Comparisons of time series of annual mean surface air temperature for China since the 1900s: Observations, model simulations and extended reanalysis

Qingxiang Li[1], Lei Zhang[1], Wenhui Xu[1], Tianjun Zhou[2], Jinfeng Wang[3], Panmao Zhai[4], Phil Jones[5]

1 China Meteorological Administration, National Meteorological Information Center, 46 Zhongguancun South Avenue, Haidian District, Beijing, China 100081

2 LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, P.O. Box 9804, Beijing, China,100029

3 LREIS, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China 100101

4 China Academy of Meteorological Sciences, 46 Zhongguancun South Avenue, Haidian District, Beijing, China, 100081

5 Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK; Also Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Revised for Submission in *Bulletin of American Meteorological Society*

April 2016

***Corresponding author:**

Dr & Prof Qingxiang Li
National Meteorological Information Center
China Meteorological Administration
46 Zhongguancun South Avenne, Haidian District,
Beijing, China 100081
Tel: 86-10-58995059
E-Mail: liqx@cma.gov.cn

Capsule

This paper assesses the similarities and differences of several annual average SAT time series for China based on historical meteorological observations since the 1900s.

Abstract

Time series of global or regional average surface air temperature (SAT) are fundamental to climate change studies. A number of studies have developed several national and regional SAT series for China, but due to the diversity of meteorological observational sites, different quality control routines for data and the inconsistency of statistical methods used, they differ in long-term trends. This paper assesses the similarities and differences of the existing time series of the annual average SAT for China that are based upon historical meteorological observations since the 1900s. The results indicate that the China average is similar to the series for the Northern Hemisphere (NH) landmass, except that the initial warming of the NH series derived from the CRUTEM3/4 datasets ends earlier (before the early 1940s) than in China's series. A major difference among the existing China average time series is the 1940s warmth, a period when there were very few observations across the country due to World War II. The SAT anomalies for China during the 1930s-1940s have been reduced by improved homogeneity assessment compared to previous estimates. The new improved time series is in better agreement with both the historical 20th century reanalysis data and the historical climate simulation of CMIP5 models. The new time series also shows the slowdown of the warming trend during the past 18 years (1998-2015). The best estimate of a linear trend for increases in temperature with a 95% uncertainty range is 0.121±0.009 °C per decade for 1900-2015, indicating that the improved homogeneity assessment for China leads to a slightly greater trend than that based on raw data (0.107±0.009 °C per decade).

Keywords: Climate change, China SAT series, Homogeneity, Observations, Reanalysis data

1. Introduction

Over the past century, global-scale climate warming has been accelerating, and climate change is drawing increasing attention from the media and the public. Surface Air Temperature (SAT) is the major subject of concern in the context of climate change. Starting in the second half of the last century (after 1950), the warming of the climate has been obvious at the global scale (Brohan et al., 2006; Hansen et al., 2010; Lawrimore et al., 2011; Jones et al., 2012; Morice et al., 2012; Vose et al., 2012a; Hartmann et al, 2013). In the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) headline statements from the Summary for Policymakers, it was noted that many changes in the climate system since 1950 are unprecedented when compared to earlier decades. From 1880 to 2012, the global average sea and land surface temperature shows an increasing linear trend of 0.85°C. From 2003 to 2012, the average temperature increased by 0.78°C (0.72 ~ 0.85) compared with that from 1850 to 1900 (Hartmann et al, 2013). Further, IPCC AR5 notes that the warming rate has slowed down recently (1998-2012). This short 15-year period has been discussed by a number of authors (Easterbrook，2008；Easterling and Wehner，2009；Kaufmann et al., 2011；Fyfe et al., 2013；Guemas et al., 2013；Slingo et al., 2013；Karl et al., 2015；Lewandowsky et al., 2015).

Brohan et al (2006) indicated that temperature series have three types of uncertainty: station errors, sampling errors and bias errors. For China, the observational (station) errors in the SAT series are relatively small and well understood. Errors caused by the homogeneity adjustment of temperature series have been evaluated and discussed (Jones et al., 2008; Li et al., 2010a). Sampling errors in China are an issue when the number of observational sites is reduced (Li et al., 2010a; Wang et al., 2014); this is especially relevant before the 1950s and in the 1940s. Bias errors, including the impacts of urbanization (Zhou et al., 2004) and/or land use changes (Zhang et al., 2005) on SAT, are particularly important issues. Many efforts have been devoted to improved understanding of these sources of error in China, focusing on either individual regions, eastern China, or China as a whole. The results indicate that urbanization effects are weakest in western China and in the smaller cities. The urbanization effects are evident in some large cities, but the size of the effect across the whole country is relatively small (an order of magnitude smaller than the longer term warming since 1900) (Jones et al., 1990; Li et al., 2004a, 2010b, 2014; Zhou et al., 2004; Zhang et al., 2005; Hua et al., 2008; Jones et al, 2008; Ren et al., 2008; Yan et al., 2009; Ren and Ren, 2011; Yang et al., 2012; Wang et al., 2015).

The development of SAT time series averaged over mainland China has been a focus in the climate change research community in China (Zhang and Li, 1982; Tu, 1984, Ding and Dai, 1994; Wang et al., 1998; Wang and Gong, 2000; Chen et al., 2004, Zhai et al, 2004; Qin et al., 2005; Ding and Ren, 2008; Wang et al., 2014). Currently, four sets of long-term time series exist that have been widely used in climate change studies in China: WangS, Tang, Li, and WangJ times series (see Table 1 for details). In this study, we supplement these four existing time series with two time series derived from CRUTEM3 (Brohan et al., 2006) and CRUTEM4 datasets (Jones et al., 2012). Analysis of these six time series leads to divergent conclusions about SAT variations due to their differences in station densities and data processing techniques.

Global climate models have shown reasonable performance on the global scale, but large uncertainties in regional scale performance still exist. The performances of the CMIP3 (the 3rd Coupled Model Intercomparison Project. Meehl et al, 2007)) and CMIP5 (the 5th Coupled Model Intercomparison Project. Taylor et al., 2012) modelsin the simulation of China SAT changes have been evaluated in many studies (for details of these simulations see the SM) (Zhou and Yu, 2006; Xu and Xu, 2012; Guo et al., 2013; Zhang et al., 2013; Zhou et al, 2013; Jin and Zhou, 2014). However, the observed time series used in these comparisons are different. The climate modeling community needs to pay more attention to the uncertainty of observed time series. In addition, the first reanalysis product at centennial scales (20CR), developed by the US National Oceanographic and Atmospheric Administration (NOAA), provides new opportunity for data-model comparisons. The near-surface air temperature in 20CR provide useful results (especially in terms of the continuity of series) (Cheng et al., 2013) that can be verified by the simulation of global SAT variation (Compo et al., 2011, 2013). In this study, the SAT time series derived from 20CR data is also compared to the observational data.

The objective of this study is to assess the similarities and differences of these annual mean SAT time series for China based on historical meteorological observations since 1900. The existing observed time series are compared to the results of CMIP5 model simulations and the 20CR reanalysis. The overall temperature trend of China SAT is investigated, and the characteristics of each type of data are discussed.

TABLE 1

2. Data and methods

High quality climate datasets are the basis for the detection and attribution of climate changes (Karl and Williams, 1987; Gullet et al., 1990; Peterson et al., 1998). For years, temperature analyses across China, especially on local or sub-regional scales, saw discrepancies due to the inhomogeneity of the baseline data (Li et al., 2004b). To solve this problem, considerable efforts have been devoted to the homogenization of the SAT station data for China (Li et al., 2004, 2009; Song et al., 2004; Li and Yan, 2009; Xu et al., 2013; and SM). The first homogenized SAT datasets for the whole country, starting at the beginning of 20th century, were published in 2010 (Li et al, 2010a). Currently, there are four sets of SAT time series that have been widely used in climate change studies in China (see Table 1 for details); these four sets of data are used in our analysis. The WangS series (Wang et al, 1998) and Tang series (Tang and Ren, 2005) are described in Cao et al. (2013); the Li series (Li et al, 2010a) and WangJ series (Wang et al, 2014) are both developed from Li et al (2010a).

In addition to the aforementioned China average time series, the CRUTEM3/4 datasets (Brohan et al., 2006; Jones et al., 2012) were also used to estimate the Northern Hemisphere/China surface temperature changes. Another homogenized global LSAT dataset (GHCN-M, Lawrimore et al., 2011), developed by NOAA’s National Center for Environment Information (NCEI), was also used for comparison with the China observational dataset. The Climate Anomaly Method (CAM, Jones and Hulme, 1996) was applied to gridded deviations from the climatological average (“anomalies”) to construct a regional average time series over all of China.

The 20th century historical climate simulations of CMIP5 models (Taylor et al., 2012) were used in our analysis. The results of 41 CMIP5 models were compared to the historical observational data. Because the historical climate simulation of CMIP5 models ends in 2005, our model-data comparison focuses on the period from 1900 to 2005. The Taylor diagram (Taylor, 2001) was used in model evaluations. First, data from the 41 models were re-gridded to 1×1 resolution grids, and the China average SAT series for each model was calculated by averaging all the anomalies series for each grid by latitude cosine weighting. This was nearly the same methodology applied by Jones and Hulme (1996) with the exception of a different resolution between the model (1×1) and observation (5×5) datasets. The China SAT series was obtained by averaging all the selected models series (41 or 8 in this paper) using the same weighting.

Data from the 20th Century Reanalysis V2 were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (accessible from their web site at http://www.esrl.noaa.gov/psd/). A China average was calculated for the years 1900 to 2005 for additional comparisons with the CMIP5 simulations (for details of these simulations see the SM).

3. Results

3.1 Comparison of observed time series

Fig. 1

A comparison of the six time series of average SAT in China (see Table 1 for details; all series plotted relative to 1961-1990 average) for the period of 1909 to 2006 (when all data are available) is shown in Figure 1. Note that all series, with the exception of the WangS series, have a large amount of missing data from western parts of China. For example, in the Li series (which has the densest observational station data coverage among the series with the exception of WangS) the length of grid-point time series to the west of 110°E is 46-85 years. The station series for western China began between 1922 and 1961. The WangS series augmented western China station data with proxy climate data. First, they derived regional annual SAT series (starting in 1880) for Northeast China, North China, East China, South China, Taiwan, Central China, Southwest China, Northwest China, Xinjiang and Tibet based on observations. In addition, they included Dunde and Guliya ice core data, historical documentary information and tree ring data. WangS then obtained China SAT time series by area-weighting the averages of all regions (Wang et al, 1998). The WangJ series differs from the others because a grid was not developed; instead, the China average was calculated as one domain.

All six SAT series for China reveal similar decadal variations since the 1900s. The period from 1909-2006 can be divided into the following epochs: 1909 to the mid-1940s (the first stage of SAT rise), the mid-1940s to the late-1960s (the stage of SAT decline), and the late-1960s to 2006 (the second stage of accelerating SAT rise). The differences among the series are mainly evident in the period before the 1950s. The Tang anomaly series is slightly lower in the first 10 years, while the WangS anomaly series is higher than the other series from 1909 to 1929. From the 1930s to the 1950s, the WangS and Tang anomaly series are slightly higher than those of Li and WangJ. After the 1950s, the six series are much more aligned with each other.

The reason for the spread among the existing time series lies in the methods and the basic data used; homogenization of the basic climate data is the key factor here. All of the homogenized series (Li, WangJ and CRUTEM3/CRUTEM4) show less variance among one other. CRUTEM3 always lies in the middle of the various series, Li and CRUTEM4 are consistently slightly below the other series between 1916 and 1951, and WangJ is slightly below all the other series around 1936 & 1968. The basic data used for establishing the Li and WangJ series are identical, but the two series differ slightly from each other due to the utilization of different statistical methods (even after the 1950s). Tang uses non-homogenized station data, but is closer to the other four series than WangS, which uses non-instrumental proxy data for the annual averages before the 1950s.

Four series are compared in Figure 2 (CRUTEM3, CRUTEM4, WangJ and Li). For additional comparison with larger-scale SAT variations, we also include the series for the Northern Hemisphere (NH_CRU), calculated using CRUTEM (here, we include both NH SAT series calculated from CRUTEM3 and CRUTEM4, which are extremely similar). The NH_CRU series also has high consistency with China's series, which is similar to the finding of Bradley et al (1987). This is likely related to the fact that China is a large mid-latitude country and is a significant part of the NH land mass (see discussions in Jones et al., 2007 and Jones et al., 2012). The China average series reveal a more pronounced warming period starting in the 1940s, with 1946 as the warmest year before 1980; however, for NH_CRU, the warmest year before 1980 is 1938. Thus, the first 20th century warming phase for NH_CRU ends earlier than in China's series, and the subsequent NH temperature decline also ends earlier, which is consistent with Bradley et al (1987).

FIG. 2

3.2 Comparative analysis of China SAT time series simulated by climate models and reanalysis

The annual mean SAT series simulated by 41 models for 1900-2005 (using 1961-1990 as the base period) is shown in Figure 3. As evident from the ensemble mean of the models (blue curve), the simulated SAT anomalies from the early 1920s to the 1960s are generally higher than those of the observed series (red curve), with an average differences close to 0.3°C before the 1940s. After the 1950s, the simulated rate of climate warming is still lower than the observations (Figure 4b). A comparison between the long-term temperature change estimated by multiple-member ensemble historical runs from 41 models and seven observational series (the eastern China series by Cao et al. (2013) is also included here) is shown in Figure 4a. The rate of change of mean SAT for China simulated by the 41 models for 1900-2005 is approximately 0.00-0.20 °C per decade, with an average of 0.06°C per decade, and the average rate simulated for 1951-2005 is 0.15°C per decade (less than the 0.19°C per decade for observation). The seven observational series exhibit different trends centered at approximately 0.09 °C per decade (red line in Figure 4a). The trends for the model series are slightly smaller than those for observational series at the centennial-scale, but the agreement is better for the last 55 years (1951-2005) (Figure 4b).

Two additional approaches are adopted to additionally evaluate the simulation of the SAT anomalies over China. First, Pearson correlation coefficients (CCs) between the observational and simulated series for the full 106 (1900-2005) years are calculated. The highest CC is 0.75, found for FGOALS-s2 by the Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences (CAS). In decreasing order of CC (larger than 0.6), the remaining 7 top-ranking series are: MPI-ESM-MR (0.69), BCC-CSM1.1 (0.67), BNU-ESM (0.67), IPSL-CM5A-LR (0.65), BCC-CSM1.1 (0.62), CESM1 (WACCM) (0.62) and CESM1 (FASTCHEM) (0.61). It is interesting to note that four out of the eight top-ranking simulated series for China are derived from models developed by Chinese institutes. Another important observation is that using the homogenized series from Li provides higher correlations (the CCs shown here are higher than those given by Zhang et al (2013), where the unadjusted (raw) observation data were used).

Figure 5 (Taylor diagram) akin to Jiang and Tian (2013), which shows the performance of the models in their historic simulations of SAT anomalies from 1900-2005. The closer the distance between model series and REF (observations here are referring to the Li series), the better the performance of the corresponding model. The top-ranking models by distance are: BCC-CSM1.1, GFDL-ESM2M, CMCC-CM, CanESM2, CESM1 (CAM5), NorESM1-M, MIROC-ESM, and MPI-ESM-MR. The average of the eight top-ranking models (whose normalized standard deviation is below 2.0) reproduces the SAT observational variations for the last 106 years with a correlation coefficient of 0.81 (Figure 6a). Compared with the observations (Li), the model series average underestimates the observations before the 1970s except for the short period in the 1950s. As observed in Figure 6b, the average of the eight top-ranking models reproduces the SAT observational variations for the last 105 years with a correlation coefficient of 0.79, which is slightly lower than in Figure 6a. However, as a whole, the average models series agrees well with observations during the entire period.

FIG. 3

FIG. 4

FIG. 5

FIG. 6

Figure 3 and Figure 6 also display the regional SAT time series of China calculated from 20CR (black bold lines). During the entire twentieth century, the long-term SAT warming trend for China is very well reflected in 20CR. For 20CR, the early temperature rise continues from the 1900s to the 1960s, followed by a sudden decline at the beginning of the 1960s and then a rise again after the mid-1960s. The average rate of rise of China's SAT for 1900-2005 calculated by 20CR data is 0.10°C/10 a, which is close to the level of the observed time series (0.09°C/10 a). In 20CR, temperatures begin to decrease from the mid-1990s. It is important to note that the analytical model of the 20CR dataset is driven by monthly average sea surface temperatures, and only surface pressure data are assimilated. Therefore, the SAT estimates obtained are completely independent from the myriad of land surface temperature observations. Without assimilating surface air temperature and/or time-varying anthropogenic aerosol data (though volcanic aerosols were included), the 20CR succeeded in reproducing the recent winter SAT cooling features in China but failed in reproducing the significant summer warming trends in southern China from 1998-2012 (Li et al., 2015; and Fig. S3 in SM).

This has led to the underestimation of China's average SAT in the past 15 years.

4. Discussion

4.1 SAT anomalies in the 1940s

1) Observations

There is little doubt about the high SAT anomaly during the 1940s over China because all six observational series collected for this study consistently show similar features (Figure 1). However, the magnitudes of the SAT anomalies differ among these datasets (Li et al., 2010a). In the WangS series, the 1910s to the 1950s are warmer than the other series, while the Tang series only shows the high temperature anomalies in the 1940s. The two series achieve comparable amplitude in 1946 (WangS and Tang are 0.68°C and 0.88°C, respectively) with some years in the most recent warm period (2000s), while the remaining series have relatively lower amplitude (0.31°C (Li), 0.38°C (WangJ) and 0.36°C (CH_CRU) in 1946). Note that the only difference between Li (and WangJ) and the Tang series is that the former two series used a homogenized dataset, while the latter did not, which indicates that data homogenization is likely to be the most significant cause of differences in the anomalies during the 1940s.

The uncertainties for SAT change series (Li) has been assessed in Li et al., (2010), Based on their evaluation, the average 95% annual uncertainty range is about 0.327 °C during 1940s (1941-1950), so the decadal uncertainty would be $0.327 / \sqrt{10} \approx 0.1\,°C$, which is smaller than 0.15°C in 1930s, and about 0.20~0.22°C in 1900s, 1910s and 1920s (Figure S2 in SM). Most of the uncertainties comes from the limited data coverage (Li et al., 2010, see their Figure 5), and it agrees well with the station numbers change during 1900s-1940s for eastern and western China (E China and W China), respectively (Figure S1 in SM). The W China has similar inter-annual and decadal changes with E China from 1930s to 2010s. This assures that the Li series can present the SAT change for whole China to a certain extent even if the series was built based on only the stations from eastern China in the earlier years (1900s-1920s).

FIG. 7

Because the observational data in the 1940s were disrupted due to World War II from 1937-1945, certain difficulties are associated with the detection of SAT variation in this time period. Li series basically reflects the average for all the 30 stations' SAT anomalies change since 1930s in the western China (Figure 7). Wang et al (1998) introduced proxy data (tree rings, ice cores and historical documents) to enhance estimates for the western half of China. But how to reliably estimate the missing instrumental observation with proxy data remains an open question. Annual averages for NH mid-latitude continental areas are principally dominated by temperatures in winter, but proxy sources are mostly indicative of the summer half of the year. And the WangS series is composed of annual averages and does not include separate series for the seasons (Wang and Gong, 2000). Although the use of proxy data has the problems above, but to increase the coverage of the data in the west, it is also a valuable endeavor, which is worth further study with better proxy data and more perfect treatment method.

2) CMIP5 and 20CR

Zhou and Yu (2006) noted that CMIP3 models do not achieve a good simulation of abnormal climate warming in China in the 1940s, which may be related to the absence of solar forcing variations that should change over time between different simulation experiments for the 20th century. For historical climate simulations using the CMIP5 models, most models still failed to reproduce the warmth of the 1940s. As evident from the comparison between the average of the simulations using all models and the observed time series in Figure 3, the warmth in the 1940s is less evident in the simulations. Using the observational average SAT time series (Li series) that has been established from the homogenized data in that paper, the amplitude of the temperature rise before the 1940s is much lower and the extent of abnormal temperature rise in the 1940s is also less significant. This result is also consistent with the 20CR data (Figure 3). 20CR is somewhat like AMIP-type experiments, but the comparison of this aspect has been addressed in Compo et al (2013). They showed that assimilating the MSLP data improved over the AMIP-type (which was just SST/Sea-Ice assimilation). It seems that the abnormal climate warming revealed by the SAT time series in the 1940s may not be as high as the WangS and Tang series indicate.

4.2 Temperature change trend estimates from 1900-2015

In this section, we consider the series that includes the most recent year (2015), so only some of the series (CRUTEM4, GHCN-M (Lawrimore et al., 2011), and Li + Xu (Li et al, 2010; the data since 2007 are derived from Xu et al, 2013)) can be used. The temperature anomaly series at the annual timescale are shown in Fig. 8. The linear change trends and 95% uncertainty ranges (Wesley, 1997) are calculated in Table 2.

Over the period from 1900 to 2015, the annual linear trend coefficient is 0.121±0.009 °C per decade, 0.130±0.009 °C per decade, and 0.114±0.009 °C per decade for Li+Xu, CRUTEM4 and GHCN-M v3, respectively. All are slightly larger than the raw ("unadjusted" in Section 3.2) data (0.107±0.009 °C per decade). The difference between the raw and adjusted data happens mainly before the 1950s.

Finally, we consider the most recent period (1998-2015) in more detail. During this period, Li et al (2015) states that a hiatus in warming occurs in most parts of China except for the southwest (also by Duan and Xiao, 2015). Although 2015 is the 2nd warmest year on record (2007 is the warmest year), the temperature change trend for 1998-2015 remains 0.079±0.13 °C per decade, which is still much smaller than estimates for 1951-2015 (0.244±0.021°C per decade) and for 1979-2015 (0.379±0.044 °C per decade) (Figure 7). This agrees well with Li et al (2015), though the recent trend is slightly larger than the 1998-2012 period trends. China averages calculated by both CRUTEM4 and GHCN-M v3 show negative trends during the 1998-2015 period, both of which are related to the slight underestimation of the temperature in recent years. It is worth noting that over this period none of the series have trends that are statistically significant at the 5% level.

FIG. 8

TABLE 2

5. Conclusion and prospect

With the motivation to recommend the best estimate of China average SAT series to the international climate research community, we have assessed the similarities and differences of the existing time series of annual average SAT for China. These time series are based upon historical meteorological observations since the 1900s, and the observed time series are further compared to the historical climate simulation of CMIP5 models and the 20th century reanalysis. The major findings are summarized below.

1) Although there are still certain differences among the existing China average time series during the 1940s warmth period due to the scant number of observations, SAT time series using homogenized temperature data have improved the agreement of long-term changes among all the datasets since the 1900s. The similarity of the China observational average with the NH series, noted by Bradley et al (1987), is confirmed, but the NH series peaks slightly earlier (1938) than China (1946).

2) SAT time series using homogenized temperature data have also improved the agreement of long-term changes between the observational and model simulation datasets since the 1900s. Several Chinese models in CMIP5 show better performance in the simulation of annual SAT variations over China in the past century in terms of correlation coefficients (all above 0.6). These models consistently show reasonable reproductions of decadal-scale variations of SAT in the past century, but tend to over-estimate recent trends during the past 15 years. 20CR also reflects the inter-decadal variations of China's SAT at the century scale, but greatly under-estimates the temperature in the past two decades.

3) A new evaluation of the long-term trend based on the homogenized observed data was achieved: SAT in China reveals a warming trend (0.121±0.009°C per decade)

slightly greater than what the raw data indicates (0.107±0.009°C per decade) from 1900-2015. The best estimations of linear change trends with 95% uncertainty ranges are 0.121±0.009, 0.244±0.021, 0.379±0.045 and 0.079±0.13°C (not significant) per decade for 1900-2015, 1951-2015, 1979-2015, and 1998-2015, respectively. Thus the average of SAT "warming" over China also shows a "slow down" in the past 18 years (1998-2015).

Acknowledgements: Sincere and special thanks go to Prof. Shaowu Wang from Peking University, who had reviewed the draft of the manuscript and provided many valuable comments and suggestions. We express our deepest condolence to his passing on January 11, 2015 due to sickness. This study is supported by: China Meteorological Administration Special Foundation for Climate Change (CCSF201338); Ministry of Science and Technology of China, No. GYHY201406016; Tianjun Zhou is supported by National Natural Science Foundation of China under Grant No. 41330423.

References

Bradley, R. S., H. F. Diaz, P. D. Jones, and P. M. Kelly, 1987: Secular fluctuations of temperature over the Northern Hemisphere land areas and mainland China since the mid-19th century. In, The Climate of China and Global Climate, D. Ye, C. Fu, J. Chao and M. Yoshino, Eds., China Ocean Press, 76-87.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, 111, D12106, doi:10.1029/2005JD006548.

Cao, L., P. Zhao, Z. Yan, P. D. Jones, Y. Zhu, Y. Yu and G. Tang, 2013: Instrumental temperature series in eastern and central China back to the 19th century. *J. Geophys. Res.*, 118, 8197-8207, doi:10.1002/jgrd.50615.

Chen, L. X., and Coauthors, 2004: Characteristics of the Climate Change and its Formation Mechanism in China in Last 80 Years. *Journal of Meteorological Research*, 62, 634-646, doi: 10.3321/j.issn:0577-6619.2004.05.010.

Cheng, X. Y., Y. H. Wang, G. C. Li, and Q. X. Li, 2013: Evaluation of three reanalysis reanalysis precipitation datasets in China. *Advances in Climate Change Research,* 9, 258-265, doi:10.3969/j.issn.1673-1719.2013.04.004.

Compo, G. P., J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, Jr., R. S. Vose, G. Rutledge, and P. Bessemoulin, 2011: Review Article The Twentieth Century reanalysis

Proiect. *Quart J Roy Meteor Soc*, 137, 1-28, doi:10.1002/qj.776.

——, P. D. Sardesmukh, J. S. Whitaker, P. Brohan, P. D. Jones, and C. McColl, 2013: Independent confirmation of global land warming without the use of station thermometers. *Geophys. Res. Lett.*, 40, 3170-3174, doi:10.1002/grl.50425.

Ding, Y. H., and G. Y. Ren (ed.), 2008: Introduction to China's Climate Change Sciences, China Meteorological Press, 281.

——, and X. S. Dai, 1994: Temperature Variation in China during the Last 100 years, Meteorology, 20, 19-26.

Duan, A. and Xiao, Z., 2015: Does the climate warming hiatus exist over the Tibetan Plateau? *Sci. Rep.* 5, 13711; doi: 10.1038/srep13711

Easterbrook, D. J., 2008: Global cooling is here—evidence for predicting global cooling for the next three decades. http://www.globalresearch.ca/index.php?context=&aid=10783.

Easterling, D. R., and M. F. Wehner, 2009: Is the climate warming or cooling? *Geophys Res Lett* 36:L08706, doi: 10.1029/2009GL037810.

Fyfe, J. C., N. P. Gillett, and F. W. Zwiers, 2013: Overestimated global warming over the past 20 years. *Nat. Clim. Change*, 3, 767–769, doi:10.1038/nclimate1972.

Guemas, V., F. J. Doblas-Reyes, I. Andreu-Burillo, and M. Asif, 2013: Retrospective prediction of the global warming slowdown in the past decade, *Nat. Clim. Change*, 3, 649–653, doi:10.1038/nclimate1863.

Gullet, D. W., L. Vincent, and P. J. F. Sajecki, 1990: Testing for homogeneity in temperature time series at Canadian climate stations, CCC Report No. 90–4, Atmospheric Environment Service, Downsview, Ontario. 43 pp.

Guo, Y., W. J. Dong, F. M. Ren, Z. C. Zhao, and J. B. Huang, 2013: Comparison of simulation of China's annual average air temperature by CMIP5 and CMIP3 models. *Advances in Climate Change Research*, 9, 181-186, doi:10.3969/j.issn.1673-1719.2013.03.004.

Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, 48, RG4004, doi:10.1029/2010RG000345.

Hartmann, D. L., A. M. G. Klein Tank, M. Rusticucci, L. V. Alexander, S. Brönnimann, Y. Charabi, F. J. Dentener, E. J. Dlugokencky, D. R. Easterling, A. Kaplan, B. J. Soden, P. W. Thorne, M. Wild and P. M. Zhai, 2013: Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hua, L. J., Z. G. Ma, and W. D. Guo, 2008: The impact of urbanization on air temperature across China. *Theor Appl Climatol*, 93, 179-194, doi: 10.1007/s00704-007-0339-8.

Jiang D B and Z P. Tian, 2013: East Asian monsoon change for the 21st century: Results of CMIP3 and CMIP5 models. *Chin Sci Bull*, 58: 1427 ・ 1435, doi: 10.1007/ s11434-012-5533-0

Jin, C. and T. Zhou, 2014: Analysis of the Interannual Variations of the East Asian Winter Monsoon

Simulation by Four CMIP5 GCMs. *Chin. J. Atmos. Sci.*, 38, 453-468, doi:10.3878/j.issn.1006-9895.2013.13180.

Jones, P.D., Osborn, T.J. and Briffa, K.R., 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate* **10**, 2548-2568.

Jones P., D. H. Lister, and Q. Li, 2008: Urbanization effects in large-scale temperature records, with an emphasis on China. *J. Geophys. Res.*, 113, D16122, doi:10.1029/2008JD009916.

——, ——, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice, 2012: Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *J Geophys Res-Atmos*, 117, doi:10.1029/2011JD017139.

——, P. Ya. Groisman, M. Coughlan, N. Plummer, W.-C. Wang, and T. R. Karl, 1990: Assessment of urbanization effects in time series of surface air temperature over land. Nature, 347, 169-172, doi:10.1038/347169a0.

Karl, T. R., A. Arguez, B. Huang, J.H. Lawrimore, J.R. McMahon, M.J. Menne, T.C. Peterson, R.S. Vose, and H. Zhang, Possible artifacts of data biases in the recent global surface warming hiatus, Science, 2015. http://dx.doi.org/10.1126/science.aaa5632

——, and C. N. Williams Jr., 1987: An approach to adjusting climatological times series for discontinuous in homogeneities. *J. Clim. Appl. Meteorol.*, 26, 1744–1763.

Kaufmann, R. K., H. Kauppi, M. L. Mann, and J. H. Stock, 2011: Reconciling anthropogenic climate change with observed temperature 1998-2008. *Proc. Nat. Acad. Sci.*, 108, 11790-11793, doi: 10.1073/pnas.1102467108.

Lawrimore, J. H., M. J. Menne, B. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, and J. Rennie, 2011: An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *Journal of Geophysical Research-Atmospheres*, 116, doi:10.1029/2011jd016187.

Lewandowsky, S., J. Risbey, and N. Oreskes, 2015: The "Pause" in Global Warming: Turning a Routine Fluctuation into a Problem for Science. *Bull. Amer. Meteor. Soc.* doi:10.1175/BAMS-D-14-00106.1, in press.

Li, Q. X., H. Zhang, J. Chen, W. Li, X. Liu and P. Jones, 2009: A mainland China homogenized historical temperature dataset of 1951-2004. *Bull. Amer. Meteor. Soc.*, 90, 1062-1065, doi: 10.1175/2009BAMS2736.1.

——, ——, X. Liu and J. Huang, 2004a: Urban heat island effect on annual mean temperature during the last 50 years in China. *Theor. Appl. Climat.*, 79, 165-174, doi: 10.1007/s00704-004-0065-4.

——, X. Liu, H. Zhang, T. C. Peterson and D. R. Easterling, 2004b: Detecting and adjusting temporal inhomogeneity in Chinese mean surface air temperature data. *Adv., Atmos., Sci.*, 21, 260-268, doi: 10.1007/BF02915712.

——, J. Huang, Z. Jiang, L. Zhou, P. Chu and K. Hu, 2014: Detection of urbanization signals in extreme winter minimum temperature changes over northern China. *Climatic Change*, doi: 10.1007/s10584-013-1013-z

——, S. Yang, W. Xu, X. Wang, P. Jones, D. Parker, L. Zhou and Y. Feng, 2015: China experiencing the recent warming hiatus, *Geophys. Res. Lett.*, 42, doi:10.1002/2014GL062773.

——, W. Dong, W. Li, X. Gao, P. Jones, J. Kennedy and D. Parker, 2010a: Assessment of the uncertainties in temperature change in China during the last century. *Chin. Sci. Bull.*, **55**, 1974-1982, doi: 10.1007/s11434-010-3209-1.

——, W. Li, P. Si, X. Gao, W. Dong, P. Jones, J. Huang, and L. Cao, 2010b: Assessment of surface air warming in northeast China, with emphasis on the impacts of urbanization. *Theor. Appl. Climatol.*, 99. doi: 10.1007/s00704-009-0155-4.

Li, Z., and Z. Yan, 2009: Homogenized daily mean/maximum/minimum temperature series for China from 1960-2008. *Atmos Ocean Sci Lett*, 2, 237–243.

Peterson, T. C., and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, 18, 1493–1517. doi: 10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T

Qin, D. H., Y. H. Ding, J. L. Su (ed.), 2005: Evolution of China's Climate and Environment (Vol.1) : Science Press, 63-103.

Ren, G. Y., and Coauthors, 2008: Urbanization effects on observed surface air temperature trends in North China. *J. Climate*, 21, 1333-1348, doi:10.1175/2007JCLI1348.1

Ren, Y. Y. and G. Y. Ren, 2011: A remote-sensing method of selecting reference stations for evaluating urbanization effect on surface air temperature trends. *J. Climate*, 24, 3179-3189, doi:10.1175/2010JCLI3658.1.

Slingo, J., and Coauthors, 2013: The recent pause in global warming parts 1-3Rep., The Met Office, FitzRoy Road, Exeter, U. K.

Song, F., Q. Hu, and W. H. Qian, 2004: Quality control of daily meterological data in China, 1951-2000: A new dataset. *Int. J. Climat.*, 24, 853-870. doi:10.1002/joc.1047

Tang G., and G. Y. Ren, 2005: Reanalysis of Surface Air Temperature Change of the Last 100 Years over China. *Climatic & Environmental Research*, 10, 791-798.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research-Atmospheres*, 106, D7, 7183-7192

Taylor, K. E., R. J. Stouffe, G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, 93, 485-498. doi:10.1175/BAMS-D-11-00094.1

Tu, Q. P., 1984: Trend and cycle of China's air temperature variation in the past century. *Journal of Nanjing Institute of Meteorology*, 7, 151-162.

Wang, F., Q. Ge, S. Wang, Q. Li, P. Jones, 2015: A new estimation of urbanization contribution to the warming trend in China, *J. Climate*, 28:8923-8938, doi: 10.1175/JCLI-D-14-00427.1

Wang, J., C. Xu, M. Hu, Q. Li, Z. Yan, P. Zhao and P. D. Jones, 2014: A New Estimate of the China Temperature Anomaly Series and Uncertainty Assessment in 1900-2006. *J. Geophys. Res.* 119, 1-9, doi:10.1002/jgrd.51046.

Wang, S., and D. Gong, 2000: Enhancement of the warming trend in China. *Geophys. Res. Lett.*, 27, 2581-2584, doi: 10.1029/1999GL010825.

——, J. L. Ye, and D. Y. Gong, 1998: Construction of mean annual temperature series for the last

one hundred years in China. Quart. *J. Appl. Meteor. Sci.*, 9, 392-401. (in Chinese)
Wesley E，A Method to Estimate the Statistical Significance of a Correlation When the Data Are Serially Correlated，J. Climate，1997，10，2147-2151.
Xu, W., Q. Li, X. L. Wang, S. Yang, L. Cao, and Y. Feng, 2013: Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. *J. Geophys. Res. Atmos.*, 118, doi:10.1002/jgrd.50791.
Xu, W., Q. Li, P. Jones, et al., 2016: A new integrated and homogenized global monthly land surface air temperature dataset for the period since 1900, Submitted to J. Climate.
Xu, Y., and C. H. Xu, 2012: Preliminary Assessment of Simulations of Climate Changes over China by CMIP5 Multi-Models, *Atmos Ocean Sci Lett*, 5, 489−494.
Yan, Z., Z. Li, Q. X. Li and P. D. Jones, 2009: Effects of Site-Change and Urbanisation in the Beijing Temperature Series 1977-2006, *Int. J. Climat.*, 30, 1226–1234, doi:10.1002/joc.1971
Yang, X. C., Y. L. Hou, and B. D. Chen, 2011: Observed surface warming induced by urbanization in east China. *J. Geophys. Res.*, 116, D14113, doi: 10.1029/2010JD015452.
Zhai, P., Q. Chao, and X. Zou, 2004: Progress in China's climate change study in the 20th century, *J. Geograph. Sci.*, 14, 3-11, doi: 10.1007/BF02841101.
Zhang, J., and Coauthors, 2005: Impact of land use changes on surface warming in China. *Adv Atmos Sci*, 22, 343-348, doi: 10.1007/BF02918748.
Zhang, L., Y. H. Ding, T. W. Wu, and X. G. Xin, 2013: The 21st century annual mean surface air temperature change and the 2℃ warming threshold over the globe and China as projected by the CMIP5 models, *Journal of Meteorological Research*, 71, 1047-1065, doi:10.11676/qxxb2013.087.
Zhang, X. G., and X. Q. Li, 1982: Some Characteristics of Temperature variation in China in the Present Century, *Journal of Meteorological Research*, 40, 198-208.
Zhou, L., and Coauthors, 2004: Evidence for a significant urbanization effect on climate in China. *Proc. Nat. Acad. Sci.*, 101, 9540-9544, doi: 10.1073/pnas.0400357101
Zhou, T., and R. Yu, 2006: Twentieth-century surface air temperature over China and the globe simulated by coupled climate models, *J. Climate*, 19, 5843-5858, doi:10.1175/JCLI3952.1.

Captions:

TABLE 1. Six centennial-scale temperature change series across China

TABLE 2. Comparisons of the raw and homogenized SAT trends during different periods (unit: °C /10a)

FIG. 1. Six SAT time series across China in the past century (1909-2006) (1961-1990 as the base period)

FIG. 2. Comparison of SAT series across China and the Northern Hemisphere (NH_CRU) (1961-1990 as the base period)

FIG. 3. China SAT series' variations simulated by 41 models in CMIP5 in the past century (1961-1990 as the base period). The red, black and blue lines show the Li series, 20CR and multi-model average, respectively.

FIG. 4. Comparison of long-term temperature change trends for 41 models in CMIP5 and 6 observational series. a) 1909-2005, b) 1951-2005 (the bar indicates the trend distribution of 41 models (the black line is a fitted line), the straight lines indicate the trends of 6 observational series, and the red line indicates the Li series trend).

FIG. 5. Taylor diagram displaying normalized pattern statistics of surface air temperature over China between the 41 climate models and observation data (Li) for the period from 1900-2005. Each number represents a Model number (see right column), and the Li series is considered the reference (REF) dataset. Standard deviation and centered root mean square differences are normalized by the reference standard deviation. The radial distance from the origin is the normalized standard deviation of a model; the correlation between a model and the reference is given by the azimuthal position of the model, with the oblique dotted line indicating the 99% confidence level; and the centered root mean square difference between a model and the reference is their distance apart.

FIG. 6. Series of China SAT variations simulated by the 8 top-ranking models in CMIP5 and 20CR datasets (1961-1990 as base period). upper panel: ranking by correlation coefficients only; bottom panel: ranking by correlation coefficients and normalized standard deviation.

FIG. 7. The distribution of the long-term stations used to construct the west China series (upper panel) and the W China series and the 30 station series. (bottom panel, 1961-1990 as the base period)

FIG.8. Anomaly time series of China's average annual SAT from 1900-2015 (1961-1990 as the base period)

TABLE 1. Six centennial-scale temperature change series across China

Series	Data coverage	Quality control	Mean temperature	Calculation method of regional series	References
WangS	Observations from CMA and proxy data	—	The arithmetic average of fixed observations (the times were different)	Sub-regional series by arithmetic means of single series and national series by area weighting average from 10 regional series	Wang et al（1998）
Tang	Observations from CMA	—	The arithmetic average of Tx and Tn	Gridding then calculating the regional series by the Climate Anomaly Method (CAM)	Tang and Ren （2006）
Li	Same as above	Homogenized station series	Same as above	CAM	Li et al (2010a)
WangJ	Same as above	Same as above	Same as above	BSHADE-MSN, which takes into account prior knowledge of geographical spatial autocorrelation and nonhomogeneity of target domains, remedies the biased sample and maximizes an objective function for the best linear unbiased estimation of the regional mean quantity	Wang et al (2014)
CRUTEM3	Temperature data CRU collected and processed	Same as above	Same as above	CAM	Brohan et al (2006).
CRUTEM4	Temperature data CRU collected and processed	Same as above	Same as above	CAM	Jones et al (2012).

TABLE 2 Comparisons of the raw and homogenized SAT trends during different periods (unit: °C /10a)

	1900-2015	1951-2015	1979-2015	1998-2015
RAW	0.107*±0.010	0.247*±0.021	0.381*±0.045	0.059±0.13
ADJ	0.121*±0.009	0.244*±0.021	0.379*±0.044	0.079±0.13
CRUTEM4	0.130*±0.009	0.243*±0.021	0.348*±0.051	-0.150±0.14
GHCN	0.114*±0.009	0.215*±0.021	0.297*±0.049	-0.093±0.14

* The trend has passed the significant test of 95%

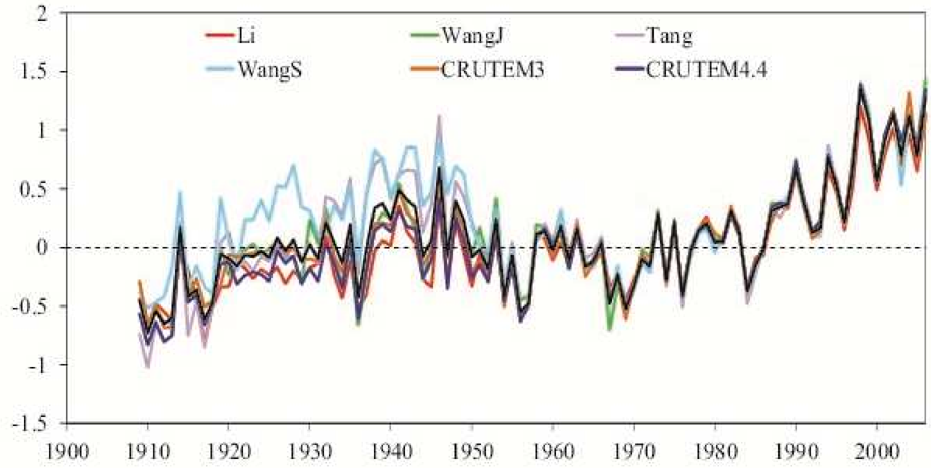


FIG. 1. Six SAT time series across China in the past century (1909-2006)
(1961-1990 as the base period)

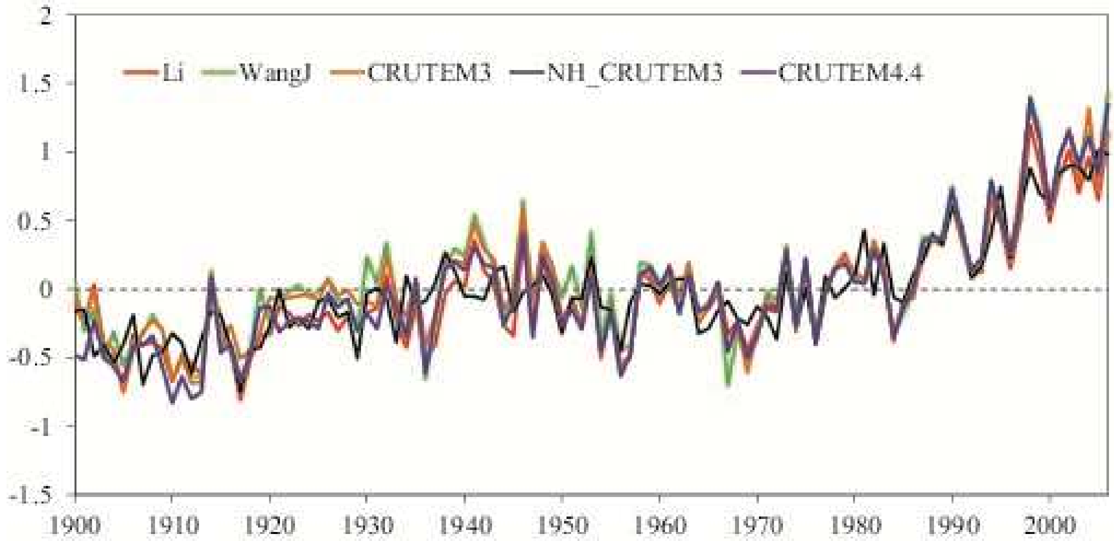


FIG. 2. Comparison of SAT series across China and the Northern Hemisphere (NH_CRU) (1961-1990 as the base period)

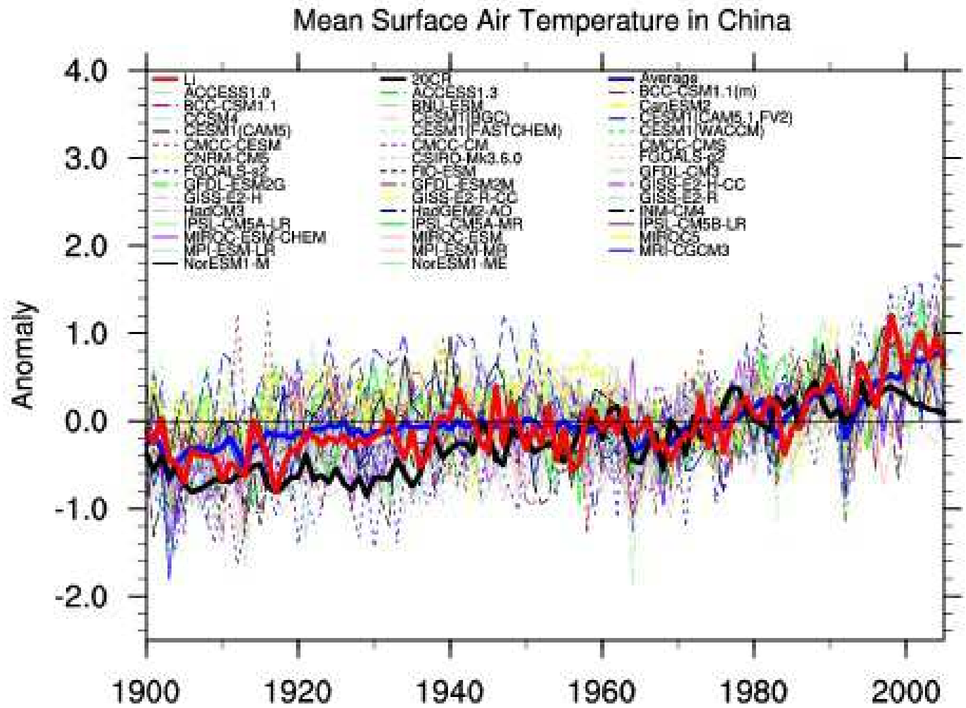


FIG. 3. China SAT series' variations simulated by 41 models in CMIP5 in the past century (1961-1990 as the base period). The red, black and blue lines show the Li series, 20CR and multi-model average, respectively.

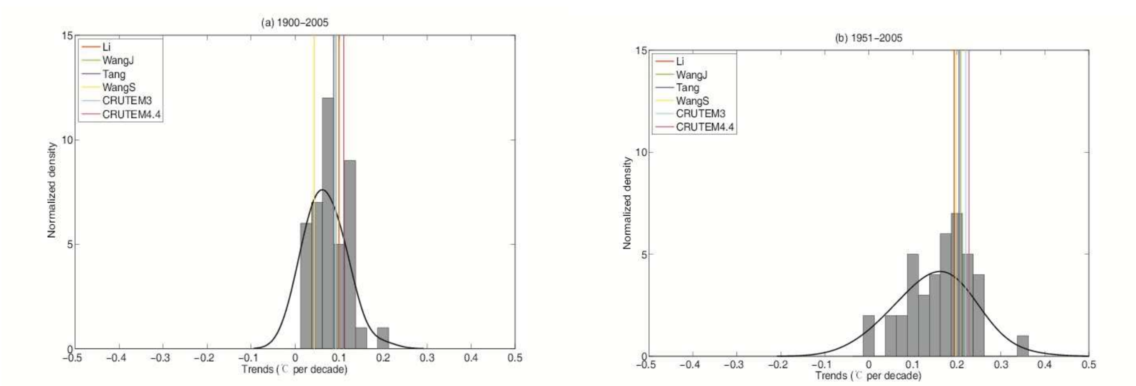


FIG. 4. Comparison of long-term temperature change trends for 41 models in CMIP5 and 6 observational series. a) 1909-2005, b) 1951-2005 (the bar indicates the trend distribution of 41 models (the black line is a fitted line), the straight lines indicate the trends of 6 observational series, and the red line indicates the Li series trend).

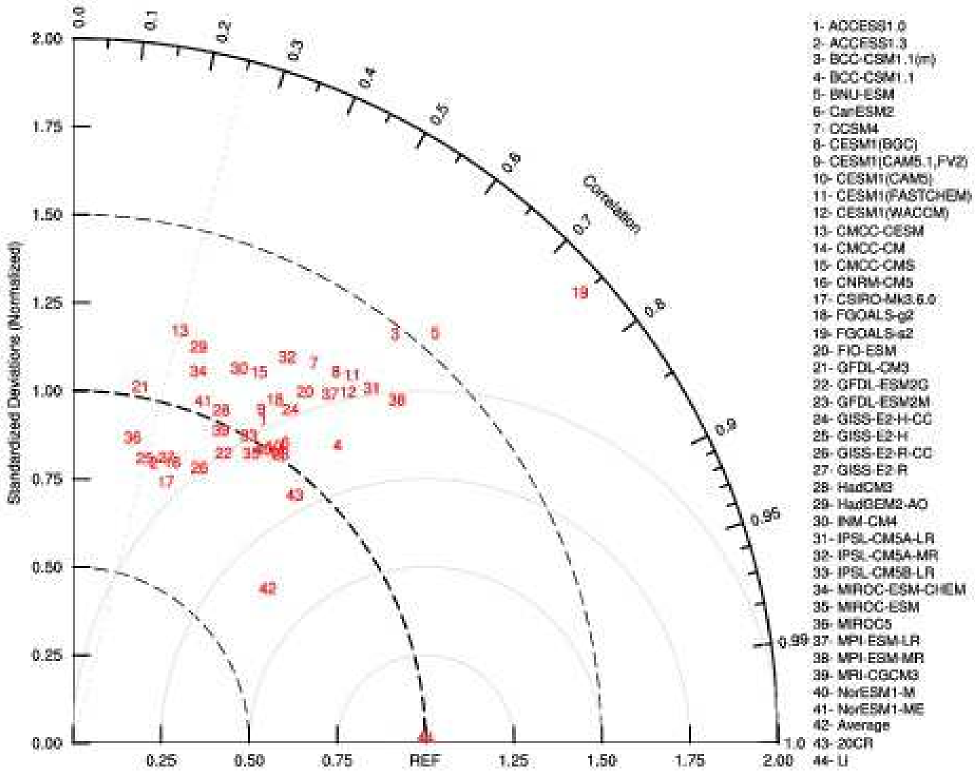
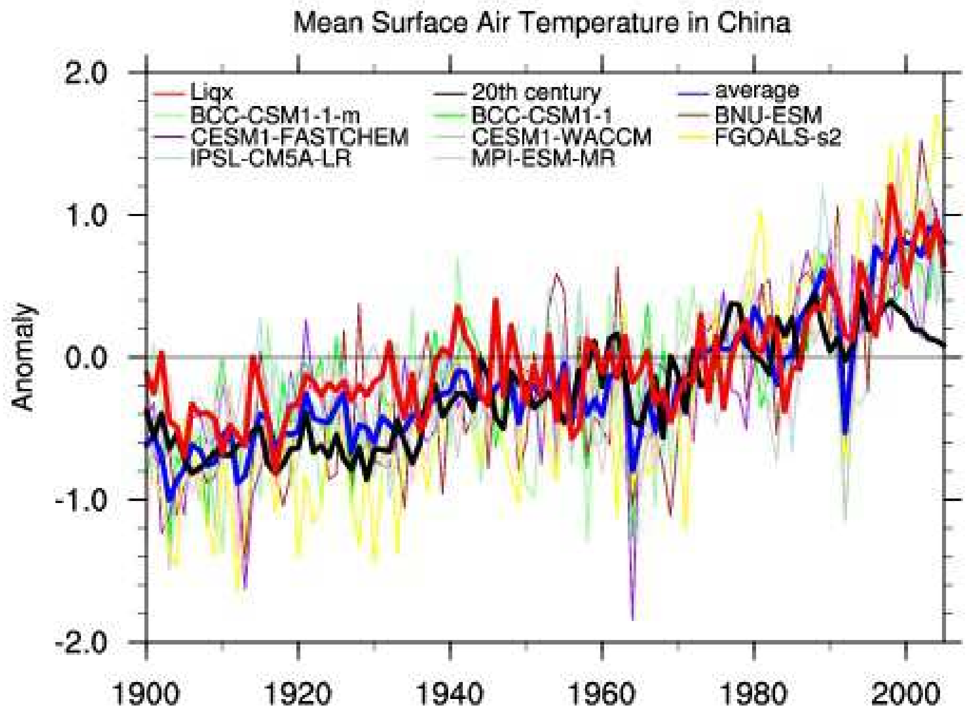


FIG. 5. Taylor diagram displaying normalized pattern statistics of surface air temperature over China between the 41 climate models and observation data (Li) for the period from 1900-2005. Each number represents a Model number (see right column), and the Li series is considered the reference (REF) dataset. Standard deviation and centered root mean square differences are normalized by the reference standard deviation. The radial distance from the origin is the normalized standard deviation of a model; the correlation between a model and the reference is given by the azimuthal position of the model, with the oblique dotted line indicating the 99% confidence level; and the centered root mean square difference between a model and the reference is their distance apart.



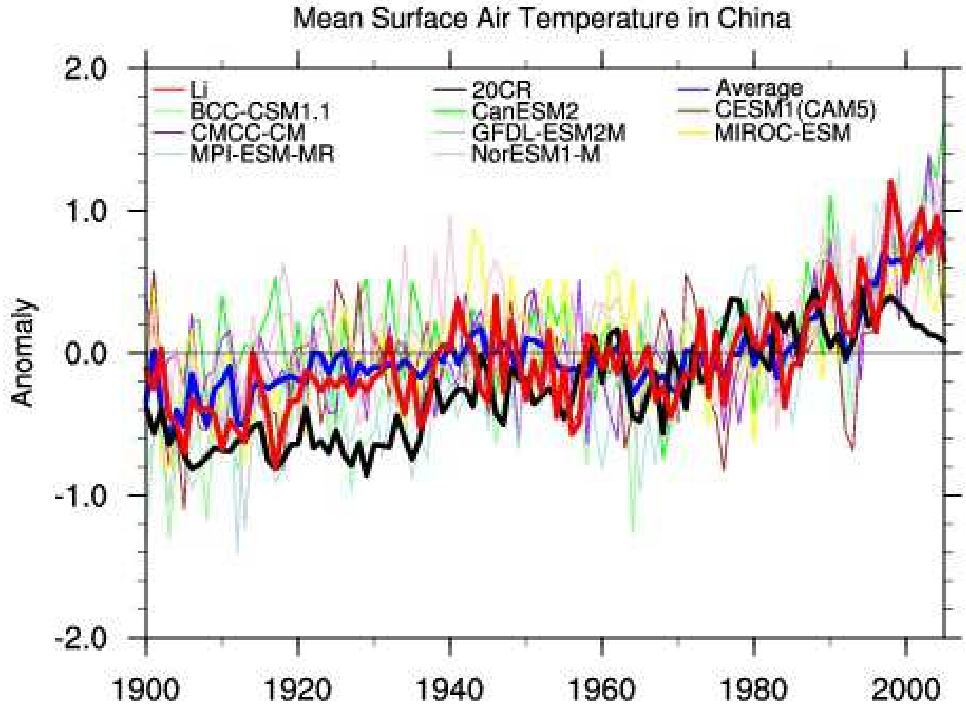


FIG. 6. Series of China SAT variations simulated by the 8 top-ranking models in CMIP5 and 20CR datasets (1961-1990 as base period). upper panel: ranking by correlation coefficients only; bottom panel: ranking by correlation coefficients and normalized standard deviation.

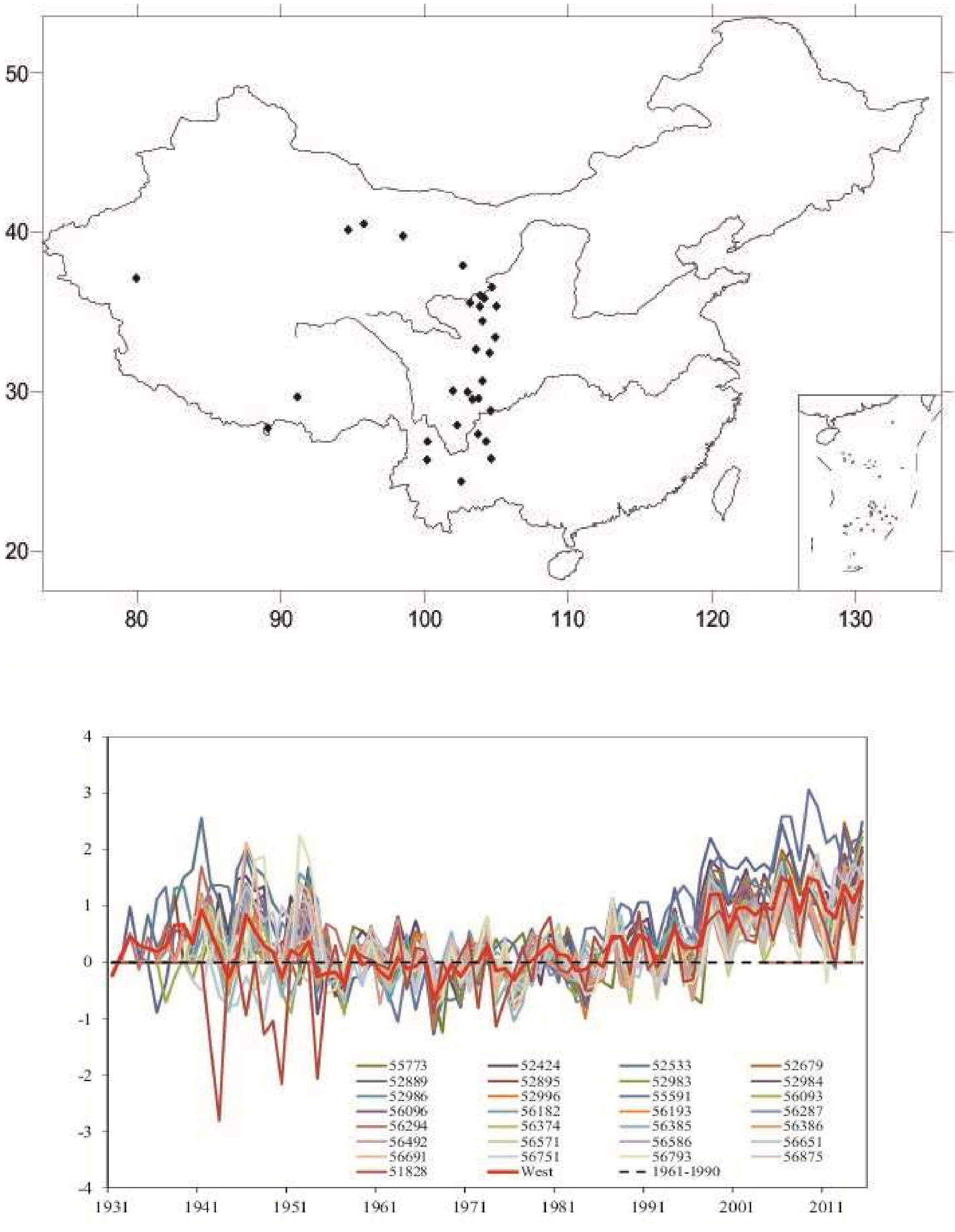


FIG. 7. The distribution of the long-term stations used to construct the western China series (upper panel) and the W China series and the 30 station series (bottom panel, 1961-1990 as the base period)

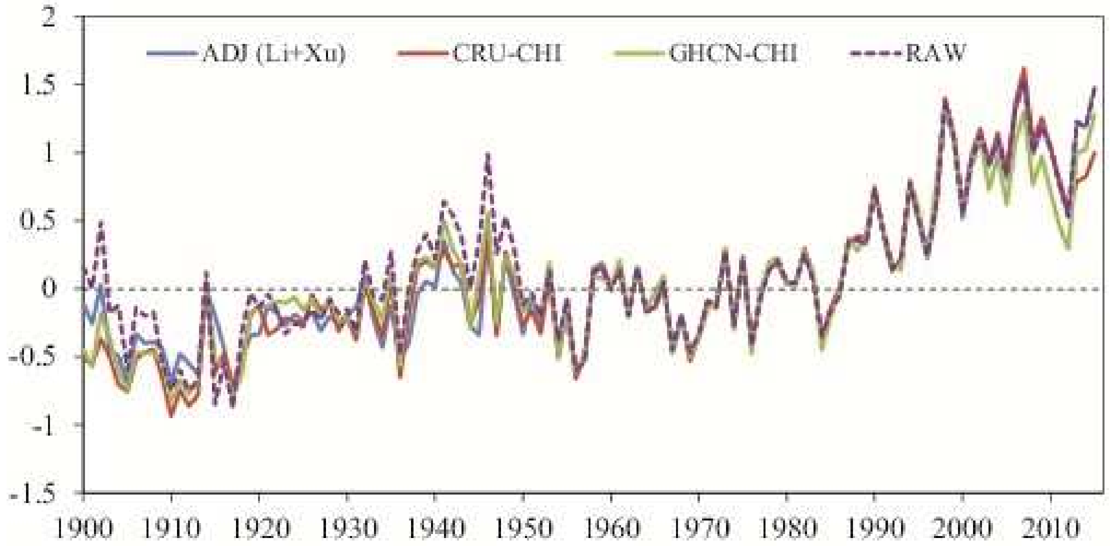


FIG. 8. Anomaly time series of China's average annual SAT from 1900-2015 (1961-1990 as the base period)