Simulation of temperature extremes in the Tibetan Plateau from CMIP5 models and comparison with gridded observations

Qinglong You [1]*, Zhihong Jiang[1], Dai Wang[1], Nick Pepin[2], Shichang Kang[3]

1. Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/ Joint International Research Laboratory of Climate and Environmental Change (ILCEC)/ Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology (NUIST), Nanjing, 210044, China;
2. Department of Geography, University of Portsmouth, PO1 3HE, U.K.
3. State Key Laboratory of Cryospheric Science, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China;

* Corresponding author E-mail address: yqingl@126.com

Resubmitted to Climate Dynamics, May 23, 2017

Abstract: Understanding changes in temperature extremes in a warmer climate is of great importance for society and for ecosystem functioning due to potentially severe impacts of such extreme events. In this study, temperature extremes defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) from CMIP5 models are evaluated by comparison with homogenized gridded observations at 0.5° resolution across the Tibetan Plateau (TP) for 1961-2005. Using statistical metrics, the models have been ranked in terms of their ability to reproduce similar patterns in extreme events to the observations. Four CMIP5 models have good performance (BNU-ESM, HadGEM2-ES, CCSM4, CanESM2) and are used to create an optimal model ensemble (OME). Most temperature extreme indices in the OME are closer to the observations than in an ensemble using all models. Best performance is given for threshold temperature indices and extreme/absolute value indices are slightly less well modelled. Thus the choice of model in the OME seems to have more influences on temperature extreme indices based on thresholds. There is no significant correlation between elevation and modelled bias of the extreme indices for both the optimal/all model ensembles. Furthermore, the minimum temperature (Tmin) is significanlty positive correlations with the longwave radiation and cloud variables, respectively, but the Tmax fails to find the correlation with the shortwave radiation and cloud variables. This suggests that the cloud-radiation differences influence the Tmin in each CMIP5 model to some extent, and result in the temperature extremes based on Tmin.

Key words: Tibetan Plateau; Temperature extreme; CMIP5; Observation

1. Introduction

According to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5) [*IPCC*, 2013], the globally averaged combined land and ocean surface temperature has shown a warming of 0.85 ℃ (0.65-1.06) over the period 1880-2012 [*IPCC*, 2013]. A warming climate has been shown to exacerbate climate extremes, which can be of particular relevance to society and ecosystems due to their severe impacts [*Coumou and Rahmstorf*, 2012; *Easterling et al.*, 2000; *IPCC*, 2013; *Rahmstorf et al.*, 2007]. Correspondingly, the demand for understanding and modelling future changes in climate extremes has increased in recent years [*IPCC*, 2013; *Sillmann et al.*, 2013a; *Sillmann et al.*, 2013b]. The Expert Team on Climate Change Detection and Indices (ETCCDI) (http://cccma.seos.uvic.ca/ETCCDI) has developed a set of indices to quantify extremes and thus facilitate an understanding of observed change [*IPCC*, 2007; 2013; *Peterson and Manton*, 2008]. These indices were widely used in IPCC AR4 [*IPCC*, 2007] and AR5 [*IPCC*, 2013].

The ETCCDI indices have been analyzed based on observational records [*Aguilar et al.*, 2005; *Alexander et al.*, 2006], reanalyses [*Fang et al.*, 2008; *You et al.*, 2014], and future climate modelling projections [*Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012; *Kharin et al.*, 2013; *Sillmann et al.*, 2013a; *Sillmann et al.*, 2013b]. Many studies have been applied at the global scale [*Alexander et al.*, 2006; *Donat et al.*, 2013; *Frich et al.*, 2002]; but also at continental scales (such as Africa [*Aguilar et al.*, 2009; *New et al.*, 2006], America [*Peterson et al.*, 2008] and Europe [*E.M. Fischer and Schaer*, 2010; *Sillmann and Croci-Maspoli*, 2009]), and regional scales (such as China [*Ren et al.*, 2011; *You et al.*, 2011; *Zhai and Pan*, 2003], the Tibetan Plateau [*You et al.*, 2008], the Asia-

Pacific Network region [*Choi et al.*, 2009] and Russia [*Bulygina et al.*, 2007]). At the global scale increases in the number of warm days/nights and decreases in the number of cold days/nights are not in dispute [*IPCC*, 2013].

Climate models have improved since IPCC AR4, and can now reproduce observed continental-scale surface temperature patterns fairly accurately, along with past trends including the rapid warming since the mid-20th century and the cooling immediately following large volcanic eruptions [*IPCC*, 2013]. Therefore models are now being used to project changes in climate extremes [*Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012; *Sillmann et al.*, 2013a; *Sillmann et al.*, 2013b; *Sillmann and Roeckner*, 2008; *T Yang et al.*, 2012]. In IPCC AR5 for example, the Coupled Model Intercomparison Project Phase 5 (CMIP5) [*Taylor et al.*, 2012] has produced a freely available multi-model dataset which has allowed evaluation of ETCCDI indices at the global scale [*Sillmann et al.*, 2013a; *Sillmann et al.*, 2013b; *Sillmann and Roeckner*, 2008]. However there are still limitations in accurately simulating regional extremes [*Easterling et al.*, 2000]. CMIP5 model discrepancies in simulating cold extremes are generally larger than those for warm extremes, and there are larger uncertainties in the tropics and subtropics [*Kharin et al.*, 2013].

No previous study has specifically addressed climate extremes on the Tibetan Plateau (TP). The TP is over 4000 m above sea level and is surrounded by large mountain ranges (i.e. the Kunlun, Qilian, Hengduan, and Karakoram). All 14 of the world's peaks over 8000m are found in the TP, and 6 of the most important rivers in the world, including the Yellow, Yangtze and Yuarlung Zangbo rivers. These feed millions of people in downstream regions [*Guo et al.*, 2016; *Kuang and Jiao*, 2016; *T Yang et al.*, 2012; *You et al.*, 2011; *You et al.*, 2016; *You et al.*, 2014]. It is therefore pivotal to understand changes in extremes over the TP [*Duan and Xiao*, 2015; *Guo et al.*, 2016; *Kuang and Jiao*, 2016; *Yan et al.*, 2016; *You et al.*, 2016]. In this study, we examine changes in temperature extremes across the plateau using CMIP5 model ensembles and compare the results with gridded observations. Such studies are essential to improve knowledge on simulations of climate extremes in the plateau region.

2. Data and methods

Homogenized daily mean (T_{mean}), maximum (T_{max}) and minimum temperatures (T_{min}) are provided at 0.5° resolution by the National Climate Center of China Meteorological Administration (NCC/CMA). Values are interpolated using an "anomaly approach" from over 2400 stations [*Wu and Gao*, 2013; *Xu et al.*, 2009]. A 30-year Tmean, Tmax and Tmin for 1971–2000 are calculated for each Julian date at each station, and further extension of the dataset can be conducted directly based on this climatology without having to recalculate it every time. Stations with more than 1/3 (10 years) missing data are excluded from the analysis [*Wu and Gao*, 2013; *Xu et al.*, 2009]. This dataset has been widely used to validate regional and global atmospheric model simulations of extreme climate indices in past studies [*Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012; *You et al.*, 2015].

The CMIP5 Project represents the latest and most ambitious coordinated international climate model intercomparison exercise [*Taylor et al.*, 2012]. Table 1 lists CMIP5 models used in this study. Further model details and information on their configuration or features can be found in the Program for Climate Model Diagnosis and Intercomparison (PCMDI) data portal (http://www-pcmdi.llnl.gov/) [*Taylor et al.*, 2012]. Outputs from the 'historical' simulations of these CMIP5 models were used by the PCMDI in IPCC AR5 [*IPCC*, 2013]. In this study daily T_{mean}, T_{max} and T_{min} simulations and observations covering 1961-2005 are selected and interpolated to a common 2.5×2.5° grid using a bi-linear interpolation procedure (http://code.zmaw.de/projects/cdo).

Sixteen indices of temperature extremes (Table 2), including some of the ETCCDI indices are used to assess intensity, frequency and duration of climate extreme events [*Aguilar et al.*, 2009; *Aguilar et al.*, 2005; *Alexander et al.*, 2006; *Donat et al.*, 2013; *Peterson et al.*, 2008; *Sillmann et al.*, 2013a; *You et al.*, 2011]. Detailed descriptions are provided in Table 2 (also see http://cccma.seos.uvic.ca/ETCCDI). 17 CMIP5 model simulations of 16 indices are chosen, and the root-mean-square error, the standard deviations and correlation betwee the model and observation are calculated. The comprehensive model rank (M_R) [*Chen et al.*, 2011; *Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012] which measures the consistency of simulations for each model is defined as:

$$M_R = 1 - \frac{1}{(1 \times m \times n)} \sum_{i=1}^{n} rank_i ,$$

Where m and n is the number of models and indices, and $rank_i$ is based on model's order of performance on each index. The M_R of the best-performing model is closer to 1, indicating higher skill [*Chen et al.*, 2011; *Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012]. Based on M_R the optimal models from 17 models are selected and the ensemble simulations were then performed. The temporal skill scores are calculaled as:

$(\frac{STDm}{STDo} - \frac{STDo}{STDm})^2$, where STDm and STDo denotes the interannual standard deviation of simulation and observations, respectively [*Chen et al.*, 2011; *Z Jiang et al.*, 2015; *Z Jiang et al.*, 2012].

The Mann-Kendall test for a trend and Sen's slope estimates are used to estimate trends [*Sen*, 1968]. This is a common method employed to compute trends in meteorological and climate extreme series [*Bulygina et al.*, 2007; *Choi et al.*, 2009; *You et al.*, 2011; *You et al.*, 2016; *Zhang et al.*, 2011]. A trend is statistically significant if $p<0.05$.

3. Results

3.1 Evaluation of temporal variability

Three assessment indices (the temporal correlation coefficient (a), the ratio of standard deviation (b) and the root-mean-square error (c) between observed and modelled extremes) are used to evaluate the ability of each model to simulate the 16 temperature extremes similar to the observed values (Figure 1). Correlation coefficients between observed and simulated extremes are nearly all positive (red cells in Figure 1a) for all 16 temperature extreme indices, and they reach over 0.5 for TXn, TN90p, TN10p and FD0 (see Table 2). This suggests that CMIP5 models can simulate much of the interdecadal variability of temperature extremes in the TP. Using the ratio of modelled to observed standard deviation (Figure 1b), a value closer to 1 means a more realistic model simulation. With the exception of duration indices such as TR20, WSDI and CSDI, most ratios are quite close to 1 and thus the models are fairly realistic. For root-mean-square errors (Figure 1c), many indices such as TNx, DTR and threshold indices such as TX90p and TN90p have fairly small values, indicating that these indices are captured relatively well by most CMIP5 models.

To synthesize the three assessment indices an M_R value is calculated for each model to illustrate their overall ranking (Figure 2). Each model is ranked from 1 (best) to 17 (worst) for each index. The length of the color column is the summary of each ranking and shorter columns mean a better model performance. The colors reprsent the ranking of each individual index. The top five CMIP5 models are MPI-ESM-MR, CCSM4, HadGEM2-ES, BNU-ESM, and GFDL-ESM2M, respectively.

3.2 Evaluation of spatial variability

The spatial success of each model in reproducing observed patterns of extreme indices can be assessed in a similar way using equivalent spatial statistics (Figure 3). In Figure 3a, the correlation coefficients between observed and modelled patterns of extremes are positive for some indices, especially DTR and WSDI. Hoewever there are also several indices with negative correlations such as TXx, TNx, SU25 and TR20. Thus compared with the temporal variability, the spatial variability of temperature extremes in the TP is only simulated well in some cases. However there are uncertainties in observations because of a lack ofstations in many sub-regions. For the ratio of modelled to observed standard deviation (Figure 3b), values near 1 are common. The exception is for TR20 which shows extremely high ratios. DTR, TX10p and TN10p are closest to 1. Root-mean-square errors are smallest for threshold indices such as TX90p and TN90p (Figure 3c) suggesting that most CMIP5 models are particularly good at simulating these. Duration indices such as SU25 and FD0 have larger root-mean-square errors.

A similar spatial ranking of overall model performance (Figure 4) shows the best models to be BNU-ESM, CanESM2, EC-EARTH, HadGEM2-ES, and ACCESS1.0, respectively.

3.3 A combined temporal and spatial ranking

The relationship between temporal and spatial ranks for each model is shown in Figure 5. Each dot represents a model, identified by its number on the right. The ranking is given a value between 0 and 1 for each model based on the three assessment indices. The correlation coefficient between the two is 0.448 meaning the inter-model consistency in simulating spatial pattern and inter-annual variability. Models closer to the top right of the diagram show better overall performance. The sum of the temporal and spatial ranking is shown in Figure 6, the top four models are: BUN-ESM (5), HadGEM2-ES (8), CCSM4 (10), and CanESM (11). These four will be defined as the optimal models. Two ensemble simulations were then performed: one with just the four optimal models, and one with all 17 models.

The difference in climatology of extreme indices between the optimal/all models ensembles and the observations are shown in Figure 7. Time series of individual indices from these three datasets (optimal/all models ensembles and observations) are represented in Figure 8. Trends and temporal skill scores for each index in each dataset are summarized in Table 3. Although patterns are complex, compared with the all models ensemble, the optimal models ensemble is shown to greatly reduce the gap between simulation and observations for both spatial and temporal patterns. This is particularly the case for the indices of TNn, SU25, TR20, WSDI and CSDI (Figures 7 and 8). The optimal model ensemble has good skill scores, and is lower than the all model ensemble score in 12 cases out of 16, showing that the optimal models ensemble is usually closest to the observations.

In order to understand the differences in the success of various CMIP5 models in simulating temperature extremes, five climate variables from each model, potentially influencing Tmax and Tmin, are selected. These are

1. the surface downwelling shortwave radiation (SDSR),
2. the SDSR at clear sky (SDSRcs),
3. the surface downwelling longwave radiation (SDLR),
4. the SDLR at clear sky (SDLRcs) and
5. the total cloud fraction (TCF).

Figure 9 shows the relationship between Tmax/Tmin and these variables for each CMIP5 model. For Tmax, there are no significant correlations with TCF, the difference between SDSRcs and SDSR, and SDSR, respectively (Figure 9a,b,c), which suggests that incoming energy balance is not simulated well and cannot account for changes in Tmax. This lack of correlation of Tmax with radiation parameters is inconsistent with previous studies which showed that CMIP5 model differences in DTR seemed to be significantly controlled by clouds, and longwave and shortwave fluxes on the global scale [*Lindvall and Svensson*, 2015].

Tmin on the other hand has significant positive correlations with TCF (R=0.34), the SDLR-SDLRcs (R=0.39) and SDLR (R=0.71), indicating that nightime cloud-radiation differences are a partial control on Tmin in most CMIP5 models. Differences in TCF, SDSRcs-SDSR, and SDLR-SDLRcs between models are related to differences in aerosol loadings.

The relationships between elevation and bias (optimal/all model ensembles minus observations) in simulations of temperature extremes are shown in Figure 10. Elevations are calculated from the 90 × 90 m SRTM (Shuttle Radar Topography Mission) DEM from the International Scientific and Technical Data Mirror Site (http://www.gscloud.cn). There is no significant correlation between elevation and any bias and thus no elevational dependancy in any bias of temperature extreme indices in the model ensembles.

4. Discussion and Conclusions

In recent decades, climate extremes have attracted much attention because of disproportionate impacts on society and ecosystems [*IPCC*, 2013]. We have examined changes in temperature extremes over the TP using standard indices defined by ETCCDI from CMIP5 models and compared these changes with those based on observations. It is informative to compare our results with past global studies to set changes in the TP in broader context. In particular it is of interest whether indices are changing in a similar way to the global scale. Since there are four main types of index: a) relative (percentile based), b) absolute, c) threshold and d) duration, we start by discussing each in turn, before considering more broad diurnal contrasts. The most comprehensive global analysis of trends in extremes in CMIP5 model simulations is that of Sillmann et al. (2013a) – hereafter S13, but unfortunately global trend magnitudes for each index are not defined in this paper which makes a direct quantitative comparison of our results difficult.

In our study the relative indices based on observations show a decrease in cold days and nights (TX10p/TN10p) and increase in warm days and nights (TX90p/TN90p). All these are consistent with warming in the same indices reported by S13 but similar patterns have also been shown in equivalent analyses of observations on a global scale [*Alexander et al.*, 2006; *Frich et al.*, 2002]. Both optimal and all ensemble models also show trends in the relative indices in our study but they are smaller in magnitude than for the observations. The difference is particularly noticeable for TN10p and TN90p where the models fail to match the rapid nighttime warming in observations over the plateau.

Previous global studies have also indicated an intensification in absolute temperature indices (TXn/TNn and TXx/TNx) in observations [*Seneviratne et al.*, 2012; *Vose et al.*, 2005], reanalyses [*You et al.*, 2013], and model simulations [*Kharin et al.*, 2013; *Rahmstorf et al.*, 2007; *Sillmann et al.*, 2013a; *Sillmann et al.*, 2013b]. In our study all absolute indices are increasing which agrees with the S13. TNn tends to have the strongest warming in the observations but TNx has in the models.

Threshold indices (FD0, ID0, SU25, TR20) can have great influence on ecosystems and human infrastructure, and small changes in the indices can have relatively large impacts [*Kang et al.*, 2010; *Kharin et al.*, 2013; *Peterson and Manton*, 2008; *Peterson et al.*, 2008; *You et al.*, 2013; *You et al.*, 2008]. Global trends in S13 show a decrease in FD0 and increase in TR20 (others not reported). Over the TP, frost days (FD0) and ice days (ID0) show rapid decreases in the observations but this is not picked up by the model ensembles. The ensembles even simulate weak increases, the reasons for which require more research.

Finally, changes in duration indices (GSL, WSDI and CSDI) are also variable. On a global scale in S13 WSDI is increasing, sometimes significantly and CSDI decreasing (albeit at a slower rate). Decreasing cold spell and increasing warm spell lengths also occur in both the observations and model ensembles in the TP, and again the increase in warm spell duration is particularly strong. Thus the TP is broadly representative of global trends, and the high elevation does not mitigate against the rapid increase in warm spells. There is however a discrepancy in our study in terms of growing season length which decreases in the model ensembles but increases in the observations. In summary the signs of the trends in most indices over the TP are in agreement with global trends reported in S13.

Taken together the relative and absolute index changes in the observations imply that nighttime warming over the TP is much stronger than daytime warming, probably because the water vapour [*Rangwala et al.*, 2009] and radiative [*Ohmura*, 2012] feedbacks critical at high elevations are enhanced at lower air temperatures [*Rangwala et al.*, 2009; *Rangwala et al.*, 2013]. Numerous other studies have shown elevation-dependent warming whereby high elevations are warming more rapidly than the global mean [*Pepin and Coauthors*, 2015; *Vuille et al.*, 2015; *Wang et al.*, 2016]. However, any elevational signal is usually clearer in nighttime observations of Tmin in comparison to Tmax [*Rangwala and Miller*, 2012; *Yan and Liu*, 2014]. Interestingly however the CMIP5 model ensembles do not reflect this over the plateau. DTR is increasing in the model ensembles (albeit insignificantly) whereas it is strongly decreasing (-0.22˚C/decade) in the observations. The decreasing DTR may also partly be the reason why frost days are increasing and the growing season is shortening in the model ensembles. The cause of the lack of nighttime warming in comparison with daytime warming in the ensembles requires further investigation. One possible theory is that it is likely to be because the CMIP5 models in general are dominated by surface based (especially snow albedo) feedback mechanisms (which should be enhanced during the day) and less influenced by water vapour and Planck feedbacks (which should be enhanced at night). To start to appreciate the relative roles of various feedback mechanisms, we also investigated the relationship between cloud and radiation variables and daily maximum/minimum temperatures in the models (Figure 9). At night there are strong relationships, again suggesting that cloud-related feedbacks are a dominant control of nighttime trends in Tmin. Although water vapour and cloud feedbacks are still relevant during the day, the situation is more complex with additional surface albedo loops due to snow/ice retreat [*Kang et al.*, 2010] and vegetation changes [*D Jiang et al.*, 2011] also being strongly important. Cryospheric change in the TP such as the shrinking of glaciers and melting of frozen ground [*Kang et al.*, 2010; *K Yang et al.*, 2014; *K Yang et al.*, 2011; *You et al.*, 2016] will preferentially enhance daytime warming. For example, more than 80% of glaciers in western China have retreated, losing 4.5% of their areal coverage since 1951 [*Kang et al.*, 2010]. Vegetation is more complex since migration of treelines upslope could encourage warming through greening (in a similar way to the Arctic [*Chapin et al.*, 2005], but this is not happening everywhere and there is also degradation in vegetation through overgrazing which could introduce other moisture-related feedback loops. The added influence of surface feedback loops (snow, vegetation) and their seasonal dependence means that the relationship between Tmax and cloud variables probably depends on season and location.

The most successful models which formed part of the optimal ensemble were BNU-ESM, CanESM2, CCSM4 and HadGEM2-ES. A comprehensive review of model performance is available in IPCC (2013), where assessed models according to the rates of change of tropospheric temperature and precipitable water for the tropics (20˚S to 20˚N) – see Figure 9.9, p774 in IPCC (2013). All the models in the optimal ensemble apart from BNU-ESM for which there is no data, showed strong warming and wetting trends, indicative of stronger water vapour feedback. Thus models with strong tropical vapour feedback appear to do well in simulating temperature extremes over the TP, the reasons for which require more research. S13 also evaluated the success of all CMIP5 models on a global scale at simulating trends in extremes and it is informative to compare their results with ours. CCSM4 and HadGEM2-ES also performed well globally, but BNU-ESM and CanESM2 showed more variable performance, and the latter was not good for TXx and TNn.

A summary of individual feedbacks for each model in the CMIP5 experiment is presented in IPCC [2013]. Unfortunately it is difficult to find characteristics that stand out for the four models in the optimal ensemble, in comparison with the other models in this table. In part this is because a lot of models have missing data on vital feedbacks.

Equilibrium climate sensitivity tends to be high for the optimal models, particularly HadGEM-E2 which has the second highest of any model at 4.6˚C. However, model feedbacks including lapse rate (negative), surface albedo (positive) and cloud feedback (positive or negative) show no strong pattern for the four best models. The absence of any obvious strong model signature or characteristics which define a "successful" model means that much more work is required to understand the physical processes associated with temperature extremes at high elevations typical of the plateau, and subsequently what feedback mechanisms are most critical in creating a successful hindcast of temperature extremes.

Understanding the mechanisms by which extreme temperatures occur, especially at high elevations, is challenging. On a global scale, several explanations have been put forward to account for changing extremes which include changes in local and global SSTs [*Alexander et al.*, 2006], changes in large scale circulation patterns [*Kysely*, 2008], and the influence of land surface change [*IPCC*, 2013]. The last factor is particularly important in controlling daytime extremes. Successful modelling of soil-moisture and land-atmosphere coupling is required for a model to simulate the influence of soil moisture anomalies on high-temperature extremes for example, and energy partitioning (sensible vs latent heat) is a critical control [*E. M. Fischer and Knutti*, 2015]. Drier conditions and absence of soil moisture leads to greater extremes (both day and night) so long-term droughts (which maybe caused by persistent circulation anomalies) are an important factor. Any long-term degradation in vegetation on the plateau [*Kang et al.*, 2010] could therefore contribute to increased extremes and needs to be part of any model. Changes in atmospheric circulation can also modify temperature extremes and their spatial distribution [*Alexander et al.*, 2006; *You et al.*, 2011; *You et al.*, 2008]. In the TP for example cold air outbreaks imported from Siberia are associated with nearly all extremely low temperature episodes. Finally there is strengthened evidence for an influence of human activity on the observed frequency of extreme temperatures [*Coumou and Rahmstorf*, 2012; *E.M. Fischer and Knutti*, 2013; *Rahmstorf et al.*, 2007].

What is missing so far from the research into temperature extremes is an appreciation of how elevation itself could influence the various controlling factors and feedbacks discussed above. The high elevation environment is often thought of as naturally extreme, with a strong dependence of surface temperature on surface energy balance and a lack of atmosphere above to buffer response to direct radiation exchange. However it is not obvious how this natural tendency towards temperature extremes manifests itself in terms of past and future trends in extreme events. Recent research is beginning to uncover the forcing mechanisms of high elevation temperature change [*Pepin and Coauthors*, 2015] and critical to future understanding is an appreciation of elevation gradients in forcing due to snow albedo [*Giorgi et al.*, 1997] and vegetation [*D Jiang et al.*, 2011] feedbacks, water vapour and downwelling long wave radiation (Rangwala et al. 2009), the surface radiation/temperature feedback [*Ohmura*, 2012], clouds and latent heat release [*Rangwala and Miller*, 2012] and aerosols [*Xu et al.*, 2016]. Isolating the response to each forcing factor in future CMIP5 model runs is an important area for future high-elevation studies.

Acknowledgments. This study is supported by the National Key Research and Development Program of China (2016YFA0601700) and State Key Program of National Natural Science Foundation of China (41230528), Jiangsu Specially-Appointed Professor, Jiangsu Natural Science Funds for Distinguished Young Scholar “BK20140047”, the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), Jiangsu Shuang-Chuang Individual and Team Award. Observations are provided by the National Meteorological Information Center, China Meteorological Administration (NMIC/CMA) (http://cdc.cma.gov.cn), and the outputs from CMIP5 models participated in IPCC AR5 are available at http://pcmdi3.llnl.gov/esgcet/home.htm.

References

Aguilar, E., et al. (2009), Changes in temperature and precipitation extremes in western central Africa, Guinea Conakry, and Zimbabwe, 1955-2006, *Journal of Geophysical Research-Atmospheres*, *114*, D02115, doi:10.1029/2008jd011010.

Aguilar, E., et al. (2005), Changes in precipitation and temperature extremes in Central America and northern South America, 1961-2003, *Journal of Geophysical Research-Atmospheres*, *110*, D23107, doi:10.1029/2005jd006119.

Alexander, L. V., et al. (2006), Global observed changes in daily climate extremes of temperature and precipitation, *Journal of Geophysical Research-Atmospheres*, *111*, D05109, doi:10.1029/2005jd006290.

Bulygina, O. N., V. N. Razuvaev, N. N. Korshunova, and P. Y. Groisman (2007), Climate variations and changes in extreme climate events in Russia, *Environmental Research Letters*, *2*(4), 045020, doi:10.1088/1748-9326/2/4/045020.

Chapin, F. S., et al. (2005), Role of Land-Surface Changes in Arctic Summer Warming, *Science*, *310*(5748), 657.

Chen, W. L., Z. H. Jiang, and L. Li (2011), Probabilistic projections of climate change over China under the SRES A1B scenario using 28 AOGCMs, *Journal of Climate*, *24*(17), 4741-4756.

Choi, G., et al. (2009), Changes in means and extreme events of temperature and precipitation in the Asia-Pacific Network region, 1955-2007, *International Journal of Climatology*, *29*(13), 1906-1925, doi:10.1002/joc.1979.

Coumou, D., and S. Rahmstorf (2012), A decade of weather extremes, *Nature Climate Change*, *2*(7), 491-496, doi:10.1038/nclimate1452.

Donat, M. G., L. Alexander, H. Yang, I. Durre, R. S. Vose, R. Dunn, K. Willett, E. Aguilar, M. Brunet, and J. Caesar (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, *Journal of Geophysical Research-Atmospheres*, *118*, 2098-2118.

Duan, A. M., and Z. X. Xiao (2015), Does the climate warming hiatus exist over the Tibetan Plateau?, *Scientific Reports*, *5*, 13711.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O.

Mearns (2000), Climate extremes: Observations, modeling, and impacts, *Science*, *289*(5487), 2068-2074.

Fang, X. Q., A. Y. Wang, S. K. Fong, W. S. Lin, and J. Liu (2008), Changes of reanalysis-derived Northern Hemisphere summer warm extreme indices during 1948-2006 and links with climate variability, *Global and Planetary Change*, *63*(1), 67-78, doi:10.1016/j.gloplacha.2008.06.003.

Fischer, E. M., and R. Knutti (2013), Robust projections of combined humidity and temperature extremes, *Nature Climate Change*, *3*(2), 126-130.

Fischer, E. M., and R. Knutti (2015), Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes, *Nature Clim. Change*, *5*, 560-564, doi:10.1038/nclimate2617.

Fischer, E. M., and C. Schaer (2010), Consistent geographical patterns of changes in high-impact European heatwaves, *Nature Geoscience*, *3*(6), 398-403, doi:10.1038/ngeo866.

Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. Tank, and T. Peterson (2002), Observed coherent changes in climatic extremes during the second half of the twentieth century, *Climate Research*, *19*(3), 193-212.

Giorgi, F., J. W. Hurrell, M. R. Marinucci, and M. Beniston (1997), Elevation dependency of the surface climate change signal: a model study, *Journal of Climate*, *10*(2), 288-296.

Guo, D., E. Yu, and H. Wang (2016), Will the Tibetan Plateau warming depend on elevation in the future?, *Journal of Geophysical Research: Atmospheres*, *121*(8), 3969-

3978, doi:10.1002/2016JD024871.

IPCC (2007), Summary for Policymakers of 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, UK: Cambridge University Press*.

IPCC (2013), Summary for Policymakers of Climate change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change *Cambridge, UK: Cambridge University Press*.

Jiang, D., Y. Zhang, and X. Lang (2011), Vegetation feedback under future global warming, *Theoretical and Applied Climatology*, *106*(1), 211-227, doi:10.1007/s00704-011-0428-6.

Jiang, Z., W. Li, J. Xu, and L. Li (2015), Extreme Precipitation Indices over China in CMIP5 Models. Part I: Model Evaluation, *Journal of Climate*, *28*, 8603-8619.

Jiang, Z., J. Song, L. Li, W. Chen, Z. Wang, and J. Wang (2012), Extreme climate events in China: IPCC-AR4 model evaluation and projection, *Climatic Change*, *110*(1-2), 385-401, doi:10.1007/s10584-011-0090-0.

Kang, S. C., Y. W. Xu, Q. L. You, W. A. Flugel, N. Pepin, and T. D. Yao (2010), Review of climate and cryospheric change in the Tibetan Plateau, *Environmental Research Letters*, *5*(1), 015101, doi:10.1088/1748-9326/5/1/015101.

Kharin, V. V., F. Zwiers, X. Zhang, and M. F. Wehner (2013), Changes in temperature and precipitation extremes in the CMIP5 ensemble, *Climatic Change*, *119*(2), 345-357.

Kuang, X., and J. Jiao (2016), Review on climate change on the Tibetan Plateau during the last half century, *Journal of Geophysical Research: Atmospheres*, *121*(8), 3979-4007.

Kysely, J. (2008), Influence of the persistence of circulation patterns on warm and cold temperature anomalies in Europe: Analysis over the 20th century, *Global and Planetary Change*, *62*(1-2), 147-163, doi:10.1016/j.gloplacha.2008.01.003.

Lindvall, J., and G. Svensson (2015), The diurnal temperature range in the CMIP5 models, *Climate Dynamics*, *44*(1-2), 405-421.

New, M., et al. (2006), Evidence of trends in daily climate extremes over southern and west Africa, *Journal of Geophysical Research-Atmospheres, 111*, D14102, doi:10.1029/2005jd006289.

Ohmura, A. (2012), Enhanced temperature variability in high-altitude climate change, *Theoretical and Applied Climatology*, *110*(4), 499-508.

Pepin, N. C., and Coauthors (2015), Elevation-dependent warming in mountain regions of the world, *Nature Clim. Change, 5*, 424-430.

Peterson, T. C., and M. J. Manton (2008), Monitoring changes in climate extremes - A tale of international collaboration, *Bulletin of the American Meteorological Society*, *89*(9), 1266-1271, doi:10.1175/2008bams2501.1.

Peterson, T. C., X. B. Zhang, M. Brunet-India, and J. L. Vazquez-Aguirre (2008), Changes in North American extremes derived from daily weather data, *Journal of Geophysical Research-Atmospheres, 113*(D7), D07113, doi:10.1029/2007jd009453.

Rahmstorf, S., A. Cazenave, J. A. Church, J. E. Hansen, R. F. Keeling, D. E. Parker, and R. C. J. Somerville (2007), Recent climate observations compared to projections, *Science*, *316*(5825), 709-709, doi:10.1126/science.1136843.

Rangwala, I., and J. R. Miller (2012), Climate change in mountains: a review of elevation-dependent warming and its possible causes, *Climatic Change*, *114*(3-4), 527-547, doi:10.1007/s10584-012-0419-3.

Rangwala, I., J. R. Miller, and M. Xu (2009), Warming in the Tibetan Plateau: Possible influences of the changes in surface water vapor, *Geophys. Res. Lett.*, *36*, L06703, doi:L06703
10.1029/2009gl037245.

Rangwala, I., E. Sinsky, and J. R. Miller (2013), Amplified warming projections for high altitude regions of the northern hemisphere mid-latitudes from CMIP5 models, *Environmental Research Letters*, *8*(2), 024040.

Ren, G. Y., Z. Y. Guan, X. M. Shao, and D. Y. Gong (2011), Changes in climatic extremes over mainland China, *Climate Research*, *50*, 105-111.

Sen, P. K. (1968), Estimates of regression coefficient based on Kendall's tau, *Journal of the American Statistical Association*, *63*, 1379-1389.

Seneviratne, S. I., et al. (2012), Chapter 3: Changes in Climate Extremes and their Impacts on the Natural Physical Environment. SREX: Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, C. B. Field, et al., Eds., *Cambridge University Press*, 109-230.

Sillmann, J., and M. Croci-Maspoli (2009), Present and future atmospheric blocking and its impact on European mean and extreme climate, *Geophysical Research Letters*, *36*, L10702, doi:10.1029/2009gl038259.

Sillmann, J., V. Kharin, F. Zwiers, X. Zhang, and D. Bronaugh (2013a), Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, *Journal of Geophysical Research-Atmospheres*, *118*, 2473-2493.

Sillmann, J., V. V. Kharin, X. Zhang, F. Zwiers, and D. Bronaugh (2013b), Climate extreme indices in the CMIP5 multi-model ensemble. Part 1: Model evaluation in the present climate, *Journal of Geophysical Research-Atmospheres*, *118*, 1716-1733.

Sillmann, J., and E. Roeckner (2008), Indices for extreme events in projections of anthropogenic climate change, *Climatic Change*, *86*(1-2), 83-104, doi:10.1007/s10584-007-9308-6.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bulletin of the American Meteorological Society*, *93*(4), 485-498, doi:10.1175/bams-d-11-00094.1.

Vose, R. S., D. R. Easterling, and B. Gleason (2005), Maximum and minimum temperature trends for the globe: An update through 2004, *Geophysical Research Letters*, *32*(23), L23822, doi:10.1029/2005gl024379.

Vuille, M., E. Franquist, R. Garreaud, W. S. Lavado Casimiro, and B. Cáceres (2015), Impact of the global warming hiatus on Andean temperature, *Journal of Geophysical Research: Atmospheres*, *120*(9), 3745-3757, doi:10.1002/2015JD023126.

Wang, Q., X. Fan, and M. Wang (2016), Evidence of high-elevation amplification versus Arctic amplification, *Scientific Reports*, *6*, 19219, doi:10.1038/srep19219.

Wu, J., and X. J. Gao (2013), A gridded daily observation dataset over China region and comparison with the other datasets, *Chinese Journal of Geophysics*, *56*(4), 1102-1111.

Xu, Y., X. J. Gao, S. Y. Yan, C. H. Xu, Y. Shi, and F. Giorgi (2009), A daily temperature dataset over China and its application in validating a RCM simulation, *Advances in Atmospheric Sciences*, *26*(4), 763-772.

Xu, Y., V. Ramanathan, and W. M. Washington (2016), Observed high-altitude warming and snow cover retreat over Tibet and the Himalayas enhanced by black carbon aerosols, *Atmos. Chem. Phys.*, *16*(3), 1303-1315, doi:10.5194/acp-16-1303-2016.

Yan, L. B., and X. D. Liu (2014), Has Climatic Warming over the Tibetan Plateau Paused or Continued in Recent Years?, *Journal of Earth, Ocean and Atmospheric Sciences*, *1*(1), 13-28.

Yan, L. B., Z. Liu, G. Chen, J. E. Kutzbach, and X. Liu (2016), Mechanisms of elevation-dependent warming over the Tibetan plateau in quadrupled CO2 experiments, *Climatic Change*, *135*(3), 509-519, doi:10.1007/s10584-016-1599-z.

Yang, K., H. Wu, J. Qin, C. Lin, W. Tang, and Y. Chen (2014), Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review, *Global and Planetary Change*, *112*, 79-91.

Yang, K., B. S. Ye, D. G. Zhou, B. Y. Wu, T. Foken, J. Qin, and Z. Y. Zhou (2011), Response of hydrological cycle to recent climate changes in the Tibetan Plateau, *Climatic Change*, *109*(3-4), 517-534, doi:10.1007/s10584-011-0099-4.

Yang, T., X. Hao, Q. Shao, C.-Y. Xu, C. Zhao, X. Chen, and W. Wang (2012), Multi-model ensemble projections in temperature and precipitation extremes of the Tibetan

Plateau in the 21st century, *Global and Planetary Change*, *80-81*, 1-13, doi:10.1016/j.gloplacha.2011.08.006.

You, Q. L., K. Fraedrich, J. Min, S. Kang, X. Zhu, G. Ren, and X. Meng (2013), Can temperature extremes in China be calculated from reanalysis?, *Global and Planetary Change*, *111*, 268-279.

You, Q. L., S. C. Kang, E. Aguilar, N. Pepin, W. A. Flugel, and Y. P. Yan (2011), Changes in daily climate extremes in China and their connection to the large scale atmospheric circulation during 1961-2003, *Climate Dynamics*, *36*, 2399-2417.

You, Q. L., S. C. Kang, E. Aguilar, and Y. P. Yan (2008), Changes in daily climate extremes in the eastern and central Tibetan Plateau during 1961-2005, *Journal of Geophysical Research-Atmospheres*, *113*, D07101.

You, Q. L., J. Min, and S. Kang (2016), Rapid warming in the Tibetan Plateau from observations and CMIP5 models in recent decades, *International Journal of Climatology*, *36*(6), 2660-2670, doi:10.1002/joc.4520.

You, Q. L., J. Min, W. Zhang, N. Pepin, and S. Kang (2015), Comparison of multiple datasets with gridded precipitation observations over the Tibetan Plateau, *Climate Dynamics*, *45*(3), 791-806, doi:10.1007/s00382-014-2310-6.

You, Q. L., J. Z. Min, K. Fraedrich, W. Zhang, S. C. Kang, L. Zhang, and X. H. Meng (2014), Projected trends in mean,maximum, and minimum surface temperature in China from simulations, *Global and Planetary Change*, *112*, 53-63.

Zhai, P. M., and X. H. Pan (2003), Trends in temperature extremes during 1951-1999 in China, *Geophysical Research Letters*, *30*(17), 1913, doi:10.1029/2003gl018004.

Zhang, Q., V. P. Singh, J. F. Li, and X. H. Chen (2011), Analysis of the periods of maximum consecutive wet days in China, *Journal of Geophysical Research: Atmospheres*, *116*, D23106.

Table 1. CMIP5 models used in this study.

No.	Model	Institution	Country	Resolution (Lon×Lat Levels)
1	ACCESS1.0	Commonwealth Scientic and Industrial Research Organisation and Bureau of Meteorology, Australia	Australia	192×145L38
2	BNU-ESM	Beijing Normal University, China	China	128×64L26(T42)
3	CanESM2	Canadian Centre for Climate Modelling and Analysis, Canada	Canada	128×64L35(T63)
4	CCSM4	National Center for Atmospheric	USA	288×192L26

		Research (NCAR), USA		
5	CESM1-BGC	National Science Foundation/Department of Energy NCAR, USA	USA	288×192L26
6	CMCC-CM	Centro Euro-Mediterraneo per I Cambia-menti, Italy	Italy	480×240L31 (T159)
7	CNRM-CM5	Centre National de Recherches Meteorologiques, Meteo-France, France	France	256×128L31 (T127)
8	CSIRO-Mk3.6.0	Commonwealth Scientic and Industrial Research Organization (CSIRO), Australia	Australia	192×96L18 (T63)
9	EC-EARTH	Royal Netherlands Meteorological Institute, Netherlands	Netherlands	320×160L62 (T159)
10	FGOALS-s2	Instute of Atmospheric Physics, Chinese Academy of Sciences, China	China	128×108L26
11	GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, USA	USA	144×90L48
12	GISS-E2-R	Goddard Institute for Space Studies, USA	USA	144×90L40
13	HadGEM2-ES	Met Office Hadley Centre, UK	UK	192×145L40
14	IPSL-CM5A-MR	Institut Pierre-Simon Laplace, France	France	144×143L39
15	MIROC5	AORI, NIES, JAMSTEC, Japan	Japan	256×128L40 (T85)
16	MPI-ESM-MR	Max Planck Institute for Meteorology, Germany	Germany	192×96L95 (T63)
17	MRI-CGCM3	Meteorological Research Institute, Japan	Japan	320×160L48 (T159)

Table 2. Definitions of temperature extreme indices calculated by RClimDEX.

Index	Descriptive Name	Definition	Units
TX10p	Cold day	Count of days when TX < 10^{th} percentile of 1961-1990	days
TN10p	Cold night	Count of days when TN < 10^{th} percentile of 1961-1990	days
TX90p	Warm day	Count of days when TX > 90^{th} percentile of 1961-1990	days
TN90p	Warm night	Count of days when TN > 90^{th} percentile of 1961-1990	days
DTR	Diurnal temperature	Annual mean difference between TX and TN	℃

	range		
TXn	Coldest day	Annual lowest TX	℃
TNn	Coldest night	Annual lowest TN	℃
TXx	Warmest day	Annual highest TX	℃
TNx	Warmest night	Annual highest TN	℃
GSL	Growing season length	Annual count of days between the first span of at least 6 days with TG > 5℃ after winter and first span after the summer of 6 days with TG< 5℃	days
FD0	Frost days	Annual count of days when TN< 0℃	days
ID0	Ice days	Annual count of days when TX <0℃	days
SU25	Summer days	Annual count when TX >25℃	days
TR20	Tropical nights	Annual count when TN >20℃	days
WSDI	Continued warm period	Count of continued days when TX > 90th percentile of 1961-1990	days
CSDI	Continued cold period	Count of continued days when TN < 10th percentile of 1961-1990	days

Note: TX is the daily maximum temperature; TN is the daily minimum temperature; TG is daily mean temperature; TN_{mean}/TX_{mean} is the mean of daily minimum/maximum temperatures for the period 1961-1990, respectively.

Table 3. Trends and temporal skill scores for each temperature extreme index from observations (OBS), the optimal models ensemble (OME), and the all-models ensemble (AME), respectively.

Indices	Trends				Temporal skill score	
	OBS	OME	AME	Unit	OME	AME
TXx	0.21	0.02	-0.02	℃/decade	5.11	6.82
TNn	0.53	0.01	-0.02	℃/decade	4.99	3.65
TXn	-0.07	0.01	-0.07	℃/decade	3.57	2.74
TNx	0.53	0.05	-0.01	℃/decade	7.28	8.72

DTR	-0.22	0.01	0.05	℃/decade	0.31	0.48
TX90p	1.43	0.84	0.95	day/decade	1.22	2.11
TX10p	-0.87	-0.76	-0.71	day/decade	1.18	3.00
TN90p	2.60	1.54	1.35	day/decade	1.30	2.39
TN10p	-2.29	-1.18	-1.07	day/decade	9.00	9.00
SU25	0.94	0.01	-0.69	day/decade	37.98	44.57
FD0	-4.00	0.23	0.75	day/decade	9.34	9.66
TR20	0.41	0.04	-0.14	day/decade	78.34	111.55
ID0	-3.08	0.33	0.30	day/decade	2.50	1.71
GSL	3.64	-0.67	-1.00	day/decade	9.18	9.37
WSDI	2.16	1.66	1.66	day/decade	0.57	0.73
CSDI	-0.99	-0.69	-0.61	day/decade	0.23	1.18

Figure

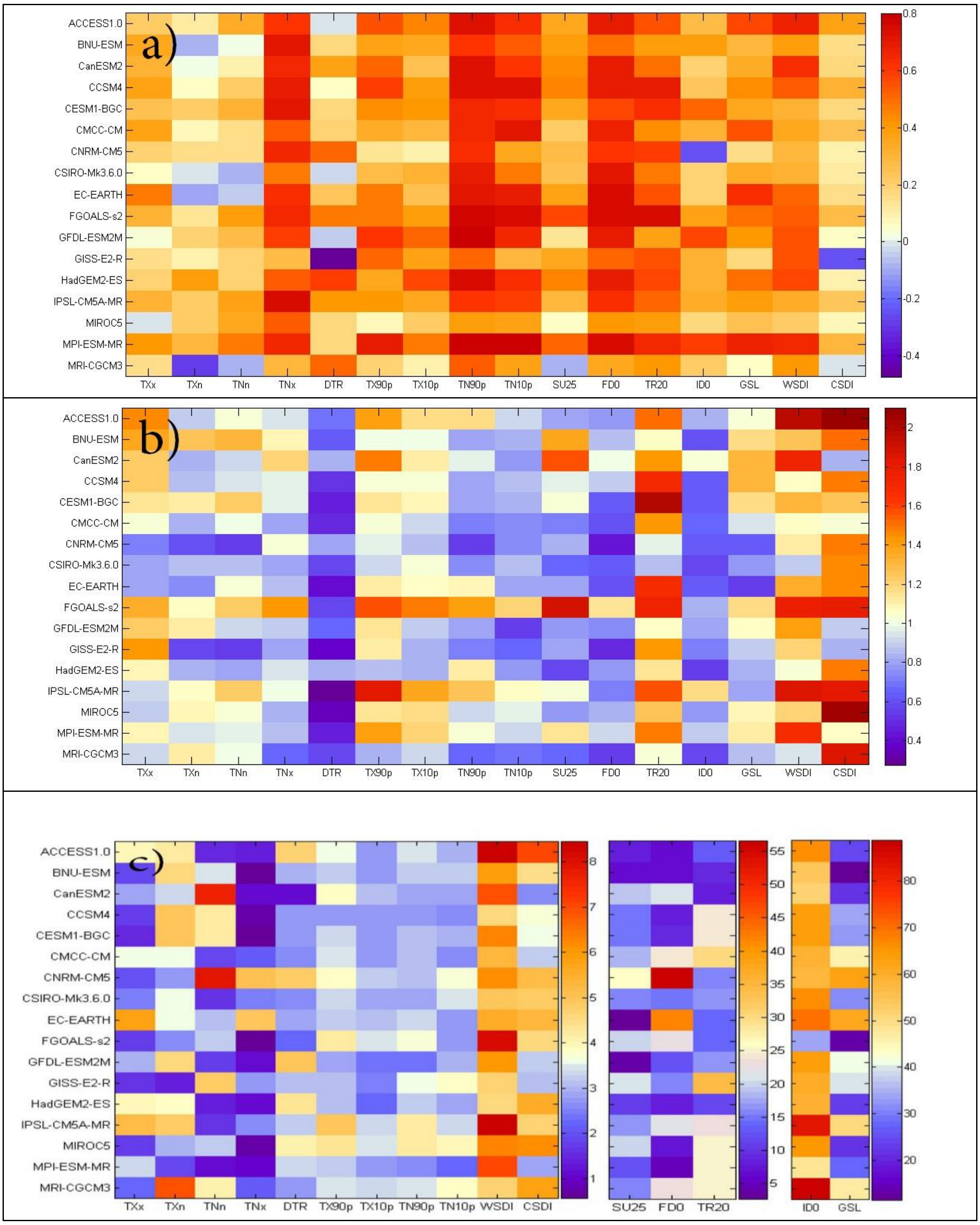


Figure 1. Portrait diagram for temporal correlation coefficient (a, top panel), standard deviation ratio (b, middle panel) and root-mean-square error (c, bottom panel) of temperature extreme indices in the Tibetan Plateau between observations and CMIP5 models.

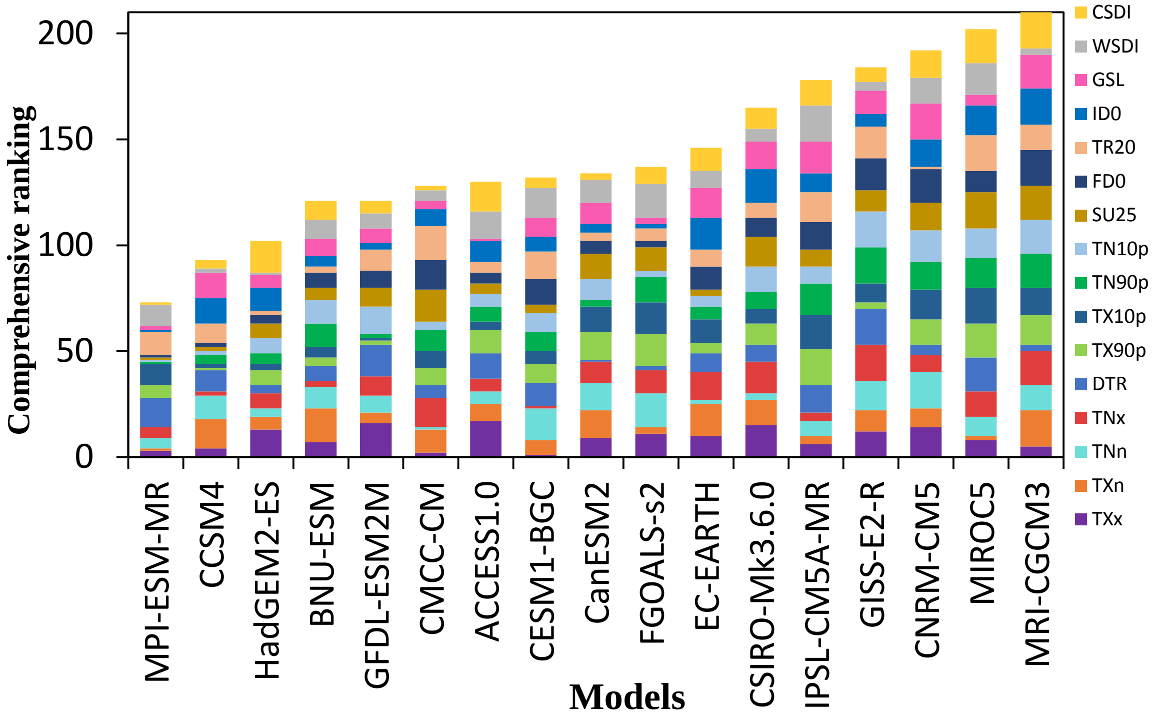
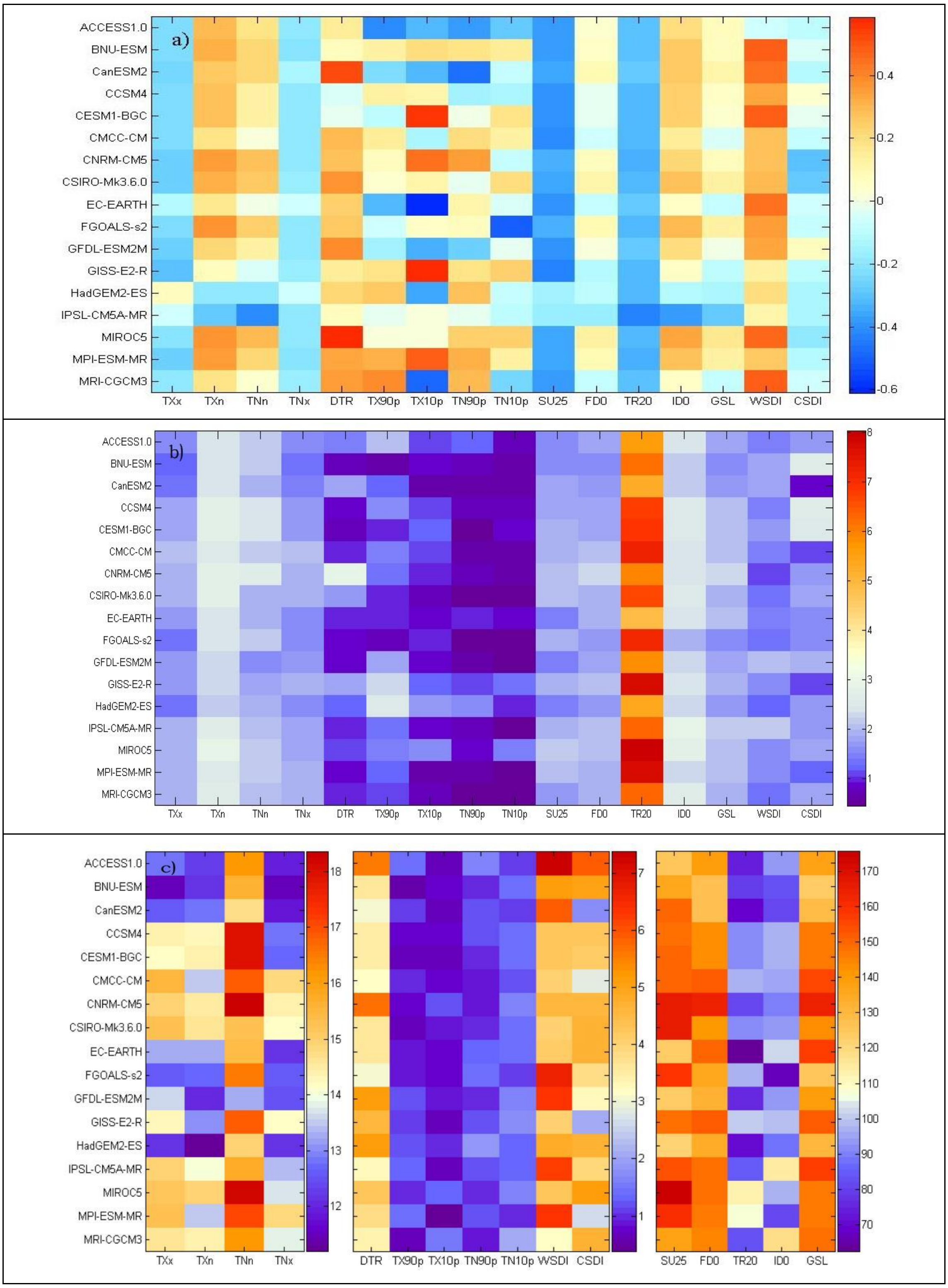


Figure 2. Comprehensive model ranking based on temporal correlation coefficient, standard deviation ratio and root-mean-square error for each temperature extreme index. The y axis is the sum of model ranking of all temperature extreme indices.



654 **Figure 3.** Same as Figure 1 but for spatial patterns.

655

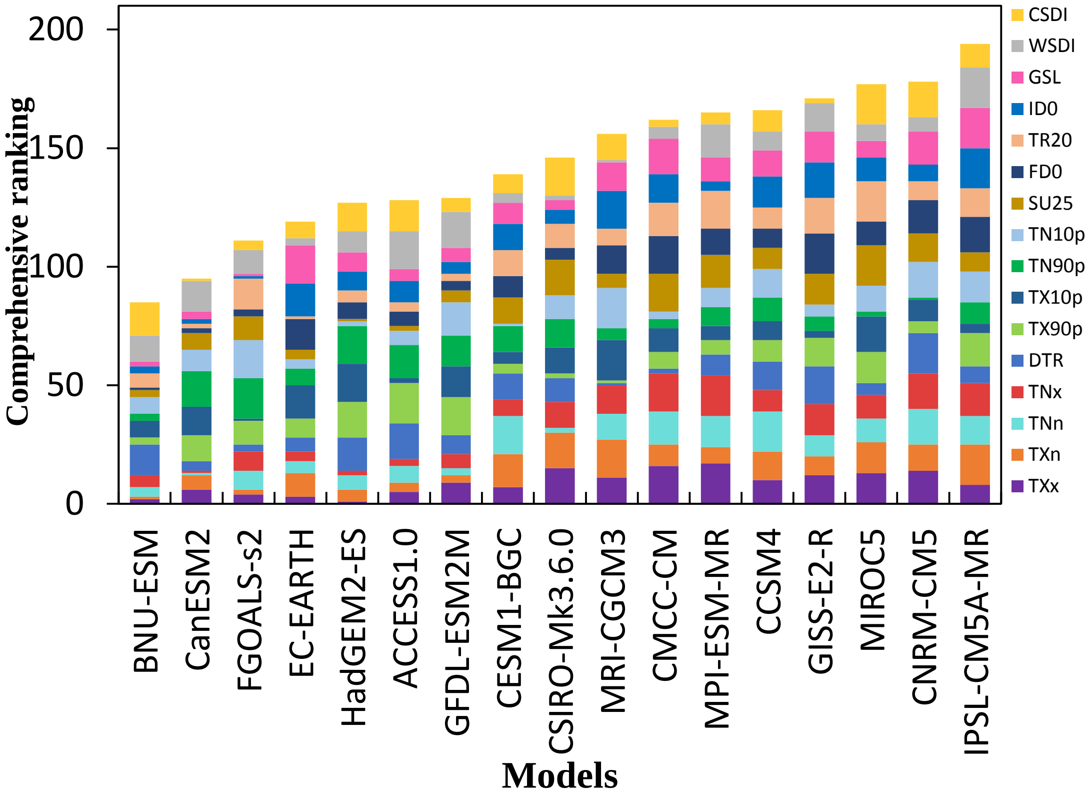


Figure 4. Same as Figure 2 but for spatial patterns.

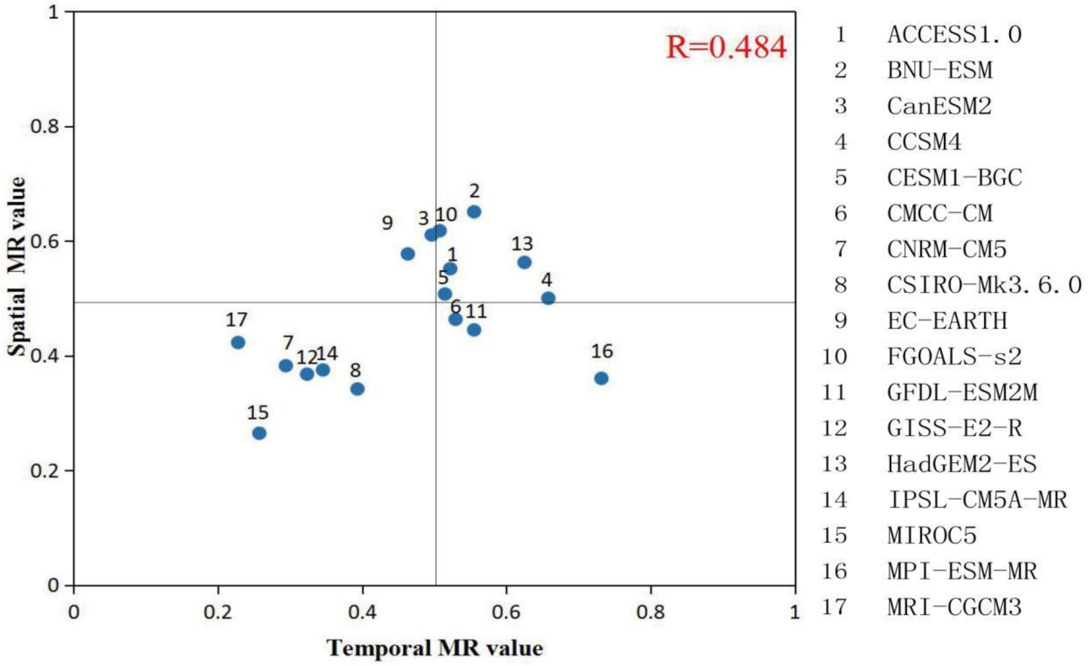


Figure 5. Scatter diagram showing the relationship between temporal and spatial model rank (M_R) value. Each dot represents a model, identified by its number on the right. The correlation coefficient between temporal and spatial M_R value is 0.448.

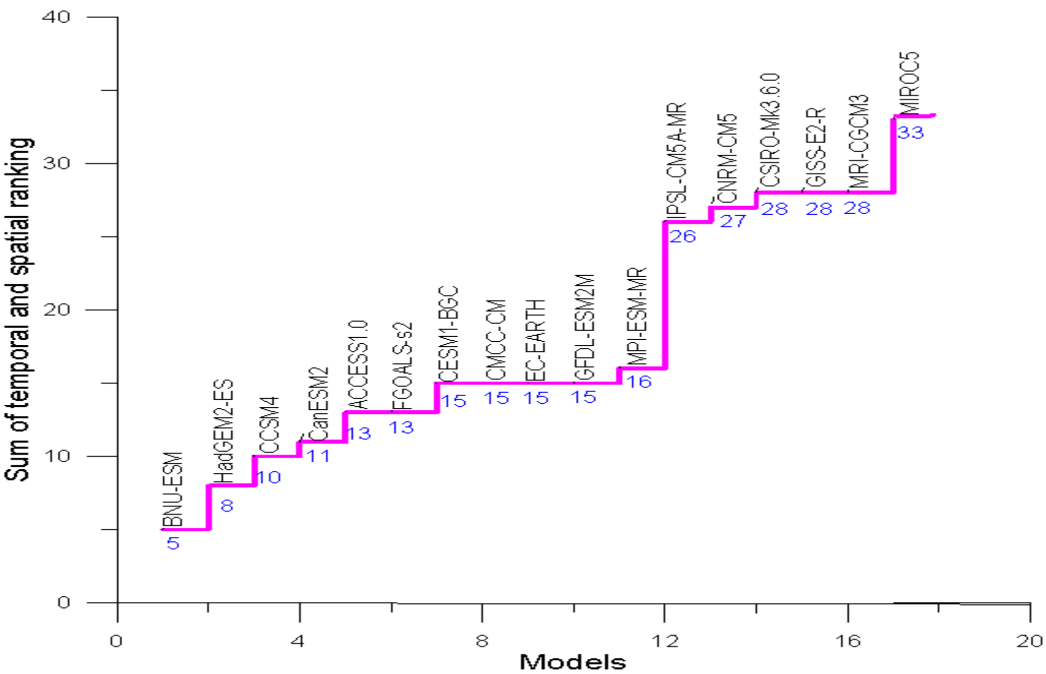


Figure 6. Comprehensive model ranking based on temporal and spatial correlation coefficient, standard deviation ratio and root-mean-square error of temperature extreme indices in the Tibetan Plateau. x axis is the number of the model, the number below each model and y axis is the sum of model ranking of all temperature extreme indices.

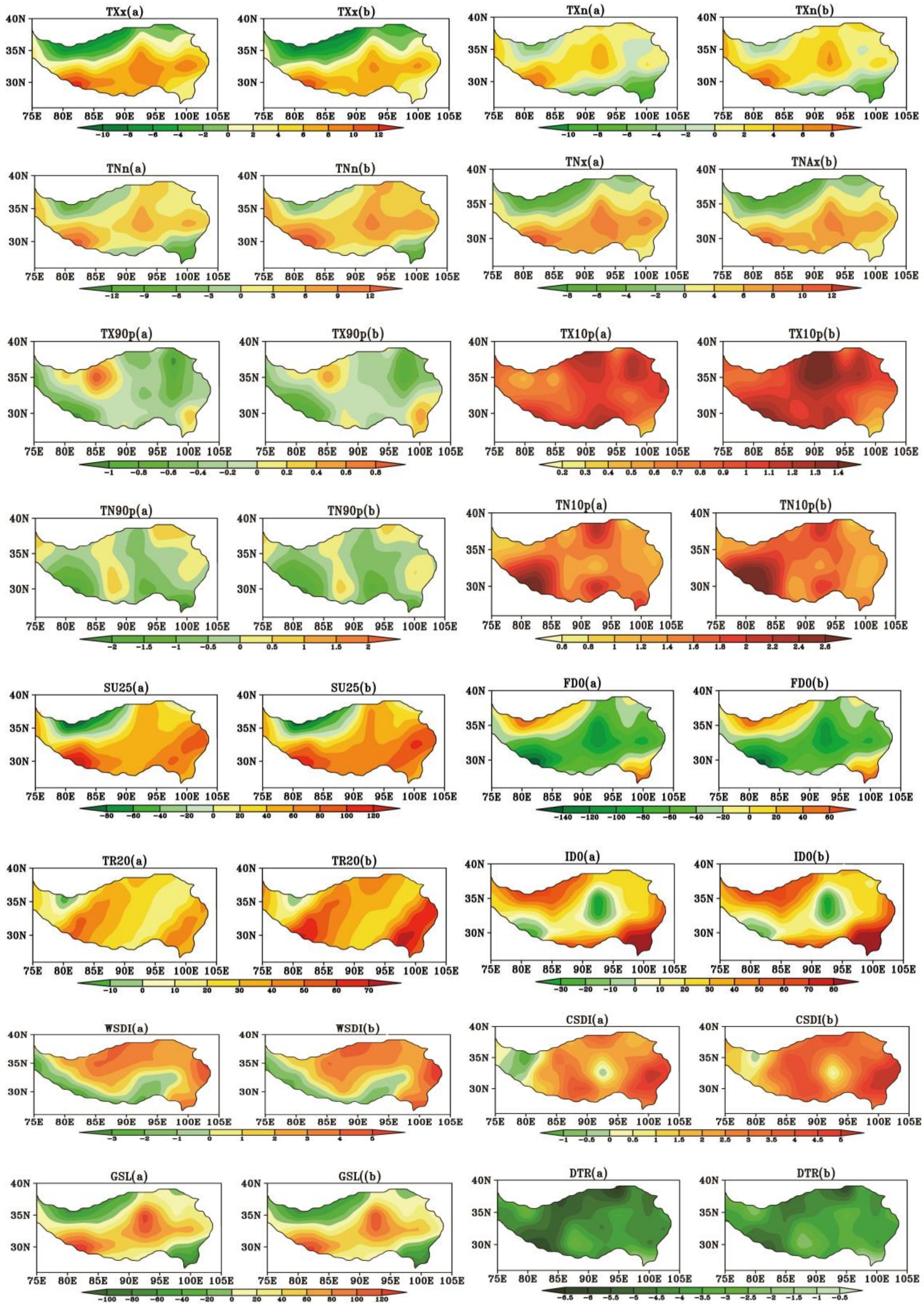


Figure 7. The climatological differences of temperature extreme indices between the optimal models ensemble (a in each panel)/all models ensemble (b in each panel) and observations in the Tibetan Plateau.

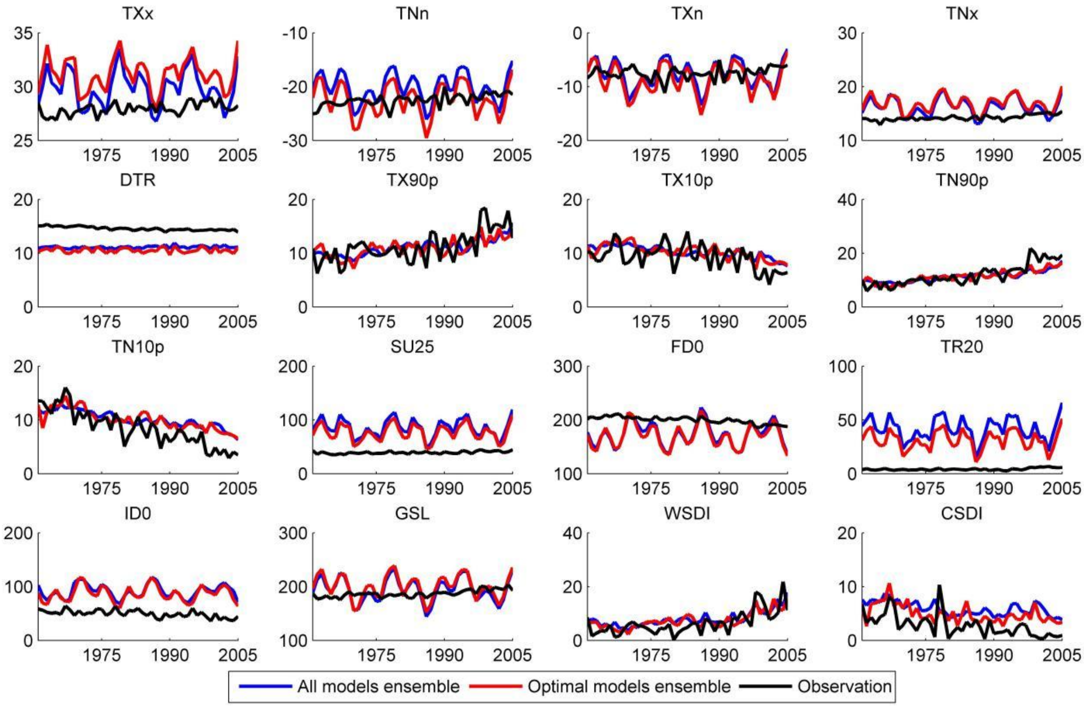


Figure 8. Time series of temperature extreme indices from the optimal/all models ensemble and observations in the Tibetan Plateau during 1961-2005.

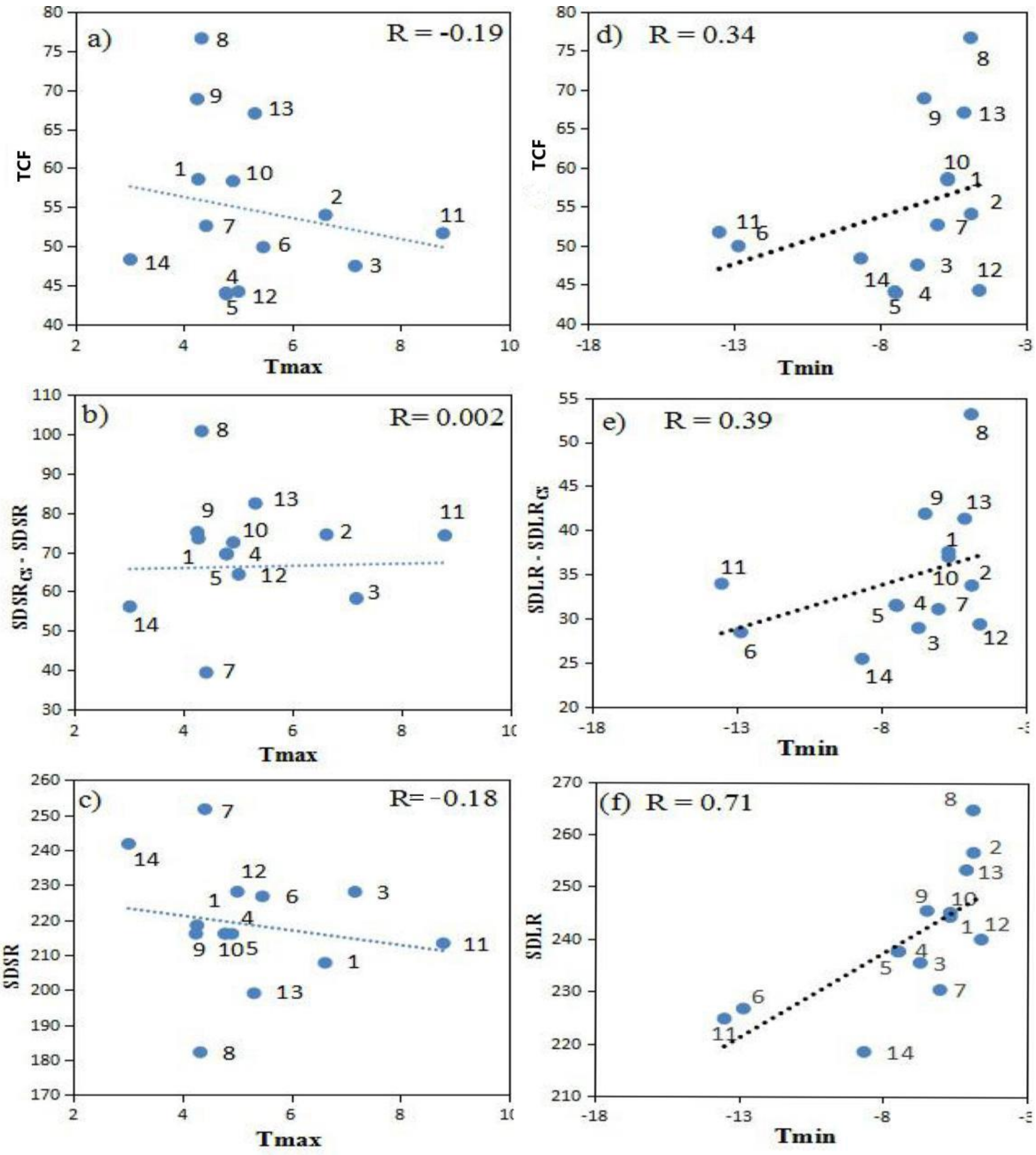


Figure 9. Relationship between the mean maximum temperature (Tmax), minimum temperature (Tmin), and climate variables from each CMIP5 model during 1961-2005 in the Tibetan Plateau on the annual basis. Climate variables are the surface downwelling shortwave radiation (SDSR), the SDSR at clear sky (SDSRcs), the surface downwelling longwave radiation (SDLR), the SDLR at clear sky (SDLRcs) and the total cloud fraction (TCF), respectively.

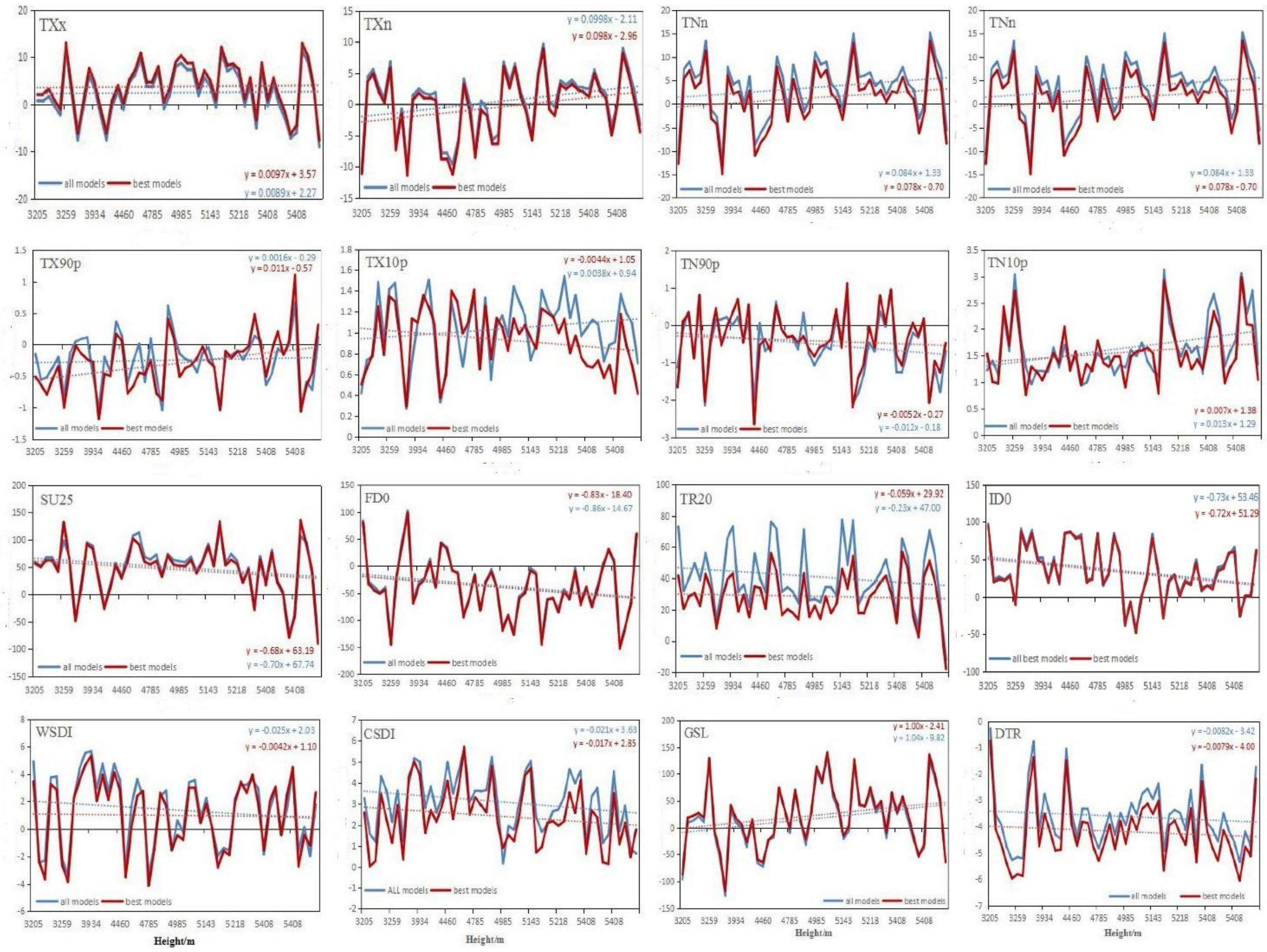


Figure 10. Relationship between elevation and bias for each temperature extreme index (optimal/all models ensemble minus observations) in the Tibetan Plateau during 1961-2005.