ~~setup of these 15 simulations) meteorological data are also shown in the monthly mean plot. For peak flow and low flow,~~ t~~he raw and simulations 1 to 3 are not shown as they are heavily biased.~~