


A new integrated and homogenized global monthly land surface air temperature dataset for the period since 1900

Wenhui Xu¹ · Qingxiang Li²  · Phil Jones^{3,4} · Xiaolan L. Wang⁵ · Blair Trewin⁶ · Su Yang¹ · Chen Zhu¹ · Panmao Zhai⁷ · Jinfeng Wang⁸ · Lucie Vincent⁵ · Aiguo Dai⁹ · Yun Gao¹⁰ · Yihui Ding¹¹

Received: 7 September 2016 / Accepted: 6 June 2017 / Published online: 22 June 2017
© Springer-Verlag GmbH Germany 2017

Abstract A new dataset of integrated and homogenized monthly surface air temperature over global land for the period since 1900 [China Meteorological Administration global Land Surface Air Temperature (CMA-LSAT)] is developed. In total, 14 sources have been collected and integrated into the newly developed dataset, including three global (CRUTEM4, GHCN, and BEST), three regional and eight national sources. Duplicate stations are identified, and those with the higher priority are chosen or spliced. Then, a consistency test and a climate outlier test are conducted to ensure that each station series is quality controlled. Next, two steps are adopted to assure the homogeneity of the station series: (1) homogenized station series in existing national datasets (by National

Meteorological Services) are directly integrated into the dataset without any changes (50% of all stations), and (2) the inhomogeneities are detected and adjusted for in the remaining data series using a penalized maximal t test (50% of all stations). Based on the dataset, we re-assess the temperature changes in global and regional areas compared with GHCN-V3 and CRUTEM4, as well as the temperature changes during the three periods of 1900–2014, 1979–2014 and 1998–2014. The best estimates of warming trends and their 95% confidence ranges for 1900–2014 are approximately 0.102 ± 0.006 °C/decade for the whole year, and 0.104 ± 0.009 , 0.112 ± 0.007 , 0.090 ± 0.006 , and 0.092 ± 0.007 °C/decade for the DJF (December, January, February), MAM, JJA, and SON seasons, respectively. MAM saw the most significant warming trend in both 1900–2014 and 1979–2014. For an even shorter and more recent period (1998–2014), MAM, JJA and SON

Electronic supplementary material The online version of this article (doi:10.1007/s00382-017-3755-1) contains supplementary material, which is available to authorized users.

✉ Qingxiang Li
liqingx5@mail.sysu.edu.cn; liqx@cma.gov.cn

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

² School of Atmospheric Sciences, Sun Yat-sen University, 135 Xingangxi Road, Guangzhou 510275, China

³ Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK

⁴ Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

⁵ Climate Research Division, Science and Technology Branch, Environment and Climate Change Canada, Toronto, ON, Canada

⁶ Australian Bureau of Meteorology, Melbourne, VIC, Australia

⁷ China Meteorological Administration, China Academy of Meteorological Sciences, 46 Zhongguancun South Avenue, Haidian District, Beijing 100081, China

⁸ LREIS, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

⁹ Department of Atmospheric and Environmental Sciences, University at Albany, SUNY, Albany, NY 12222, USA

¹⁰ China Meteorological Administration, 46 Zhongguancun South Avenue, Haidian District, Beijing 100081, China

¹¹ China Meteorological Administration, National Climate Center, 46 Zhongguancun South Avenue, Haidian District, Beijing 100081, China

show similar warming trends, while DJF shows opposite trends. The results show that the ability of CMA-LAST for describing the global temperature changes is similar with other existing products, while there are some differences when describing regional temperature changes.

Keywords CMA-LSAT dataset · Surface air temperature · Homogenized · Climate change

1 Introduction

Land surface air temperature (LSAT) change is a primary measure of global climate change (Hartmann et al. 2013; Jones 2016). IPCC AR5 has cited four global monthly land surface temperature datasets, including CRUTEM4 from the Climatic Research Unit (CRU) of the University of East Anglia, the Global Historical Climatology Network (GHCN) dataset from NOAA's National Center for Environment Information [NCEI; formerly National Climatic Data Center (NCDC)], GISTEMP from NASA's Goddard Institute of Space Studies, and the Berkeley Earth Surface Temperature (BEST) dataset. Of these datasets, the first version of the GHCN monthly air temperature dataset was developed by NCEI in the early 1990s (Vose et al. 1992). GHCN-V3, containing 7280 stations, was released in 2011, with improved quality control of duplicate data, climate anomalies, and spatial inconsistencies (Durre et al. 2007; Lawrimore et al. 2011). Homogeneity testing and adjustment of the temperature series was conducted by the automatic paired alignment approach (Menne and Williams 2009). An independent effort in the United Kingdom produced a first release of CRUTEM in the late 1980s. Today, a global dataset of over 6000 stations is still maintained in the fourth iteration of CRUTEM (CRUTEM4, Jones et al. 2012). GISTEMP (Hansen et al. 1999) was developed on the basis of the original data of GHCN-V2, introducing data from several stations in Antarctica and data from the homogenized US Historical Climatology Network (USHCN) consisting of more than 1200 stations. The BEST team combined 16 datasets to build an integrated dataset of global monthly surface air temperature, and a new algorithm has been developed that can utilize short and/or discontinuous data. After the removal of duplicate records, this dataset incorporates 36,000 stations with various record length (Rohde et al. 2013). These datasets have provided the scientific basis for quantifying and detecting climate change over land. A large dataset, containing about 32,000 stations, has also been released as part of the International Surface Temperatures Initiative (ISTI) (Rennie et al. 2014). Whilst the ISTI dataset itself is not homogenized, the data it contains have been used as the basis for a number of datasets which have performed their own

homogenization, including that of Karl et al. (2015) and the next update of the GHCN dataset (GHCNv4, not yet published at the time of writing). The Japan Meteorological Agency (JMA) also maintains a global LSAT dataset at its website (http://ds.data.jma.go.jp/tcc/tcc/products/gwp/temp/ann_wld.html), but no detailed information about the data quality control and homogenization can be found in the scientific literature, and it is only independent of GHCN after 2001. Exploiting reanalyses and using satellite data, as done in Simmons et al. (2017) and Cowtan and Way (2014) have the potential to refine the conventional approach into the future but are outside the scope of this paper.

Although methodologies for developing all four station datasets differ (for example, CRUTEM4 incorporates homogenized national data sets where available, whereas GHCN and BEST carry out their own homogenization), they all exhibit close agreement with respect to large-scale changes in global and hemispheric LSAT. However, there are differences with respect to regional climate change, especially in South America, Africa, and Asia, and for periods before 1950 (Jones 2016). Also, because they differ in terms of data collection, processing techniques and focus, these datasets exhibit subtle differences when describing global-average LSAT changes, although they have generally become closer to each other during the twentieth century (IPCC 2007; Hartmann et al. 2013). All the groups use much the same input data, but there are still some differences between regions and continents in terms of stations and data quality. Moreover, they employ different approaches to ensure data quality, to interpolate and develop gridded products, to construct the global/regional climate change series, and to calculate/communicate the error assessment. All four datasets have relatively low station densities (through GTS, detailed discussion in Sect. 6) over Asia compared with the United States and Europe, especially for large countries such as China, Russia, and India (Li et al. 2016; Jones 2016). The China Meteorological Administration (CMA) has proposed a plan to develop global baseline temperature and precipitation datasets to fulfill the needs for regional (especially Asian) and global climate monitoring and climate change studies (Li 2013).

In this study, a new dataset of integrated and homogenized monthly surface air temperature over global land (referred to as the China Meteorological Administration global Land Surface Air Temperature dataset, CMA-LSAT) is developed and applied to build a new global homogenized LSAT change series for the period since 1900.

The CMA-LSAT dataset follows a similar philosophy to the CRUTEM4 dataset, in that it employs homogenized datasets developed at the national or regional level where possible, with other sources being used only where suitable homogenized national or regional datasets do not exist. This recognizes that national

institutions are likely able to perform homogenization more effectively than is possible in a global dataset, principally due to better access to metadata, and access to more potential reference stations (as most countries have substantial domestic networks whose data are not transmitted internationally). It is considered that the advantages of being able to draw on this national-level information outweighs any disadvantage which may arise from the differences in the actual homogenization techniques used in different parts of the dataset.

The most substantial advance in this dataset, relative to other global datasets cited earlier, is the improved station coverage, especially in Asia. Compared to existing datasets, CMA-LSAT shows significant improvements in data coverage in most countries of Asia, especially in China and its neighboring regions, while its station coverage in Africa and South America is comparable to that of the existing global datasets such as CRUTEM4 and GHCN-V3.

The remainder of this article describes the development of the dataset and provides comparisons with other existing global datasets. Section 2 describes the main data resources and the integration principle. Data quality control and homogenization are highlighted in Sects. 3 and 4, respectively. Section 5 summarizes and discusses the characteristics of the merged datasets at the regional scale. Section 6 summarizes the new evaluation of large scale surface air temperature change trends. Discussions and conclusions are presented in Sect. 7 and Appendix, respectively, along with plans for future updates.

2 Data sources and pre-processing

2.1 Data sources and merging hierarchy

Considerable efforts have been devoted to homogenize, create and compare climate datasets over China by scientists from CMA, Chinese universities, and the Chinese Academy of Sciences (Ding and Dai 1994; Li et al. 2004a, 2009, 2010a, 2015, 2016; Feng et al. 2004; Zhai et al. 2004; Li and Yan 2010; You et al. 2011; Xu et al. 2013; Cao et al. 2013; Sun et al. 2014; Wang et al. 2014; Yin et al. 2015). However, relatively few of these studies (Wang and Zhou 2015) have extended the global or hemispheric domain. In 2015, the National Meteorological Information Center (NMIC) released a station dataset of monthly LSAT over global land (referred to as CMA-LSAT). In this paper, we document this dataset, perform a preliminary analysis, and report results from early comparisons with other data groups.

A total of fourteen sources of LSAT station data has been collected and integrated to develop the CMA-LSAT dataset, including three global sources (CRUTEM4, GHCN-V3, and BEST), three regional sources, and eight national sources. Table 1 summarizes all the sources currently in the CMA-LSAT dataset. Some sources only provide data at the monthly timescale, while others have daily data (which were used to derive monthly data used in CMA-LSAT). Most station records include daily maximum (T_{\max}) and minimum (T_{\min}) temperatures, but in some cases only daily mean temperature data are available. At the time this project started, the ISTI database was unavailable to us, thus it was not included in the current version

Table 1 Summary of data sources used in CMA-LSAT

Sources	Priority	Name/Country	Time scale	Raw/QC/Hom	T_{\max}	T_{\min}	T_{avg}	References
Country	1	China	Monthly	Hom	Y	Y	Y	Xu et al. (2013) Li et al. (2010a)
		USA	Monthly	Hom	Y	Y	Y	Menne and Williams (2009)
		Russia	Daily	QC	Y	Y	Y	–
		Canada	Monthly	Hom	Y	Y	Y	Vincent et al. (2012)
		Australia	Daily	Hom	Y	Y	Y	Trewin (2013)
		Korea	Daily	Raw	Y	Y	Y	–
		Japan	Monthly	QC	Y	Y	Y	–
		Vietnam	Daily	Raw	Y	Y	Y	–
Regional	2	SCAR	Monthly	Raw	N	N	Y	–
		ECA&D	Monthly	Raw	Y	Y	Y	Klein et al. (2002)
Global	3	HISTALP	Monthly	Hom	N	N	Y	Auer et al. (2007)
	4	GHCN-M v3 (raw)	Monthly	QC	Y	Y	Y	Durre et al. (2007)
	5	BEST	Monthly	QC	Y	Y	Y	Rohde et al. (2013)
Global	6	CRUTEM4	Monthly	QC/Hom	N	N	Y	Jones et al. (2012)

QC quality controlled, Hom homogenized

of CMA-LSAT. We plan to include ISTI data in our future versions. Most of the Chinese station data are available only after 1900, and pre-1900 temperature data are also subject to a higher level of uncertainty as many countries had not yet standardized their instrument shelters (Trewin 2010; Parker 1994). Because of these, we chose 1900 as the start year of the CMA-LSAT dataset, which is updated continuously in near-real time (see “Appendix” for details) although data up to 2014 were used in this paper.

A number of different methods have been used by various countries to calculate daily (and hence monthly) mean temperature (T_m). While many countries use the mean of the T_{\max} and T_{\min} , others use the mean of the values at fixed hours, e.g., the mean of the temperatures at 0, 06, 12 and 18 h local time. There are systematic differences between mean temperatures calculated by these methods (Trewin 2004) and hence it is preferable to use a consistent method where possible. However, in some cases, only the mean temperature (not the maximum or minimum) are available and hence it is necessary to use whichever method was used for the calculation of the mean temperature source data. But we calculated T_m whenever the T_{\max} and T_{\min} are available, even if the T_m is also available.

Given the historical nature of data collection, sharing, and rescue, there are many cases where an individual station identifier exists in multiple data sources (potentially duplicate station records). In addition, owing to different collection and reprocessing techniques, the duplicate records do not necessarily have identical temperature values for the same station even though they are based upon the same fundamental measurements (Rennie et al. 2014). Therefore, a hierarchy of all the source datasets must be created before merging. Sources with higher priority take precedence over lower priority sources when more than one record for the same station and same time period exists. The priority order is based on a number of criteria. Due to the emphasis this dataset places on regional climate change, national weather sources (National Meteorological Service, NMSs) and hydrological service sources (National Meteorological and Hydrological Service, NMHSs, country sources) are the most desirable and are assigned the highest priority. Among them, the American (Menne and Williams 2009), Japanese, Canadian (Vincent et al. 2012), Australian (Trewin 2013), and Russian datasets have been developed or published by their national meteorological data centers, whereas the Vietnamese and South Korean datasets are obtained by exchange between countries. For the country sources, different countries incorporate different stations, so they do not include all the same stations, so the eight sources are all given the highest priority of one (Table 1). In this case, a higher priority is given to the regional source. ECA&D (Klein et al. 2002; Wijngaard et al. 2003) and HISTALP (Auer et al. 2007) both cover the

European region. ECA&D was recently updated when the importance of such provenance was explicitly recognized and given high priority of two. HISTALP is a multinational dataset of high quality, which has been put together by the Austrian meteorological service; it was given the priority of three because it only contained the monthly mean average temperature. The global sources are assigned a lower priority than above sources. GHCN-M is given the priority of four for its regular updates with monthly mean maximum and minimum temperatures. BEST is given a higher priority of five than CRUTEM4 because CRUTEM4 does not provide monthly means of daily maximum and of daily minimum temperatures (but CRU does incorporate these variables in their CRU TS series of datasets, Harris et al. 2014). These are preferred over monthly mean temperature because they can be directly used to calculate the monthly mean on a globally consistent basis, and because there is compelling evidence that many data artifacts affect daily maxima and minima differently (Williams et al. 2012).

2.2 Data integration

Once all of the sources had been collected and formatted, the data were merged into a single comprehensive dataset. The merge process was based upon metadata matching and data equivalence criteria. The merge process started from the highest priority data source (target) and ran progressively through the other sources (candidates). Each candidate station was compared to all target stations in two steps.

In the first step, each candidate station was run through all the target stations and two metadata criteria were calculated for identifying matching stations. The first metadata criterion considers the likelihood that the same station from two sources has different values for the longitude, latitude, and elevation of the station (e.g., coordinates rounded to one or two decimal places). Therefore, using the latitudes and longitudes, the geographical distance between the candidate station and each of the target stations was computed. If the distance was less than 5 km and the height difference was less than 50 m, the first criterion was met. The second metadata criterion considers the likelihood that the station names also differ, particularly for countries that were once colonial and have subsequently gained independence, or where the phonetic spelling of names may differ between sources. Therefore, the station name similarity was calculated using the Jaccard Index (JI) (Jaccard 1901; Rennie et al. 2014), which is defined as the intersection divided by the union of two sample sets, A and B:

$$JI = \frac{A \cap B}{A \cup B}. \quad (1)$$

If *JI* reaches 0.8, the second criterion is met. If the candidate station meets both of the above metadata criteria, it

is considered to match well with the target dataset and data comparison is continued to the second step. Otherwise, the candidate station is determined to be unique and is added to the target dataset as a new station. This process is not perfect (e.g., it is possible that two duplicate records may be added as different stations if the station names used differ substantially enough not to meet the *Jl* criterion), but to refine further would require substantial manual intervention and, in some cases, access to locally-held metadata unavailable to the authors.

In the second step, a data comparison was made between the candidate station and a target station for certain stations that passed the metadata threshold. These stations were mainly from sources that had not been adjusted or could not be adjusted regularly. For example, Korean or Vietnamese national sources had higher priority, but the final year of data for these sources was only 2007 or 2011. As these data were obtained by exchange between countries, they could not be updated at present. Therefore, a data comparison was performed between the same stations in the Korean source and other lower priority sources such as GHCN, CRUTEM4, or BEST. For reliable data comparison, there was a minimum overlap threshold of 60 months between the two stations (Rennie et al. 2014). If this threshold was met, data comparison was performed using the index of agreement (*IA*) (Willmott 1981). A modified version of the *IA* (Willmott et al. 1985; Legates and McCabe 1999), where the squared term was removed, was used:

$$IA = 1.0 - \frac{\sum_{i=1}^n |T_i - C_i|}{\sum_{i=1}^n |C_i - \bar{T}| + |T_i - \bar{T}|}, \tag{2}$$

where T_i and C_i are the corresponding monthly values for the target and candidate stations, respectively, and \bar{T} is the mean value for the target station. If the *IA* reaches 0.8, the candidate station is merged with the target station, with the lower-priority source being used only where the higher-priority source is unavailable.

3 Quality control

Despite quality control, the use of various methods leads to different quality problems in integrated datasets.

Similar to the process used for GHCN-V3 (Lawrimore et al. 2011), we employed a three-step quality control process; Table 2 shows the results.

Step 1: check for climate outliers. Monthly anomalies (relative to the 1961–1990), higher than five times the standard deviation of the monthly mean of the raw data at each station were considered as outliers, which accounted for 54 (0.0007%), 39 (0.0008%), and 129 (0.0026%) station months, respectively, for T_m , T_{max} , and T_{min} for all the station data. These anomalies were treated as missing data. In this study, normals get calculated with at least 10 of 30 years.

Step 2: check for spatial consistency. At a given time, Z_i is considered as an outlier and excluded if

$$|Z_i - \bar{Z}_{ij}| > 3.5\sigma_i, \text{ and} \tag{3}$$

$$|Z_i - \bar{Z}_{ij}| > 2.5, \tag{4}$$

where Z_i is the normalized air temperature anomaly in °C at the target station i , Z_{ij} is the normalized air temperature anomaly in °C at neighboring station j (not exceeding 20) within 500 km of the target station, \bar{Z}_{ij} is the mean averaged over the selected neighboring stations, and σ_i is the standard deviation of the normalized air temperature anomalies at all the selected neighboring stations at that time. The test results showed that monthly T_m , T_{max} , and T_{min} have spatial inconsistency problems for 349 (0.004%), 170 (0.003%), and 505 (0.010%) station months among all station data, respectively. These values were treated as missing data.

Step 3: check for internal consistency. Most data sources contain monthly T_m , T_{max} , and T_{min} . Despite some data sources having been quality controlled or homogenized, internal inconsistencies may arise for some data, such as a T_m being lower than the T_{min} or higher than the T_{max} . The test results showed that internal inconsistency occurred for approximately 1544 station months, accounting for 0.03% of the full database. Where the internal consistency check was failed, T_m , T_{max} and T_{min} were all treated as missing data.

Table 2 Results for each quality control step (unit: station month)

Steps	Results of QC (unit: station month)		
	T_m	T_{max}	T_{min}
First step (check for outliers)	54 (0.0007%)	39 (0.0008%)	129 (0.0026%)
Second step (spatial consistency check)	349 (0.004%)	170 (0.003%)	505 (0.010%)
Third step (internal consistency check)	1544 (0.020%)	1544 (0.030%)	1544 (0.031%)

4 Data homogenization

It is important in observational studies that the data used are homogenized, i.e., not containing artificial changes due to changes in instruments, sampling time, or station location (Jones et al. 1985; Peterson and Vose, 1998; Li and Dong 2009; Dai et al. 2011; Trewin 2013; Wang et al. 2014; Vincent et al. 2015). The following procedure was performed to ensure homogeneity in our dataset. This section describes the procedures used to homogenize our dataset.

As described in Sect. 3, a higher priority was given to national or regional data sources because we believed that individual nations or regions are most authoritative in terms of their own climate data. Then, for data sources with homogenization already applied, such as USHCN, China, Canada, and Australia, and HISTALP, or those parts of CRUTEM4 for which the data has been homogenized by NMSs/NMHSs, or considered as homogeneous series by CRU, we adopted the data directly without additional homogenization. We realize that these various sources applied different homogeneity adjustment methods. The quantile-matching adjustment method (Wang et al. 2010; Wang and Feng 2010) was applied to produce the Chinese (Xu et al. 2013) and Canadian (Vincent et al. 2012) homogenized datasets of daily temperatures, and the percentile-matching method was used to produce the Australian homogenized dataset of daily temperatures (Trewin 2013), whereas mean adjustment was applied to produce other homogenized datasets. However, we believe that these are the best homogenized datasets; and it is impossible for us to do better because we don't have the metadata, the expert knowledge about the data, and other data that were used to produce these national homogenized datasets (for example, concurrent hourly data were used to adjust for effects of the change in the definition of the climatological day before statistical methods were used to produce the Canadian homogenized dataset). We also believe that the differences induced by using different homogenization methods are trivial in comparison with the differences in observing practices, instrumentation, and post-observation processing used by different nations/countries.

As has been described (e.g., Houghton et al. 2001; Jones and Moberg 2003), adjusting the central tendency or mean state of a climate variable is usually sufficient to homogenize monthly time series and provide reliable estimates of trends and variability for temperature. Therefore, the following homogenization process and mean adjustment were applied

to the remaining station series from other data sources, such as GHCN-V3, BEST, Russia, Japan, Korea, Vietnam, and SCAR. In all, there were 4143 station series needed to be homogenized for T_{\max} and 3732 for T_{\min} . T_m was obtained from the average of T_{\max} and T_{\min} . However, at another 1682 stations where T_{\max} and T_{\min} were unavailable, T_m values were directly assessed for homogenization.

4.1 Reference construction

It is very important to derive a homogeneous reference series that well represents the same climatic variations as those in the candidate series. However, it is often difficult to find a homogenous representative reference data series. We put all the stations, including homogenized data and raw data, in the data pool for the construction of reference series. There are three steps to derive the reference series.

In the first step, a reference series should be chosen within a representative distance (Xu et al. 2013). For this purpose, we divided the world into seven continents to determine different spatial representative distances, in a similar way to Li et al. (2010a) (Table 3). The regional spatial representative distances were obtained by the following steps: (1) calculate the correlation of the first difference series between any two stations within 1000 km of each other; and (2) then fit a curve through it to determine the distance at which the correlation goes below 0.8. From Table 3, the representative distances of the regions were all within 450 km; that is, reference stations within this distance should be chosen. The exception was found in Antarctica; although the representative distance (derived from a small sample) of the region is 220 km, there were very few stations within this distance, so it was impossible to homogenize most Antarctic stations.

In the second step, the method of Peterson and Easterling (1994) was used to establish the reference time series (R0). Peterson and Easterling (1994) showed that the reference series can be constructed by spatially averaging the neighboring first difference series, weighted by the inverse square of distance, from at least three nearby stations that are highly positively correlated with the candidate series. The averaged difference series was analyzed to remove any abnormal data points, after which it was converted back to data series for use as the reference series. The advantage of this approach is that it reduces the impact from individual jumps on the re-converted reference series, which is affected by inconsistent

Table 3 Average distance between any two stations within 1000 km of each other for which the correlation of available annual temperature data during 1900–2014 reached 0.8 across the seven regions

Region	North America	Europe	Asia	South America	Africa	Australia	Antarctica
Average distance	389	342	392	421	429	396	220

lengths of nearby series and some anomalous values or undetected inhomogeneities in individual sequences.

In the third step, the inhomogeneity and representativeness of R0 is tested to ensure its suitability. The R0 series is first tested by visual checks or by using a penalized maximal F (PMF) test (Wang 2008a, b). If R0 is obviously inhomogeneous, it is reconstructed by adjusting the potential reference stations to make a second homogeneous reference series R1. Then, the correlation between the first difference series of R1 and the target series is calculated. If this correlation reaches 0.8, the R1 series is used as the reference series. In our analysis, there were about 20 and 25 stations (0.48 and 0.67%) of T_{\max} and T_{\min} , respectively, which were excluded because of no suitable reference series available to allow the construction of an R0 or R1 series meeting the correlation criterion.

4.2 Methodology for discontinuity detection and adjustment

There are many studies on benchmarking of the data homogenization methods; many methods are found to have exhibited good performance in different aspects (Kuglitsch et al. 2012; Venema et al. 2012; Domonkos et al. 2013) with rankings of specific methods often dependent on the metric(s) used for assessment. In this study, we used the RHtestsV4 software package (Wang and Feng 2010) to homogenize the monthly temperature series. This package includes two algorithms for detecting unknown change-points: the PMTred algorithm (Wang 2008a), which is based on the penalized maximal t test (Wang et al. 2007) and requires a reference series; and the PMFred algorithm (Wang 2008a), which is based on the penalized maximal F (PMF) test (Wang 2008b) and can be used without a reference series. The RHtestsV4 package and its previous versions have been widely used to homogenize climate data (e.g., Zhang et al. 2005; Alexander et al. 2006; Wan et al. 2010; Dai et al. 2011; You et al. 2011; Kuglitsch et al. 2012; Vincent et al. 2012; Wang et al. 2013; Xu et al. 2013).

The potential change points from the RHtestsV4 were then synthesized and identified using available metadata and three other criteria. The three criteria were the timescale consistency, spatial consistency, and elements consistency. For the timescale consistency, the monthly scale breakpoints were compared with those on an annual scale. For the spatial consistency, the climate trends in the base series were compared with nearby stations to determine the breakpoints in the candidate series. For the elements consistency, the adjusted values as well as the sensitivity of the three (T_m , T_{\min} , T_{\max}) time series to artificial changes were used as the criteria to determine the breakpoints. In general, metadata was the most direct and solid evidence;

that is, if a change point is supported by metadata, it will be retained for adjustment. We consider a detected change point to be documented when metadata indicate a documented change within 1 year before or after the detected change point. Unfortunately, at present, detailed metadata (for those stations not derived from national datasets) is unavailable outside the mainland of China. For other criteria, change points supported by at least two pieces of evidence will be retained for adjustment.

4.3 Statistics of the adjustment

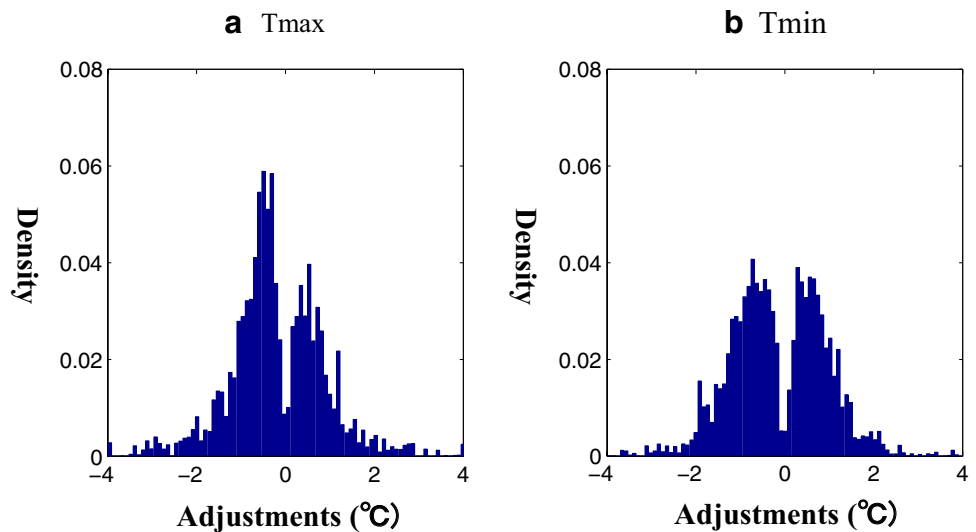
As described above, our homogenization was not applied to the data sources that have already been homogenized. In addition, stations with a length shorter than 20 years were included in the station dataset after being compared/tested either visually or by statistics. In total, there were 4143 stations that remain to be homogenized for T_{\max} and 3732 for T_{\min} . Table 4 lists the number of stations without any shift or with shifts that have been identified and adjusted for T_{\max} and T_{\min} . At the 5% significance level, the T_{\max} and T_{\min} series were found to be homogeneous at 2935 and 2456 stations (71 and 66%), respectively. A total of 1447 change points were detected in 1208 T_{\max} time series and 1750 change points were detected in 1276 T_{\min} time series; for these stations, T_m was calculated by averaging the homogenized T_{\max} and T_{\min} . In addition, other directly collected raw T_m series for 1682 stations were also assessed for homogeneity; they were found to be homogeneous at 1092 (65%) stations, and a total of 736 change points were detected in 590 T_m time series.

Figure 1 shows the probability density of all mean-adjustments applied to the monthly T_{\max} and T_{\min} identified as having inhomogeneities. The vast majority of the adjustments are between -1 and 1°C for both T_{\max} and T_{\min} , with more negative than positive adjustments for T_{\max} . The extreme adjustments mostly range from -3 to 3°C for T_{\max} and from -2 to 2°C for T_{\min} . The mean of the adjustments is -0.2322°C for T_{\max} and -0.1386°C for T_{\min} . The bimodal distribution of the results is similar

Table 4 Statistics of breaks for the T_{\max} , T_{\min} , and T_m temperature series during 1900–2014

Breaks	T_{\max}	T_{\min}	T_m
None	2935	2456	1092
One	1002	982	462
Two	184	242	113
Three	14	33	12
Four or more	8	19	3
Total breaks	1447	1750	736
Total stations	4143	3732	1682

Fig. 1 Probability density of all mean-adjustments applied to the T_{\max} and T_{\min} temperature series



to those discussed in Brohan et al. (2006) and Lawrimore et al. (2011).

Data inhomogeneities also decrease the spatial consistency of estimated trends in annual mean temperatures; this could be improved with homogenization (Li et al. 2004a; Dai et al. 2011; Vincent et al. 2012; Trewin 2013; Xu et al. 2013). Figure 2 shows the distribution of the original and adjusted T_{\max} and T_{\min} trends for those stations that have been detected as inhomogeneous, with the trends being calculated over the available record during 1900–2014. The temperature trends derived from the homogenized data series have improved spatial consistency compared to those derived from the original data series. The improvement is particularly noticeable in regions with large warming or cooling trends in North America, South America, Africa, Europe, and Asia for both T_{\max} (upper panel) and T_{\min} (lower panel). It is worth noting that we have not made any adjustments to the Antarctic data, so there is no change for this area due to a lack of available neighbors for using the approach discussed in Sect. 5.1.

4.4 Urbanization

It is well known (e.g., Oke 2004) that urban sites are generally warmer than their rural surroundings. The nature of the urban–rural temperature difference has been widely documented in the literature, with much of the focus on the largest differences between individual sites at sub-daily timescales.

For studies of the potential urban influence on global-scale temperature trends, our interest is not in extreme differences, or in assessments of the urban heat island (UHI) for individual locations, but rather in the way that changes in urban influences over time may affect long-term trends in large area averages based on many station

records (Peterson 2003; Parker 2004, 2006). Peterson and Owens (2005) also note that, in the context of a large data set, any UHI influence operates in conjunction with other influences on local-scale climate observations, including elevation differences, distances from major water bodies, observation time changes and weather types.

Existing global data sets take a range of approaches to urbanization influences. The GISTEMP data set currently includes an urbanization adjustment, based on a comparison of trends at urban stations with those at nearby rural stations (Hansen et al. 2010). Other global data sets do not include an explicit urbanization adjustment, in many cases, urban influences on temperatures at a station will be manifested as a step change (e.g. when a new building is constructed near an observing site) and will be adjusted for as part of the general homogenization process.

At the global scale, urbanization impacts on estimates of global mean land surface temperatures have been found to be negligible. The urbanization adjustments applied to GISTEMP only influence global mean LSAT by a maximum of 0.01 °C (Hansen et al. 2010), whilst Berkeley Earth (Wickham et al. 2013) also found a minimal influence, reinforcing similar conclusions in AR4 of IPCC (IPCC 2007).

Urbanization impacts on temperature can be regionally important, especially in areas such as eastern China where rapid urban growth is occurring. This has been assessed by a number of authors (e.g., Li et al. 2004a, b, 2010b; Zhou et al. 2004; Jones et al. 2008; Yan et al. 2009; Ren et al. 2005; Wang et al. 2015) with varying conclusions, but all find that urbanization has partially contributed to observed warming over China, with estimates ranging from 5 to 40% of the total observed warming. Hansen et al. (2010) also found an enhanced urbanization influence in the southwest United States, another area which has experienced rapid

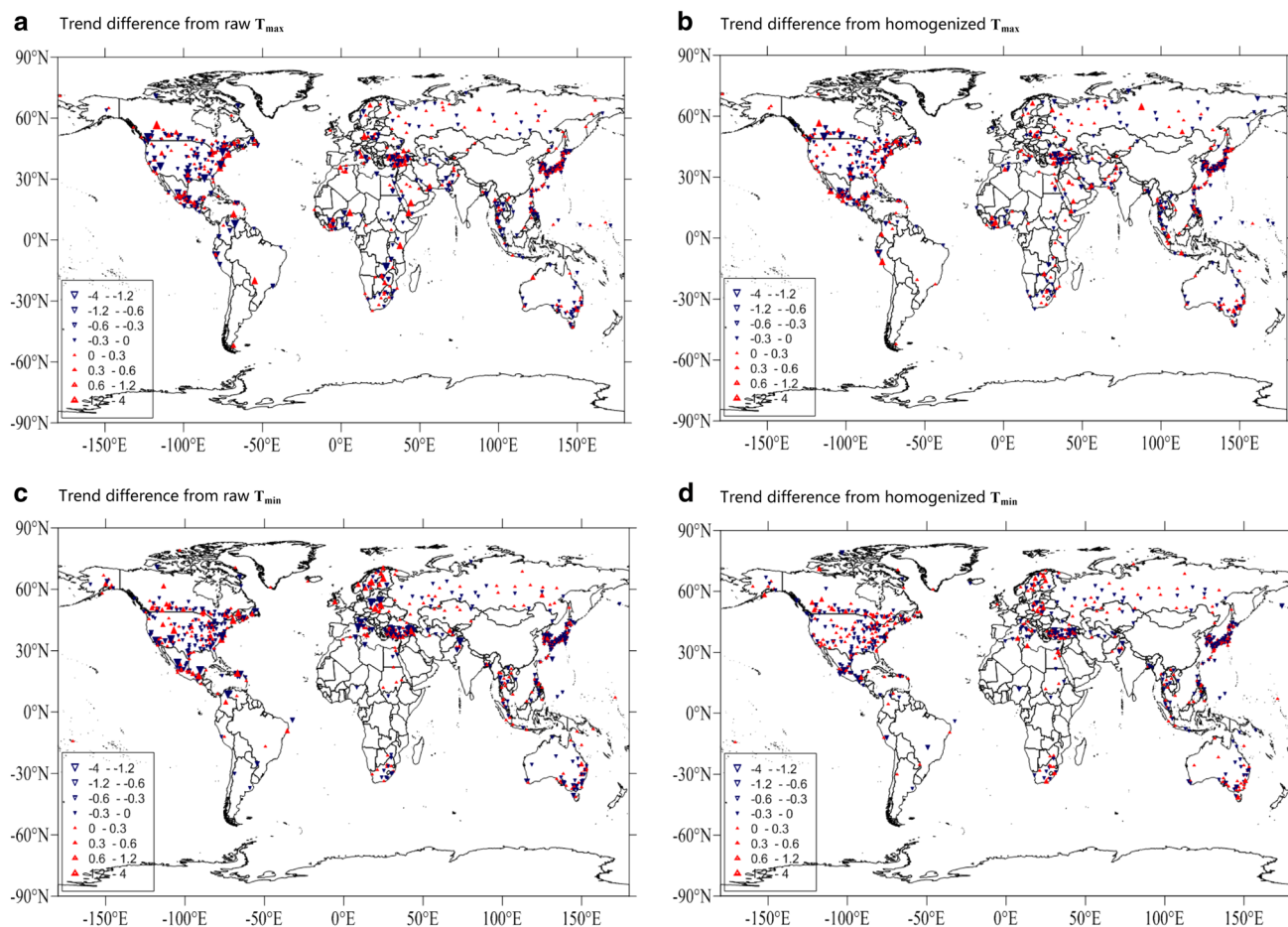


Fig. 2 Trend difference between each station and its nearest neighbor in time series of annual means of the raw and homogenized monthly means of daily maximum (T_{\max}) and of daily minimum (T_{\min}) temperatures (homogeneous stations are not shown; trends are

calculated over whatever length of record is available). The *triangle size* is proportional to the magnitude of trend difference. Units: $^{\circ}\text{C}/\text{decade}$

urbanization in recent decades, with an influence of up to 0.1°C on area averages.

Considering the evidence that urbanization has a minimal effect on LSAT at global and continental scales, the CMA-LSAT data set does not include any explicit urbanization adjustments. However, regional analyses based on the data set should take possible urbanization influences at the local scale into account in heavily urbanized areas.

5 Data assessment and discussion

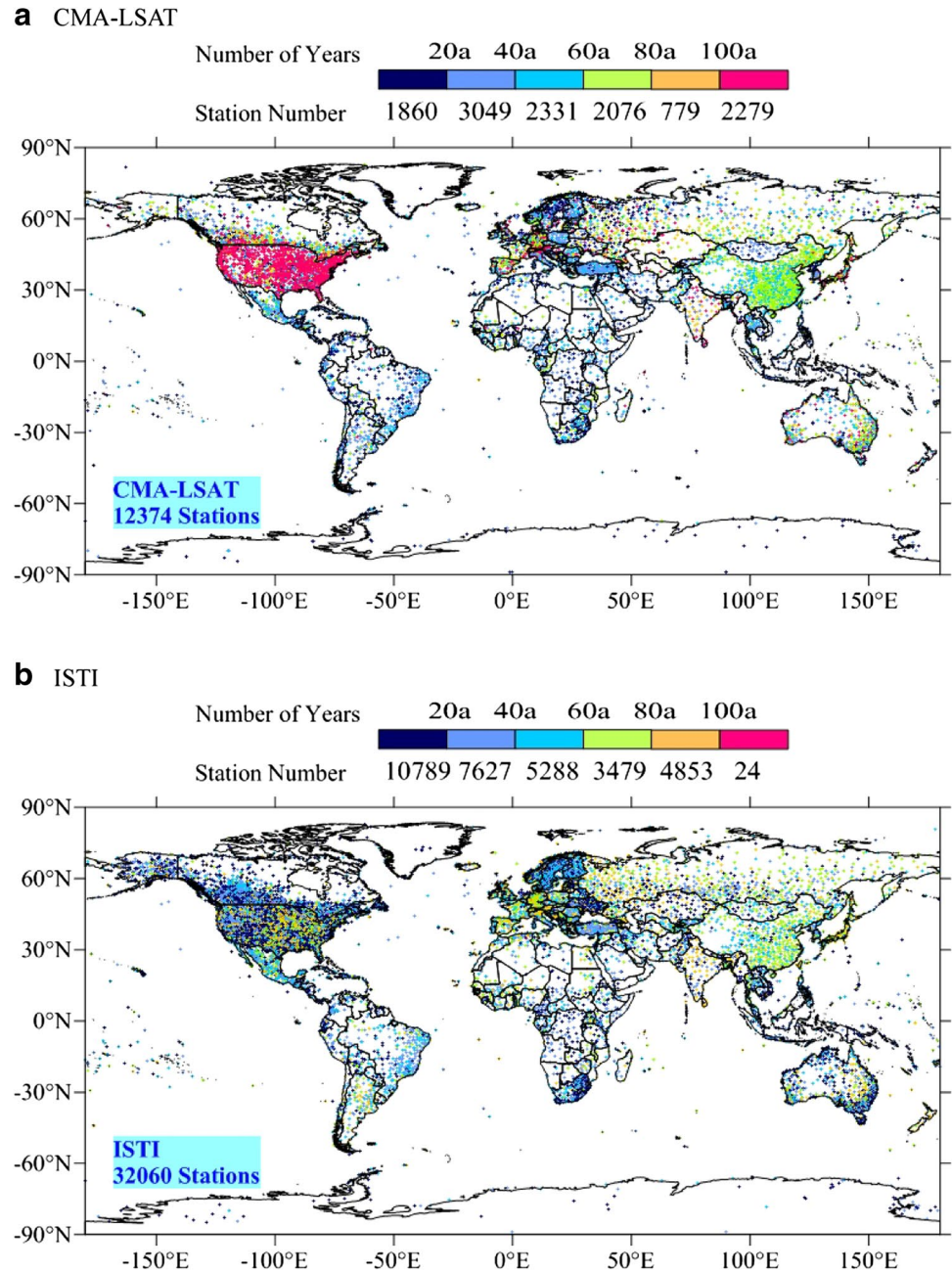
5.1 Station number and spatial–temporal distribution

Using the source hierarchy (Table 1) and data integration (Sect. 2.2), approximately 12,374 unique stations with monthly T_m series, 8273 stations with T_{\max} series, and 7655 stations with T_{\min} series were identified. Figure 3 shows the spatial distribution of the stations with monthly T_m

data during 1900–2014 and their record length from CMA-LSAT and other datasets. Although the station density in CMA-LSAT is still lower than ISTI (Fig. 3a, b), it is higher than that in GHCN-V3 and CRUTEM4 (Fig. 3c, d), particularly over Europe and Asia, where the length of data series increases the most in CMA-LSAT (Fig. 3e, f).

A comparison of station counts is given in Fig. 4. Since 1900, there are consistently more stations in CMA-LSAT than in GHCN-V3 or CRUTEM4. Moreover, there is a significant drop in the number of stations in 1990, but this is ameliorated by many of the new sources, illustrating that this station drop-off is simply due to access, not due to widespread station closures. (The ISTI dataset also shows substantial improvement over GHCN-V3 and CRUTEM4 in recent years). It is worth noting that for Europe, CMA-LSAT has fewer long station series than CRUTEM4, because we use ECA&D and GHCN-V3 as the main sources and the longer series in CRUTEM4 were not given preference due to their not having T_{\max} and T_{\min} data. This

Fig. 3 Spatial distribution of the stations with monthly mean temperature T_m data, and their length from the CMA-LSAT and other datasets during 1900–2014 (unit: year)



could become an important issue if CMA-LSAT is to be extended before 1900.

A comparison of grid boxes is given in Fig. 5. Coverage is defined by one or more stations within each $5^\circ \times 5^\circ$ grid box. There is an increase in the global station coverage for all time periods comparing with CRUTEM4 and GHCN-V3, especially for 1990–2010s (the coverage of CMA-LSAT is only slightly lower than that of ISTI, which shows both of them have comparable data resources for near real-time updating). Compared with GHCN-V3, CMA-LSAT's coverage in the Northern Hemisphere (NH) is higher by about 10–15% from 1900 to 1950, 20% during from the

1960s to the 1980s, and 30–40% during the 1990s and 2000s (Fig. 5b). The improvement for The Southern Hemisphere (SH) coverage varies, for the most part, between a 15 and 25% increase, with larger increases of approximately 50% over the past 20 years (Fig. 5c).

5.2 Characteristics at the continental scale

As described in Sect. 2.1 we spent considerable efforts to obtain a higher density of stations over Asia, for example, by exchanging data with Korea, Vietnam, and Japan. As shown by Fig. 6a, the number of Asian stations with

Fig. 3 (continued)

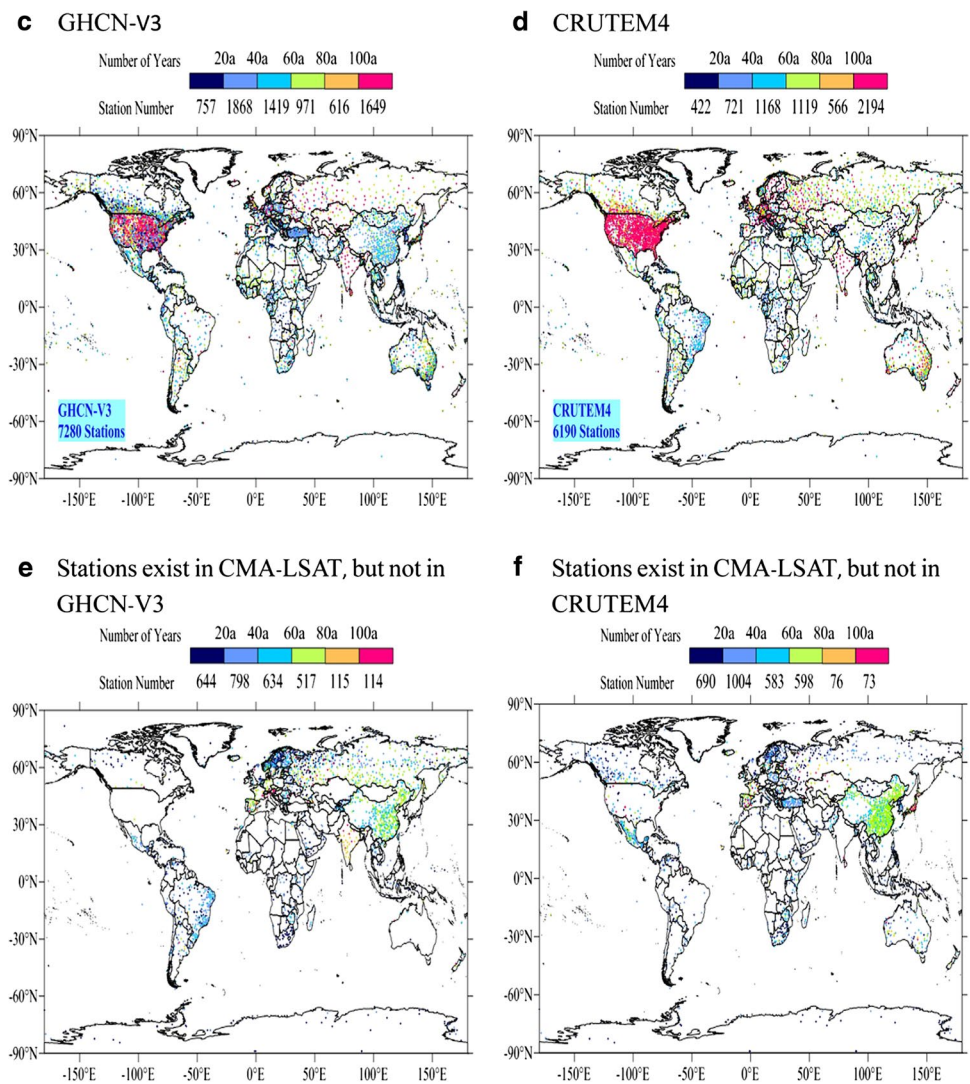
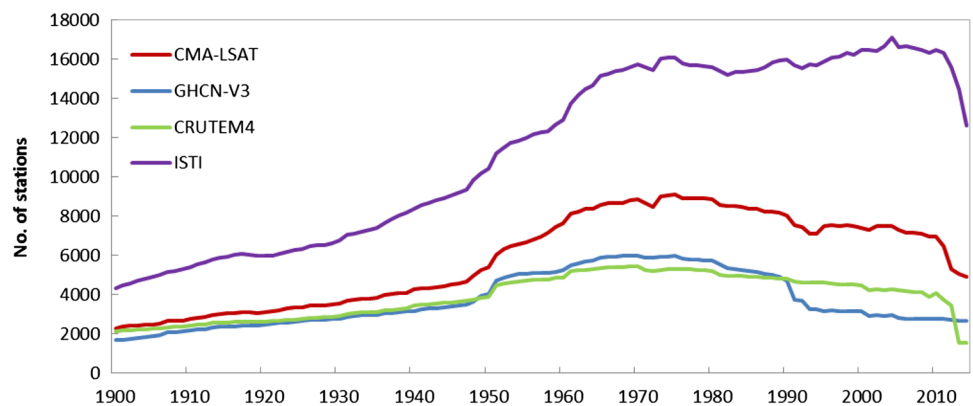


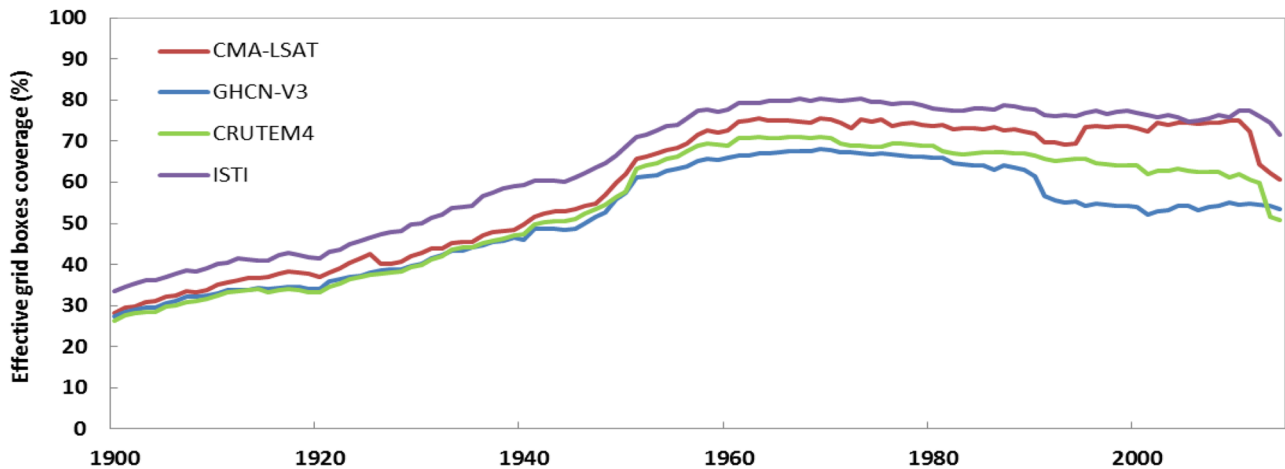
Fig. 4 Station count comparison from the CMA-LSAT and other datasets



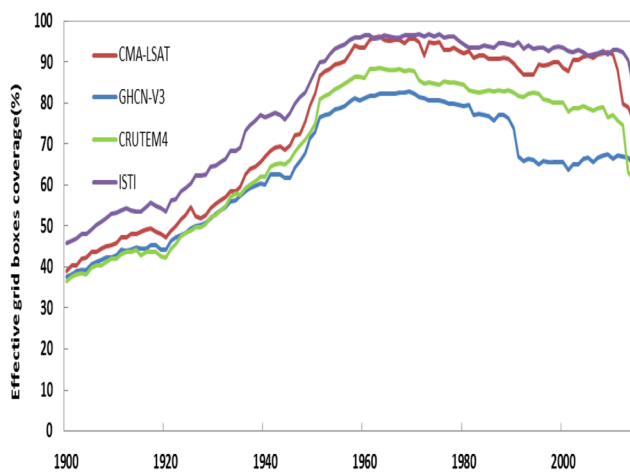
different record lengths in CMA-LSAT is significantly higher than that in other datasets: 202 stations in CMA-LSAT have record lengths of 120–150 years, compared with 93, 57, and 117 stations in GHCN-V3, CRUTEM3, and CRUTEM4, respectively. Figure 6b shows that the

number of Asian stations in CMA-LSAT is significantly higher than those in GHCN-V3 and CRUTEM4 for the entire analysis period from 1900 to 2014, but less than those for short (less than 20 years) and long (longer than 80 years) series in ISTI. The number of Asian stations after

a Global



b Northern Hemisphere



c Southern Hemisphere

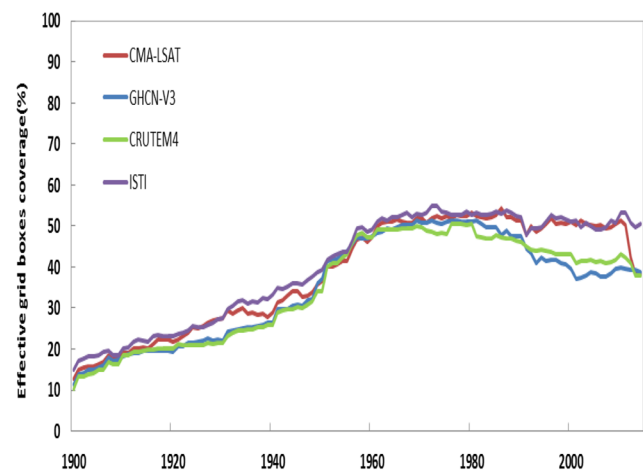


Fig. 5 Grid box numbers for the CMA-LSAT compared to GHCN-V3 CRUTEM4 and ISTI: (*upper*) global, (*left*) Northern Hemisphere, and (*right*) Southern Hemisphere

the 1990s is significantly higher in CMA-LSAT than in the other two datasets. Similar statistical results are found in both Africa and South America, where the station densities are relatively low (Fig. 6c–f). Not surprisingly, the station numbers in CMA-LSAT are fewer than those in ISTI in Africa and South America. From this point of view, ISTI shows strong potential as a new data source for future upgrades of CMA-LSAT.

The area averages are compared with two other analyses. For this purpose, monthly anomalies were calculated relative to the reference period of 1961–1990, and only those stations with annual mean values available for at least 10 years during 1961–1990 were used. Comparing with CRUTEM4 use of stations with lengths of at least 15 years data, stations with lengths of at least 10 years were chosen in CMA-LSAT to expand the coverage of normal and

use more stations. This resulted in a total number of 9765 stations (8300 in the NH and 1465 in the SH). Following Jones (1994), gridding of the temperature anomalies was made by averaging all station anomaly values within each 5° latitude \times 5° longitude grid. The use of a base period in this type of study appears to eliminate a large number of series, but almost all the series that cannot produce the 1961–1990 average are mostly short and recent series (the average length of these short stations is 18 years, comparing with 61 years when those longer ones were used). Regional mean LSAT anomaly time series were constructed based on the method of Jones (1994) and Jones et al. (1999) by averaging with area-weights, using the cosine of the central latitude of each grid box as the weight coefficient. Eight regions are defined, following Jones and Moberg (2003), for the seven continents of the world (Asia, Africa, South

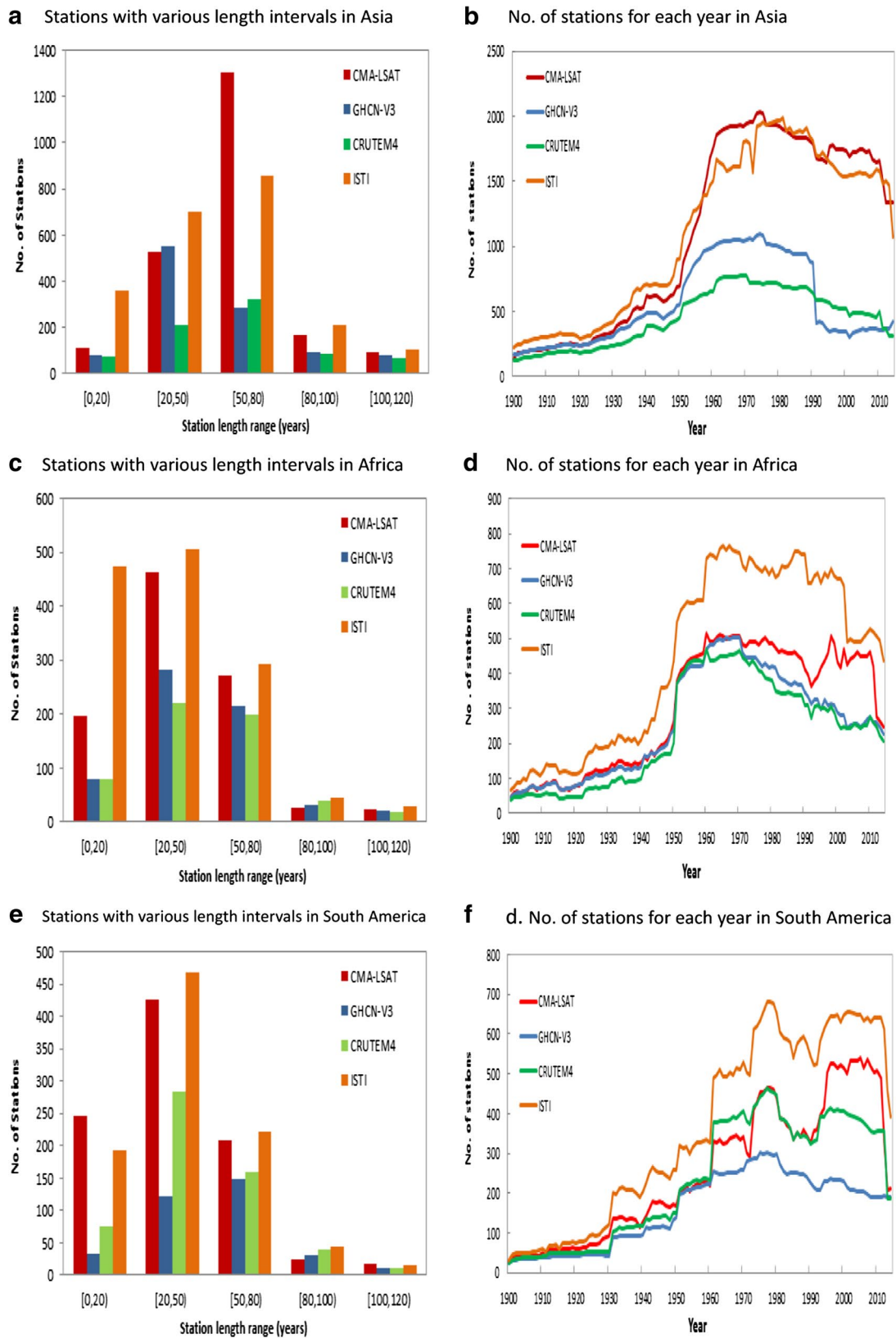


Fig. 6 Variations in station numbers at various length intervals in different regions (Asia, Africa and South America) from the CMA-LSAT and other datasets

America, Europe, North America, Australia, and Antarctic) plus the Arctic. Here, the regional series of the eight continents/regions are first chosen for comparison (Fig. 7a–h). The following global or hemispheric mean LSAT anomaly time series are established using the same method.

Figure 7a shows the Asian SAT anomaly time series during 1900–2014. Compared with the other two analyses, the CMA-LSAT Asia series is slightly cooler during the 1920–1930s and slightly warmer during recent decades. However, in general, the SAT change trends are very similar (0.120, 0.114, and 0.118 °C/decade for CMA-LAST, CRUTEM4, and GHCN, respectively) (Table 5), except for constant, very slight underestimates in recent years for CRUTEM4 and GHCN.

Figure 7b–h show the SAT anomaly time series for the other seven regions during 1900–2014. For Africa (Fig. 8b; Table 5), the greatest differences between the analyses occur at the beginning of the records. The CMA-LSAT series is slightly warmer during 1900–1950 (about 0.2 °C warmer during 1900–1920s), but is slightly cooler since the 1990s. For South America (Fig. 7c; Table 5), there is a greater divergence between the series, except in the climate normal period (1961–1990). The CMA-LSAT series is more consistent with CRUTEM4 than GHCN-V3, which is mainly due to the sharp drop in the number stations after 1990 in GHCN-V3. For Europe (Fig. 7d; Table 5), CMA-LSAT used the same homogeneity dataset as the main data source as CRUTEM4, so CMA-LSAT is more consistent with CRUTEM4. The notable differences from the GHCN-V3 arise from the fact that GHCN-V3 has fewer stations in Europe than HISTALP and CRUTEM4. For Australia (Fig. 7e; Table 5), there are slight differences between the other two datasets, which is likely due to the homogenization of the data by different groups [e.g., CMA-LSAT uses homogenized ACORN data (Trewin 2013) from the Australia Bureau of Meteorology, as does CRUTEM4]. For North America (Fig. 7f; Table 5), CMA-LSAT is consistent with the other two datasets (except at the beginning and end of the records) because all three datasets use the USHCN homogenized dataset as the main data source and CMA-LSAT uses the newly developed, second generation of homogenized data for Canada (Vincent et al. 2012). It is worth pointing out that these high latitude areas tend to show a faster temperature increase, however, the meteorological observations in high latitude areas tend to be shorter, and the shorter-term trends for the SAT in higher latitude tend to be larger. If these short-term data series have been added to the whole dataset directly, this could cause some of the warming bias in global/hemisphere SAT trends. Thus CMA-LSAT does not contain that many more stations in high latitude regions like Arctic and Antarctic. For the Arctic (Fig. 7g), the CMA-LSAT series is more consistent with CRUTEM4 than GHCN-V3 during most

of 1900–2014. For the Antarctic (Fig. 7h; Table 6), CMA-LSAT is consistent with CRUTEM4 after 1960. There are many fewer stations in the Antarctic, and therefore larger variances, before 1960; thus, as for GHCN, we only retain the series after the 1940s. CRUTEM4 before the mid-1940s is based upon a single station (Orcadas).

In summary, although CMA-LSAT shows some differences for the continental/regional SAT series compared with GHCN and CRUTEM, the three sources still reflect good consistency. With detailed metadata and a full understanding of regional climate changes, the direct use of homogenized data produced by domestic meteorological data centers would be most likely to improve the accuracy of regional/national climate change detection.

6 Comparisons of large-scale surface air temperature changes

6.1 Annual

Annual mean LSAT anomaly time series for both hemispheres and the entire globe during 1900–2014 are shown in Fig. 8. The temperature anomaly curves show extremely high similarity with those reported in previous studies (Hartmann et al. 2013). The linear trends in annual mean LSAT for the NH, the SH and the whole globe are 0.107, 0.083, and 0.102 °C/decade, respectively; all are statistically significant at the 1% confidence level (Table 6). Much of the hemispheric and global warming occurred in two distinct periods: from the 1910s to the late 1930s and from the early 1980s to the mid-2000s. The relatively cool periods or stable periods appeared in the 1900s, the 1940–1970s, and over the last 10 years (2005–2014). These results are very similar to those found in previous studies (e.g., Hansen et al. 2006; Smith et al. 2008; Jones et al. 2012; Jones 2016).

The annual warming is larger in the NH (0.107 °C/decade) than in the SH (0.083 °C/decade). However, the SH lands exhibit a slight warming from the early 1950s to the early 1970s, in contrast to the NH lands, which witness a slight cooling. The hemispheric warming that began in the early 1980s is much more remarkable in the NH than in the SH. It is also clear that the global mean LSAT change depends to a larger extent on that from the NH, owing to there being more land (and hence grid boxes in our LSAT dataset) in the NH than in the SH, as well as a higher proportion of land grid boxes with available data.

From 1979 to 2014, the mean LSAT anomalies in the NH, the SH, and the entire globe experienced unprecedented and highly significant annual warming trends, reaching 0.305, 0.142, and 0.272 °C/decade, respectively.

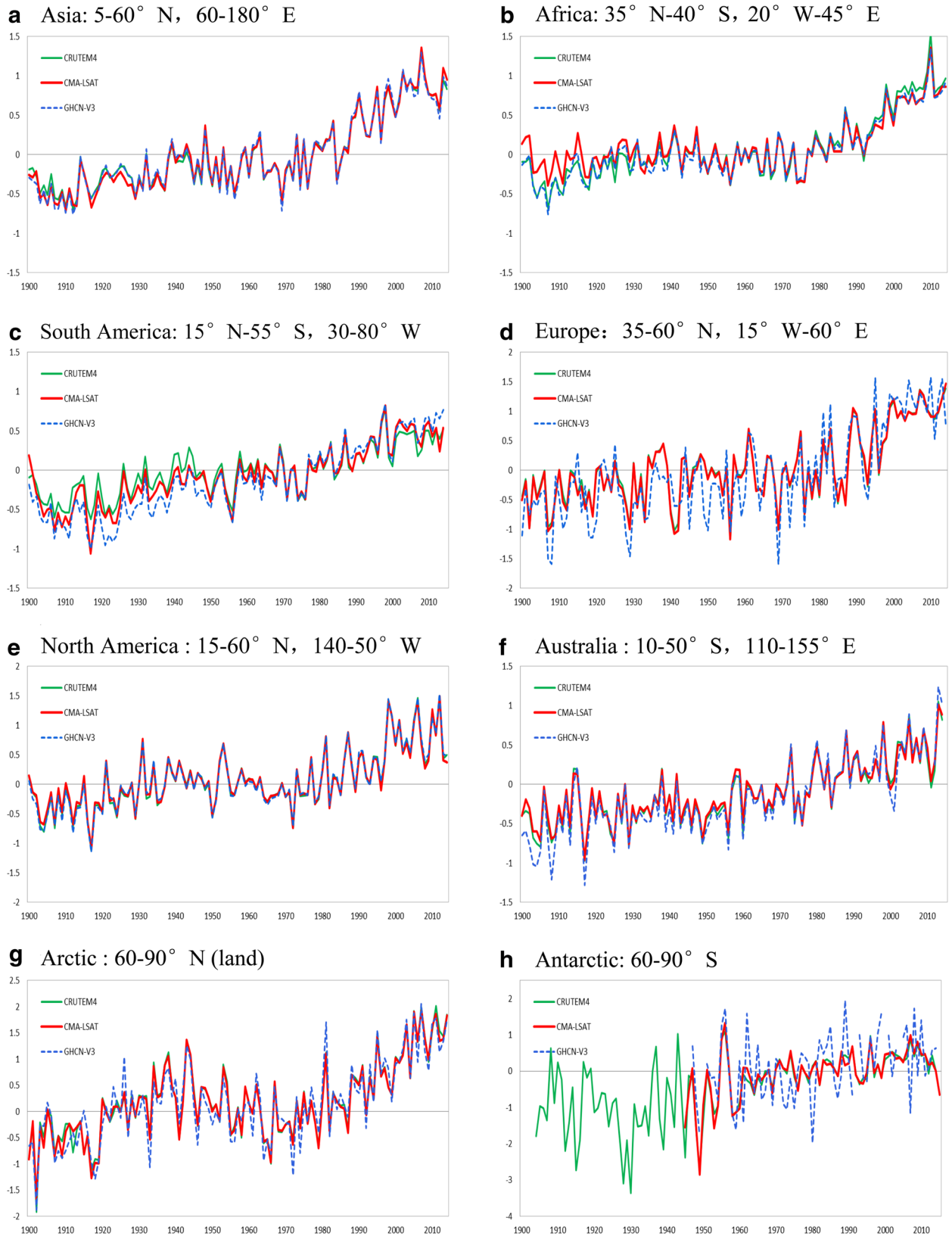


Fig. 7 Annual land surface air temperature anomalies for different regions (Asia, Africa, South America, Europe, North America, Australia, Arctic, and Antarctic (a–h), respectively) during 1900–2014 (relative to the 1961–1990 mean)

Table 5 Linear trends of annual mean LSAT for different regions during different periods. All trends are significant at the 5% level

	1900–2014						1951–2014	
	Asia	Africa	South America	Europe	North America	Australia	Arctic	Antarctic
CMA-LAST	0.1203	0.0681	0.0952	0.1284	0.0814	0.0835	0.1318	0.1281
CRUTEM4	0.1143	0.0992	0.0740	0.1221	0.0874	0.0874	0.1309	0.1404
GHCN-V3	0.1183	0.0913	0.1241	0.1577	0.0944	0.1054	0.1376	0.1500

Unit: °C/decade

Table 6 Linear trends and their 95% confidence ranges of annual mean LSAT for the Southern Hemisphere, Northern Hemisphere, and entire globe during three different periods

	1900–2014			1979–2014			1998–2014		
	NH	SH	Glo	NH	SH	Glo	NH	SH	Glo
DJF	0.113 ± 0.011 ^a	0.074 ± 0.006 ^a	0.104 ± 0.009 ^a	0.281 ± 0.052 ^a	0.126 ± 0.024 ^a	0.244 ± 0.042 ^a	-0.140 ± 0.147	0.147 ± 0.088	-0.107 ± 0.123
MAM	0.124 ± 0.009 ^a	0.080 ± 0.007 ^a	0.112 ± 0.007 ^a	0.366 ± 0.037 ^a	0.079 ± 0.037 ^a	0.284 ± 0.031 ^a	0.149 ± 0.095 ^a	-0.005 ± 0.143	0.147 ± 0.087 ^a
JJA	0.091 ± 0.007 ^a	0.083 ± 0.005 ^a	0.090 ± 0.006 ^a	0.300 ± 0.031 ^a	0.152 ± 0.026 ^a	0.267 ± 0.026 ^a	0.225 ± 0.086	0.039 ± 0.095	0.178 ± 0.075
SON	0.095 ± 0.009 ^a	0.087 ± 0.006 ^a	0.092 ± 0.007 ^a	0.314 ± 0.042 ^a	0.193 ± 0.024 ^a	0.287 ± 0.035 ^a	0.142 ± 0.106	0.252 ± 0.068 ^a	0.183 ± 0.082
ANN	0.107 ± 0.007 ^a	0.083 ± 0.005 ^a	0.102 ± 0.006 ^a	0.305 ± 0.030 ^a	0.142 ± 0.021 ^a	0.272 ± 0.025 ^a	0.150 ± 0.056	0.120 ± 0.081	0.124 ± 0.057

Unit: °C/decade

^aIndicate trends that are significant at the 5% level

Table 6 also gives the linear trends of annual mean LSAT for the periods 1998–2014, since a number of studies have looked at this period because of the debate about “hiatus”. Slower warming trends are observed in the 1998–2014 period for the NH, SH and globe (0.150, 0.120 and 0.124 °C/decade, respectively). These warming trends, although still stronger than those observed over the full 1900–2014 period, had been interpreted by many as a “hiatus” (e.g., Slingo et al. 2013), although some more recent analyses (e.g. Karl et al. 2015; Lewandowsky et al. 2015; Mann et al. 2016) question the existence of such a “hiatus” in any significant sense, or note that any slowdown was driven primarily by the oceans (Dai et al. 2015) and was less evident in LSAT. More pronounced “slowdowns”, or even local cooling, in the 1998–2014 period are evident in specific regions, such as parts of North America, central and eastern Asia, and northern Australia (Fig. 9c), with China showing a particularly pronounced slowdown (Li et al. 2015; Duan and Xiao 2015; Zhai et al. 2016).

The warmest years in the CMA-LSAT temperature record are concentrated in the later part of the data set, with 15 of the 16 warmest years in the global LSAT record being the 15 years from 2000 to 2014. Furthermore, the World Meteorological Organization has reported that 2015 was warmer than any year of the pre-2015 period, and that 2016 is very likely to be warmer still. 2015 was also reported to be the warmest year on record in China’s national data set (Zhai et al. 2016).

The spatial distributions of annual mean LSAT trends for the periods 1900–2014, 1978–2014, and 1998–2014

are shown in Fig. 9. There is spatially coherent warming at the global scale during 1900–2014, although the warming rates in most regions are below 0.2 °C/decade (Fig. 9a). During 1979–2014, however, the global land surface warming trends are clearly higher than those of the entire time period, with particularly large trends occurring at high latitudes of the NH (Fig. 9b). During the recent period of slower warming (1998–2014), a strong incoherence in the global LSAT changes can be seen, with abnormal warming in Arctic areas neighboring the Eurasian Continent and the North Atlantic Ocean, and substantial cooling in North America, eastern and central Asia, northern Australia, and southern Africa (Fig. 9c), which is quite similar with the previous studies based on both the other observational and satellite datasets (Cowtan and Way 2014). However, trends over such a short period have a large uncertainty associated with them, especially in regions with large interannual temperature variability such as continental interiors at mid- to high latitudes in Asia and North America, and hence these geographic patterns should be interpreted with caution.

6.2 Seasonal

Seasonal mean LSAT anomaly time series for the globe during 1900–2014 are shown in Fig. 10. For the first distinct global warming period, from the 1910s to the late 1930s, JJA and MAM express warming characteristics; whereas DJF and SON appear relatively stable (DJF shows a little cooling). All seasons express distinct warming characteristics across the globe from the early 1980s to the

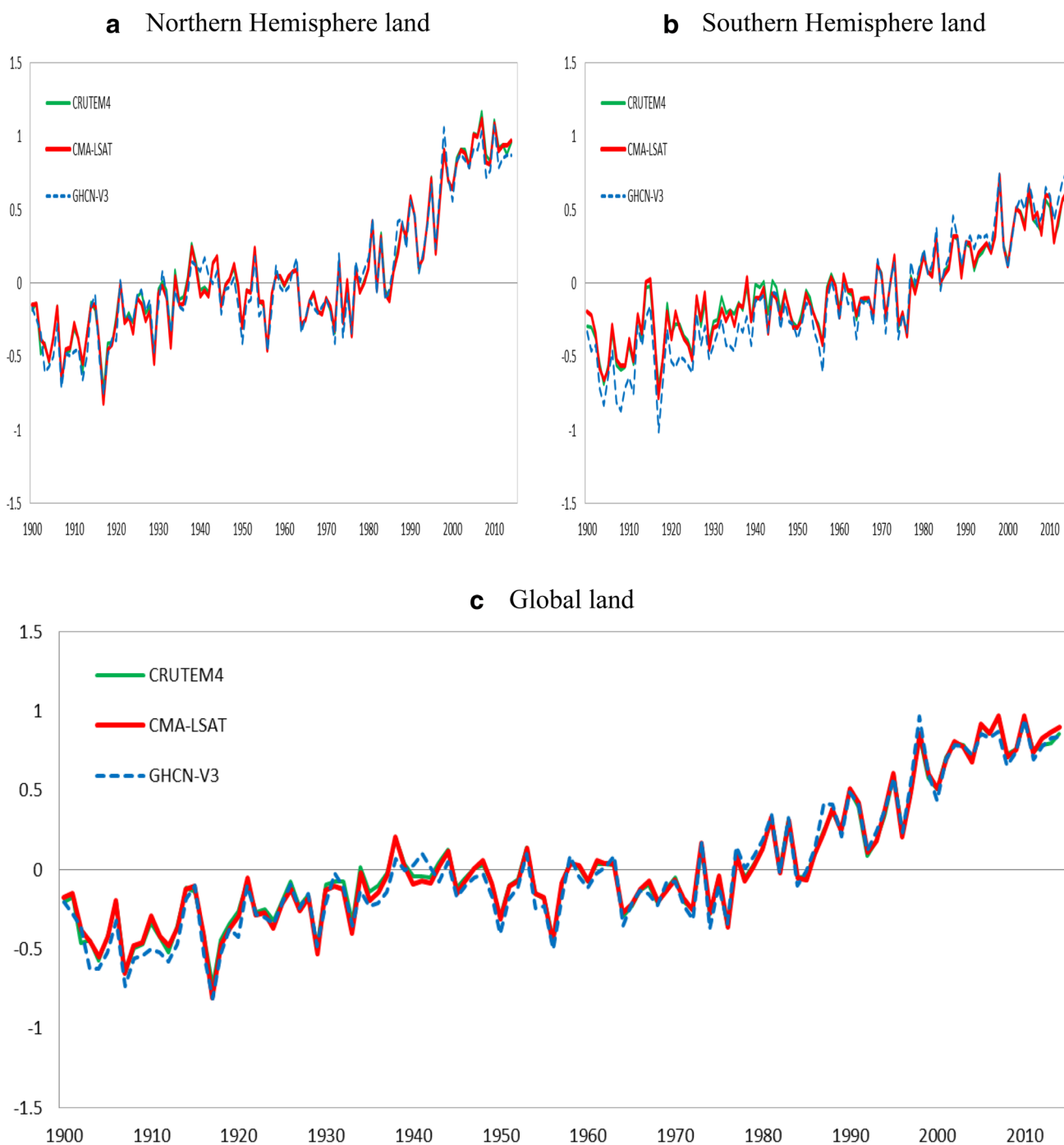


Fig. 8 Annual mean LSAT anomalies ($^{\circ}\text{C}$) during 1900–2014 for the Northern Hemisphere (a), Southern Hemisphere (b), and entire globe (c) (compared with GHCN-V3 and CRUTEM4)

mid-2000s. All seasonal series show weak trends during the 1940–1970s.

For the NH (Fig. 11), year-to-year variability is greatest during DJF and lowest in JJA. All seasonal series show comparable century-scale warming from the beginning of the twentieth century, but there are differences between them in terms of timing. Warming is significant in all

seasons during 1900–2014; it is greatest during MAM and lowest in JJA. For the SH (Fig. 12), year-to-year variability shows greater similarity between the seasons, as the SH land areas are more influenced by the oceans than for the NH. Warming is greatest during JJA and lowest in DJF.

Table 6 also gives the linear trends of seasonal mean LSAT for both hemispheres and the entire globe for the

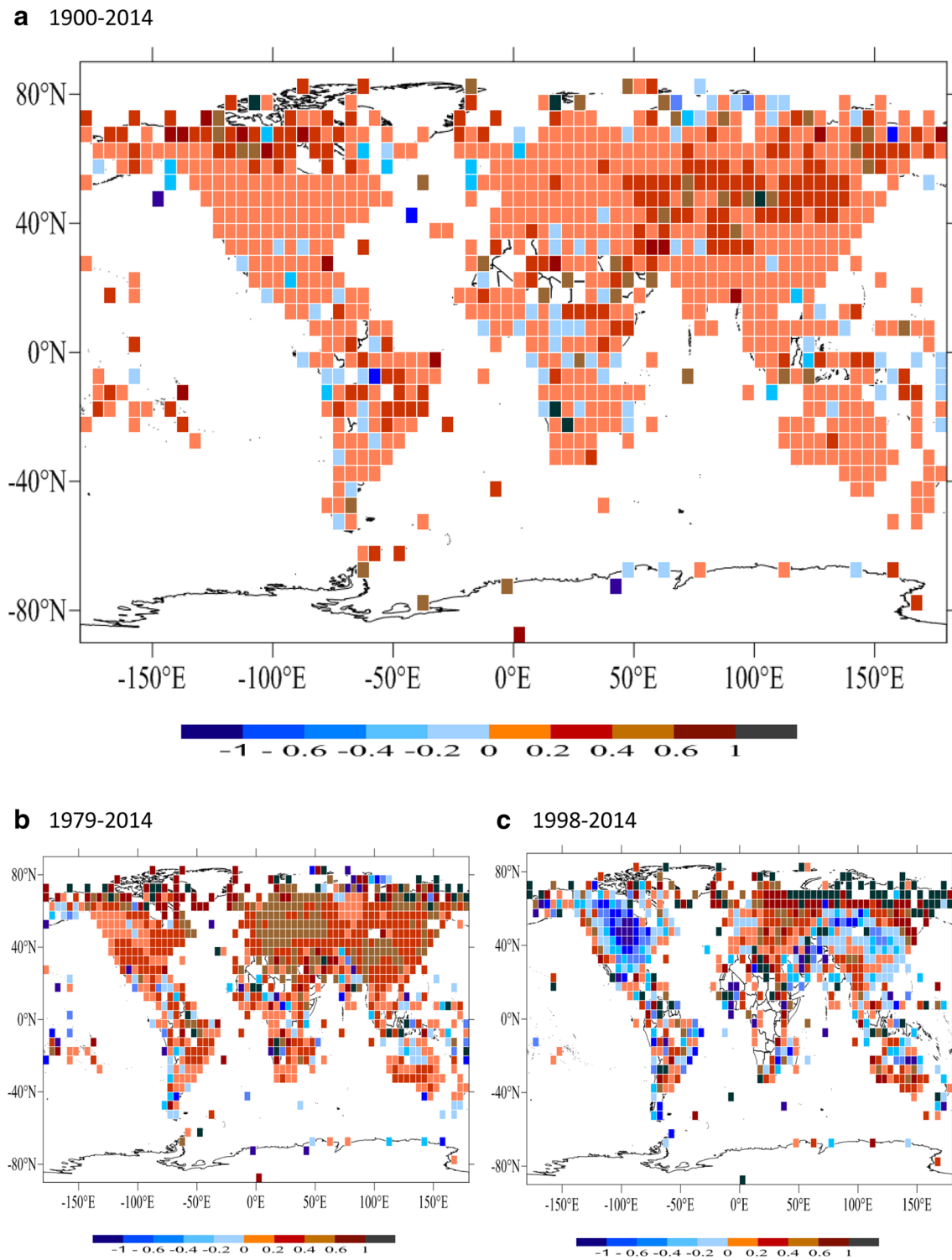


Fig. 9 Trends in global land-surface air temperature from CMA-LSAT over three different periods (white area indicates missing data): 1900–2014 (a), 1979–2014 (b), and 1998–2014 (c). Units: °C/decade

periods 1979–2014 and 1998–2014. From 1979 to 2014, all seasonal mean LSAT trends for the SH are weaker than those for the corresponding season/period in the NH. In

all seasons (except for MAM in the SH) and annual series for both hemispheres and the entire globe, warming rates are faster for 1979–2014 than for 1900–2014. For the NH,

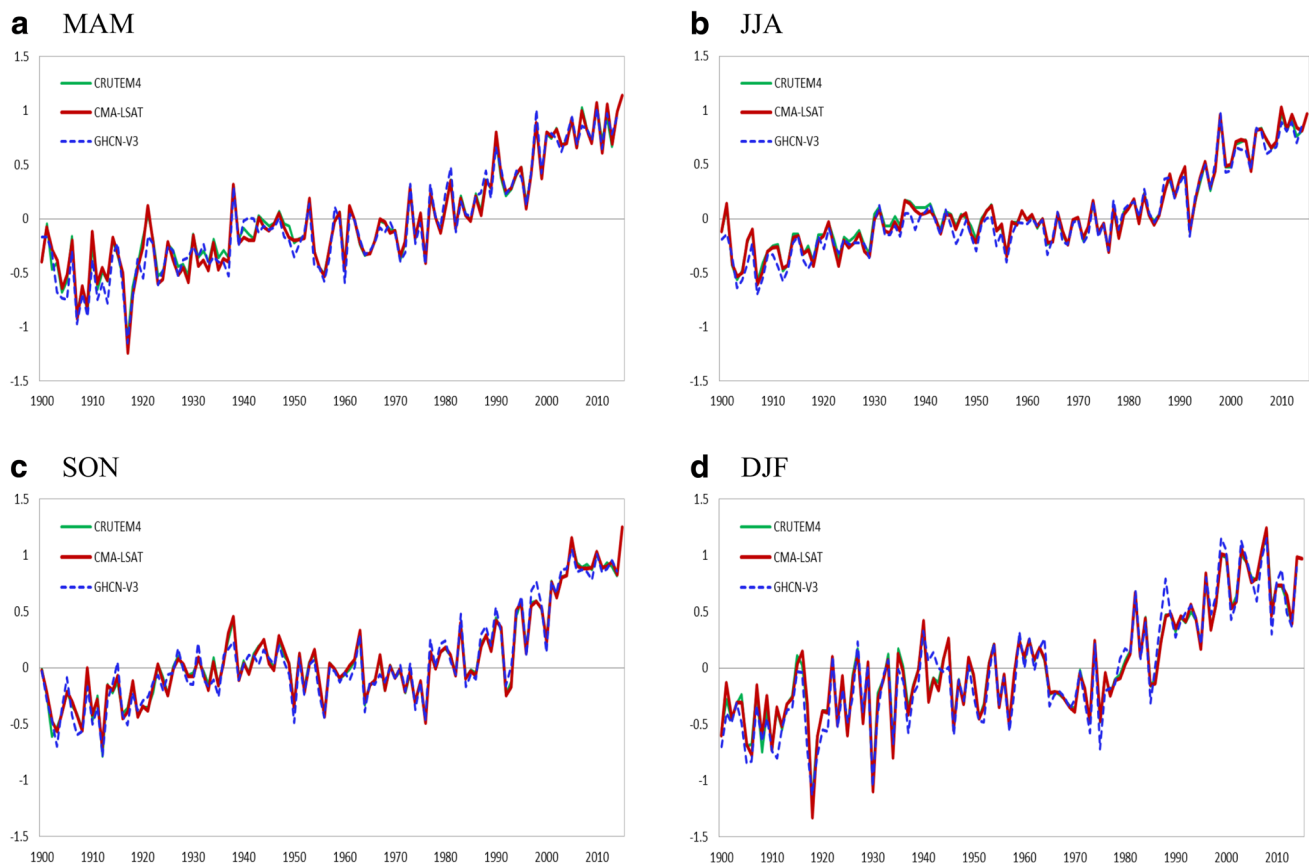


Fig. 10 Seasonal mean LSAT anomaly time series for the entire globe during 1900–2014

the warming trends for 1979–2014 show slight differences between the seasons. Warming is greatest during MAM and lowest in DJF, while for the SH, warming is the lowest during MAM and greatest during SON.

DJF shows a cooling trend over the 1998–2014 period (Table 6) whilst the other three seasons show warming. The DJF cooling is confined to the Northern Hemisphere, with the Southern Hemisphere showing warming in this season. The Southern Hemisphere shows strong warming for 1998–2014 in SON, and weak trends in MAM and JJA.

7 Summary

Motivated by the need to improve station coverage over Asia and to provide real-time monitoring of LSAT, in this paper we have detailed an effort to develop the CMA-LSAT dataset of monthly LSAT from 1900 to present. This data set, which is freely available from the NMIC of CMA, is a collaborative product of scientists at the CMA and many developers of some global, regional, national homogenized SAT datasets. This new data set benefits from these datasets

and from the improvements to the dataset described in this paper. The main characteristics of the dataset include:

1. The new database used 14 sources of LSAT and it contains 12,374 stations, of which 9765 could be used in gridding (the average start year is 1938) because they contain sufficient data for the 1961–1990 reference period to calculate the average. The records from the remaining 2609 stations are relatively short (average length is 18 years) and start later (the average start year is 1966). Spatial coverage is improved compared to two other datasets (GHCN-V3 and CRUTEM4) during 1900–2014 and reaches a maximum during the 1951–1990 period. It is worth noting that there have been consistently more stations in the CMA-LSAT dataset than the other cited datasets since 1990.
2. For the new dataset, we used homogenized data. About 50% of these data collected by CMA (including the “doubtful” and “suspect” stations from ECA&D) and homogenized by us using the PMT method (Wang 2008a, b), while the rest were obtained from existing national and regional homogenized datasets for Australia, Canada, China, USA, HISTALP and the “useful”

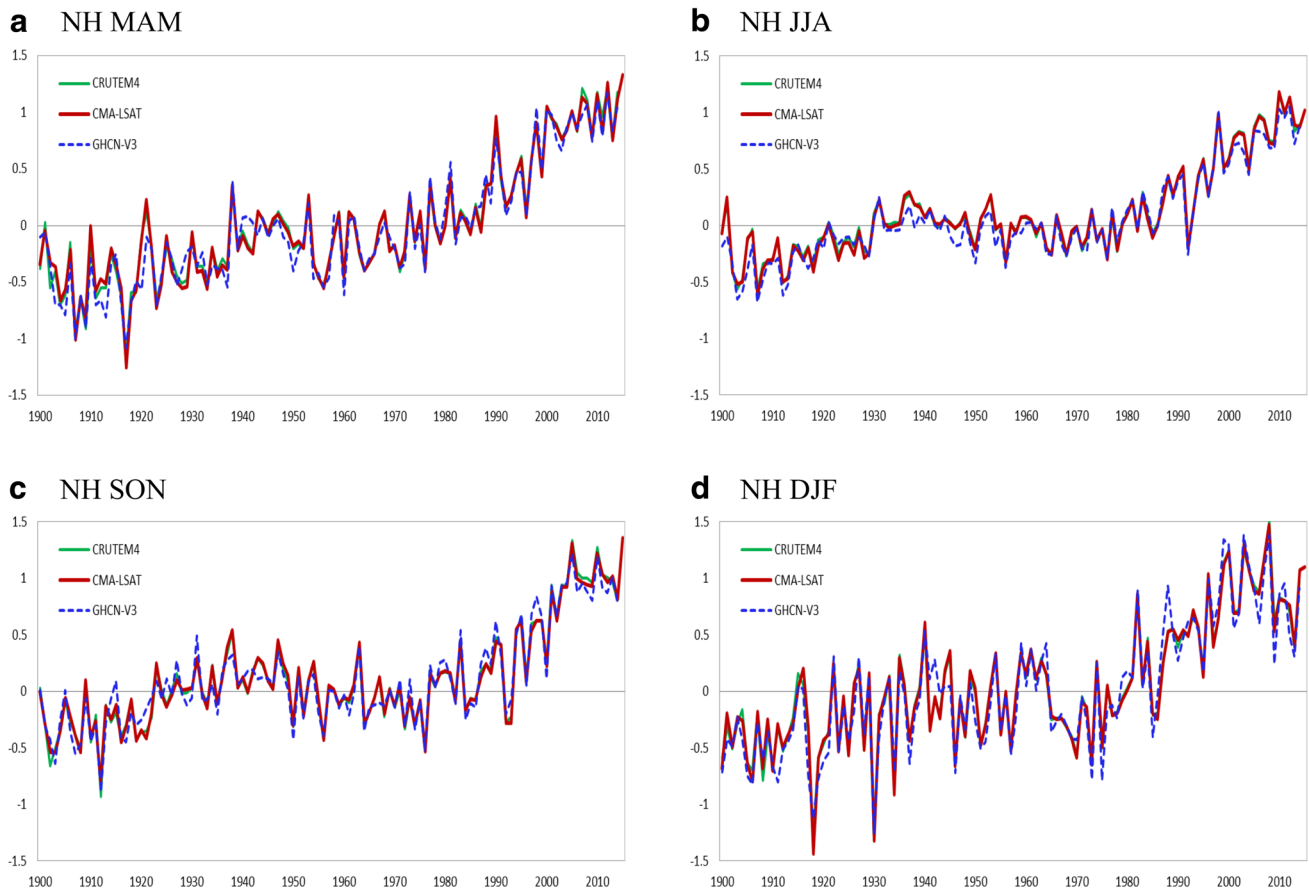


Fig. 11 Seasonal mean LSAT anomaly time series for the Northern Hemisphere during 1900–2014

station series from ECA&D. The improved homogeneity and regional coverage over Asia and other regions increase the usefulness of this dataset for regional climate change assessment.

3. The global and hemispheric average series developed from the CMA-LSAT dataset were also compared with the results from two other centers (NCEI: Peterson et al. 1998; and CRUTEM4: Jones et al. 2012). As reported elsewhere (Peterson et al. 1998; Jones et al. 1999; Folland and Karl 2001), the trends for the NH agree very well, whereas slightly larger differences occur in the SH (CMA-LSAT shows more similarity to CRUTEM4 than to GHCN3). Based on CMA-LSAT, the best evaluation of the trends for the global, NH, and SH land are, respectively, 0.102 ± 0.006 , 0.107 ± 0.007 , and 0.083 ± 0.005 °C/decade during 1900–2014; 0.272 ± 0.025 , 0.305 ± 0.030 , and 0.142 ± 0.021 °C/decade during 1979–2014; and 0.124 ± 0.057 , 0.150 ± 0.056 , and 0.120 ± 0.081 °C/decade during 1998–2014.

Acknowledgements This study was supported by the National Special Public Welfare Research Fund (Nos. GYHY201406016 and GYHY201206012), China Meteorological Administration Special Foundation for Climate Change (CCSF201438), and the Natural Science Foundation of China (Nos. 91546117 and 71373131). We thank the many contributors of data, which made establishment of the dataset possible.

Appendix

All the datasets have been made available through the national meteorological services website (<http://10.1.64.154/portal/web-home.index>).

Monthly updates will be based on a three-step procedure (Fig. 13), as follows. (1) The first round updates are based on hourly real-time data from the WMO Global Telecommunications System (GTS) and Integrated Surface Database (ISD) from NCEI. The daily values are calculated every day from the merged (GTS and ISD) and quality-controlled hourly datasets, and then the monthly temperature is

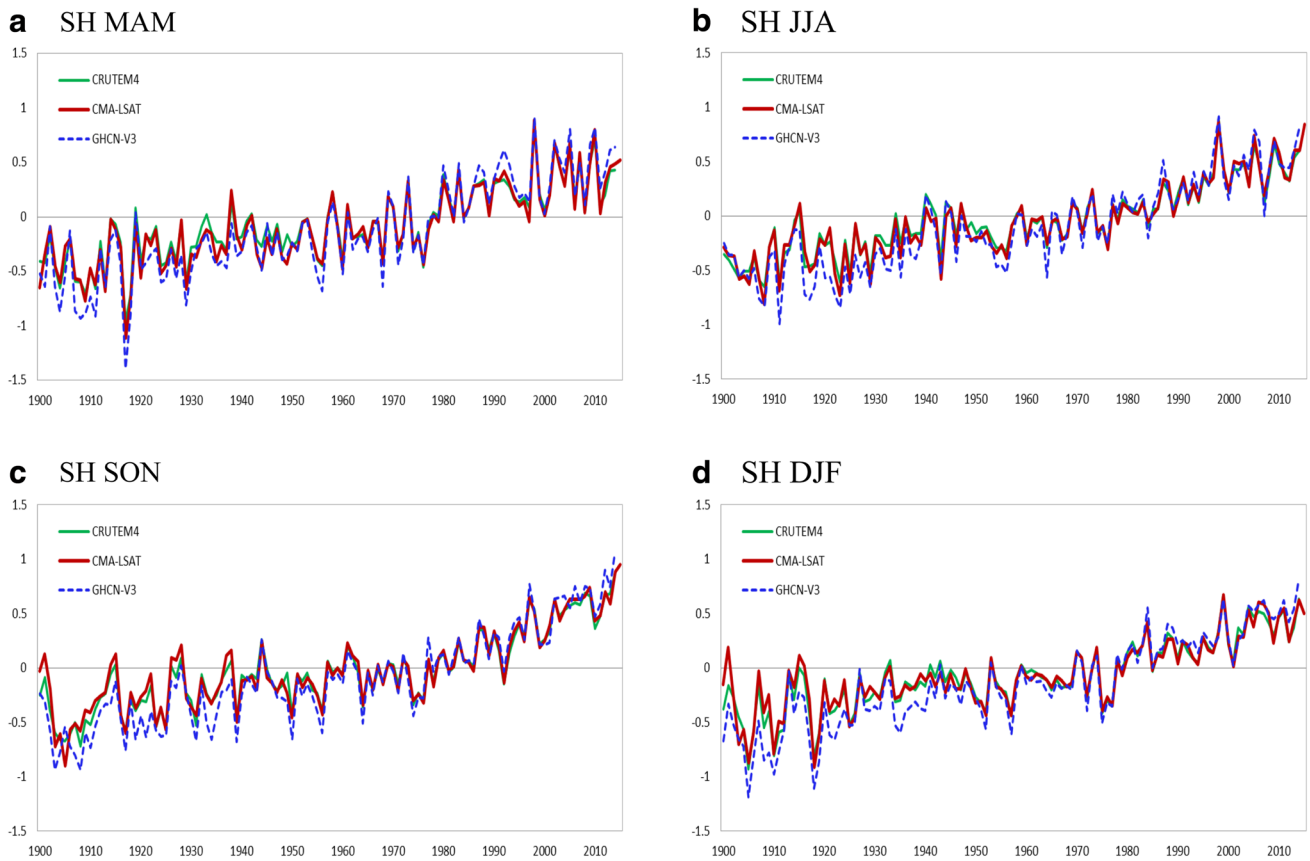


Fig. 12 Seasonal mean LSAT anomaly time series for the Southern Hemisphere during 1900–2014

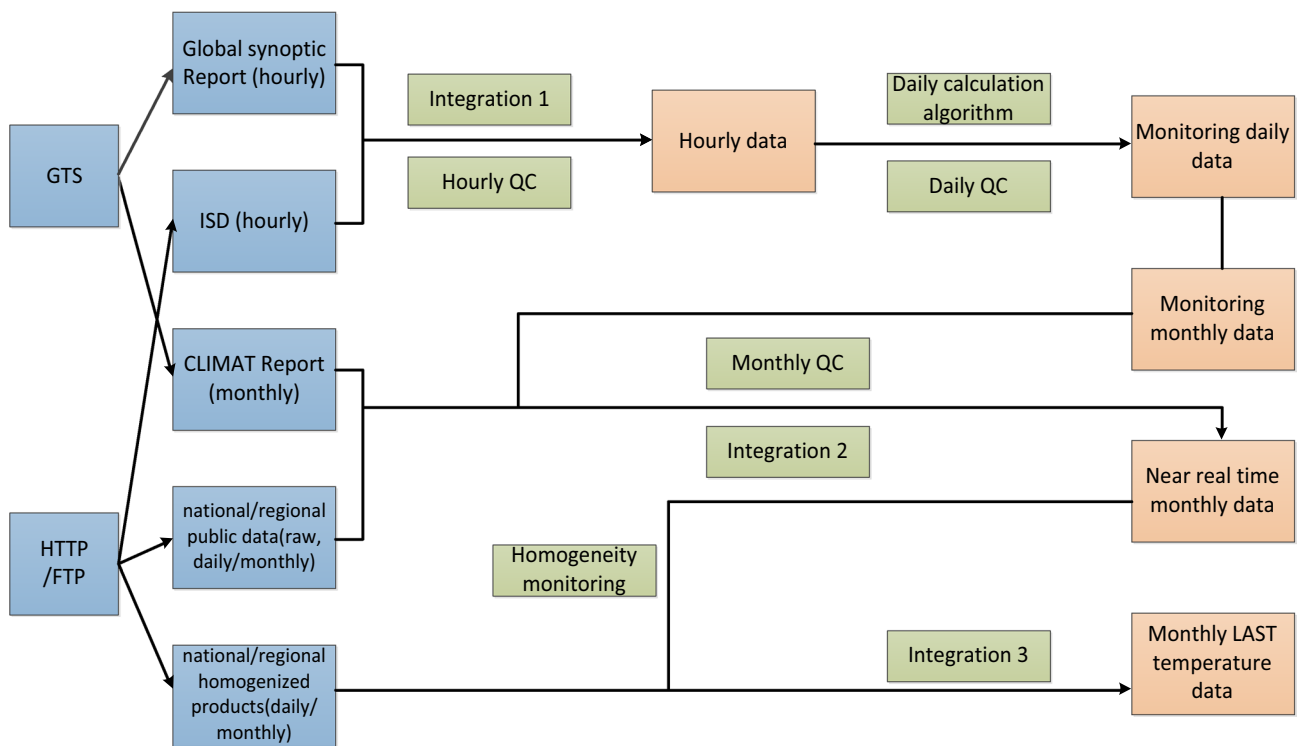


Fig. 13 Near-real-time integration and updating procedures for CMA-LSAT

obtained and updated in the dataset at the beginning of the following month. (2) Monthly mean values from CLIMAT message through GTS are provided to update the dataset on about the 20th day of the following month; this is the second round of dataset update. In the two steps, only the WMO stations are matched to be updated. The identification of the same station is performed by the WMO station number. (3) Other monthly data from existing homogenized datasets of the data sources listed in Table 1 are accessed and updated to the dataset. This round of updates is not carried out regularly and must be based on the update of each homogenization data source. Generally speaking, the last round of data update replaces the previous two rounds if the stations are found to be the same in previous rounds of data update. In the step, the identification of the same station is performed by the data source symbol and the station number. It is worth noting that, there is no new merging algorithm to be used during the process of updating.

References

- Alexander LV et al (2006) Global observed changes in daily climate extremes of temperature and precipitation. *J Geophys Res* 111:D18116. doi:[10.1029/2005JD006290](https://doi.org/10.1029/2005JD006290)
- Auer I et al (2007) HISTALP—Historical instrumental climatological surface time series of the greater Alpine region 1760–2003. *Int J Climatol* 27:17–46
- Brohan P, Kennedy J, Harris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *J Geophys Res* 111:D12106. doi:[10.1029/2005JD006548](https://doi.org/10.1029/2005JD006548)
- Cao LJ, Zhao P, Yan ZW et al (2013) Instrumental temperature series in eastern and central China back to the 19th century. *J Geophys Res Atmos* 118(15):8197–8207
- Cowtan K, Way RG (2014) Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q J R Meteorol Soc* 140(683):1935–1944. doi:[10.1002/qj.2297](https://doi.org/10.1002/qj.2297)
- Dai A, Wang J, Thorne PW, Parker DE, Haimberger L, Wang XL (2011) A new approach to homogenize daily radiosonde humidity data. *J Clim* 24:965–991. doi:[10.1175/2010JCLI3816.1](https://doi.org/10.1175/2010JCLI3816.1)
- Dai A, Fyfe J, Xie S, Dai X (2015) Decadal modulation of global surface temperature by internal climate variability. *Nat Clim Change* 5:555–559. doi:[10.1038/nclimate2605](https://doi.org/10.1038/nclimate2605)
- Ding Y, Dai X (1994) Temperature variation in China during the Last 100 years. *Meteorology* 20:19–26
- Domonkos P, Venema V, Mestre O (2013) Efficiencies of homogenisation methods: our present knowledge and its limitation. In: Proceedings of the Seventh seminar for homogenization and quality control in climatological databases. <http://www.wmo.int/pages/prog/wcp/wcdmp/documents/finalWCDMP78.pdf>
- Duan A, Xiao Z (2015) Does the climate warming hiatus exist over the Tibetan Plateau? *Sci Rep* 5:13711. doi:[10.1038/srep13711](https://doi.org/10.1038/srep13711)
- Durre I, Menne MJ, Vose R (2007) Strategies for evaluation quality assurance procedures. *J Appl Meteorol Climatol*. doi:[10.1175/2007JAMC1706.1](https://doi.org/10.1175/2007JAMC1706.1)
- Feng S, Hu Q, Qian W (2004) Quality control of daily meteorological data in China, 1951–2000: a new dataset. *Int J Climatol* 24:853–870
- Folland CK, Karl TR (2001) Recent rates of warming in marine environment meet controversy. *Eos Trans Am Geophys Union* 82(40):453–461
- Hansen J, Ruedy R, Glasco J et al (1999) GISS analysis of surface temperature change. *J Geophys Res* 104:30997–31022
- Hansen J, Sato M, Ruedy R, Lo K, Lea DW, Medina-Elizade M (2006) Global temperature change. *Proc Nat Acad Sci* 103(39):14288–14293
- Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. *Rev Geophys* 48:RG4004. doi:[10.1029/2010RG000345](https://doi.org/10.1029/2010RG000345)
- Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution monthly grids of monthly climatic observations: the CRU TS 3.10 dataset. *Int J Climatol* 34:623–642. doi:[10.1002/joc.3711](https://doi.org/10.1002/joc.3711)
- Hartmann DL, AMG Klein Tank, Rusticucci M, Alexander LV, Brönnimann S, Charabi Y, Dentener FJ, Dlugokencky EJ, Easterling DR, Kaplan A, Soden BJ, Thorne PW, Wild M, Zhai PM (2013) Observations: atmosphere and surface. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Miedgley PM (eds) *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 University Press, Cambridge
- Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) (2001) *Climate change 2001, in the scientific basis*. Cambridge University Press, Cambridge
- IPCC (2007) *Climate Change (2007) The Physical Science Basis*. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.) *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Jaccard P (1901) Etude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull de la Soc Vaud des Sci Nat* 37:547–579
- Jones PD (1994) Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *J Clim* 7(11):1794–1802
- Jones PD (2016) The reliability of global and hemispheric surface temperature records. *Adv Atmos Sci* 33(3):269–282. doi:[10.1007/s00376-015-5194-4](https://doi.org/10.1007/s00376-015-5194-4)
- Jones PD, Moberg A (2003) Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. *J Clim* 16:206–223
- Jones PD, Raper SCB, Santer B, Cherry BSB, Goodess C, Kelly PM, Wigley TML, Bradley RS, Diaz HF (1985) A grid point surface air temperature data set for the northern hemisphere, TR022. Department of Energy, Washington
- Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its changes over the past 150 years. *Rev Geophys* 37(2):173–199
- Jones PD, Lister DH, Li Q (2008) Urbanization effects in large-scale temperature records, with an emphasis on China. *J Geophys Res* 113:D16122. doi:[10.1029/2008JD009916](https://doi.org/10.1029/2008JD009916)
- Jones PD, Lister DH, Osborn TJ, Harpham C, Salmon M, Morice CP (2012) Hemispheric and large-scale land-surface air temperature variations: an extensive revision and an update to 2010. *J Geophys Res* 117:D05127. doi:[10.1029/2011JD017139](https://doi.org/10.1029/2011JD017139)
- Karl TR, Arguez A, Huang B, Lawrimore JH, McMahon JR, Menne MJ, Peterson TC, Vose RS, Zhang H-M (2015) Possible artifacts of data biases in the recent global surface warming hiatus. *Science* 348:1469–1472. doi:[10.1126/science.aaa5632](https://doi.org/10.1126/science.aaa5632)
- Klein Tank AMG, Wijngaard JB, Konnen GP et al (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. *Int J Climatol* 22:1441–1453

- Kuglitsch FG, Auchmann R, Bleisch R, Broennigmann S, Martius O, Stewart M (2012) Break detection of annual Swiss temperature series. *J Geophys Res* 117:D13105. doi:[10.1029/2012JD017729](https://doi.org/10.1029/2012JD017729)
- Lawrimore JH, Menne MJ, Gleason BE, Williams CN, Wuertz DB, Vose RS, Rennie J (2011) An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *J Geophys Res* 116:D19121. doi:[10.1029/2011jd016187](https://doi.org/10.1029/2011jd016187)
- Legates DR, McCabe GJ (1999) Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model evaluation. *Water Resour Res* 35:233–241. doi:[10.1029/1998WR900018](https://doi.org/10.1029/1998WR900018)
- Lewandowsky S, Risbey J, Oreskes N (2015) The “Pause” in Global Warming: turning a routine fluctuation into a problem for science. *Bull Am Meteorol Soc* 97(5):723–733. doi:[10.1175/BAMS-D-14-00106.1](https://doi.org/10.1175/BAMS-D-14-00106.1)
- Li Q (2013) Development of homogenized data sets in China and the possible contributions to global/regional products, International Workshop on Climate Data Requirements and Applications, March 4–8, Nanjing
- Li Q, Dong WJ (2009) Detection and adjustment of undocumented discontinuities in Chinese temperature series using a composite approach. *Adv Atmos Sci* 26(1):143–153
- Li Z, Yan ZW (2010) Application of multiple analysis of series for homogenization (MASH) to Beijing daily temperature series 1960–2006. *Adv Atmos Sci* 27(4):777–787
- Li Q, Liu X, Zhang H et al (2004a) Detecting and adjusting on temporal inhomogeneity in Chinese mean surface air temperature datasets. *Adv Atmos Sci* 21:260–268
- Li Q, Zhang HZ, Liu XN, Huang J-Y (2004b) Urban heat island effect on annual mean temperature during the last 50 years in China. *Theor Appl Climatol* 79:165–174
- Li Q, Zhang H, Chen J, Li W, Liu X, Jones P (2009) A mainland China homogenized historical temperature dataset of 1951–2004. *Bull Am Meteorol Soc* 90:1062–1065. doi:[10.1175/2009BAMS2736.1](https://doi.org/10.1175/2009BAMS2736.1)
- Li Q, Dong W, Li W, Gao X, Jones P, Kennedy J, Parker D (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](https://doi.org/10.1007/s11434-010-3209-1)
- Li Q, Li W, Si P et al (2010b) Assessment of surface air warming in northeast China, with emphasis on the impacts of urbanization. *Theor Appl Climatol*. doi:[10.1007/s00704-009-0155-4](https://doi.org/10.1007/s00704-009-0155-4)
- Li Q, Yang S, Xu WH et al (2015) China experiences recent warming hiatus. *Geophys Res Lett*. doi:[10.1002/2014GL062773](https://doi.org/10.1002/2014GL062773)
- Li Q, Zhang L, Xu W et al (2016) Comparisons of time series of annual mean surface air temperature for China since 1900: observations, model simulations and extended reanalysis. *Bull Am Meteorol Soc*. doi:[10.1175/BAMS-D-16-0092.1](https://doi.org/10.1175/BAMS-D-16-0092.1) (in press)
- Mann ME et al (2016) The likelihood of recent record warmth. *Sci Rep* 6:19831. doi:[10.1038/srep19831](https://doi.org/10.1038/srep19831)
- Menne MJ, Williams JR (2009) Homogenization of temperature series via pairwise comparisons. *J Clim* 22:1700–1717
- Oke TR (2004) Initial Guidance to Obtain Representative Meteorological Observations at Urban Sites, Instruments and Methods of Observation Program, IOM Report No. 81, WMO/TD 1250, World Meteorological Organization, Geneva
- Parker DE (1994) Effects of changing exposure of thermometers at land stations. *Int J Climatol* 14:1–31
- Parker DE (2004) Large-scale warming is not urban. *Nature* 432:290–290
- Parker DE (2006) A demonstration that large-scale warming is not urban. *J Clim* 19:2882–2895
- Peterson TC (2003) Assessment of urban versus rural in situ surface temperatures in the contiguous United States: no difference found. *J Clim* 16:2941–2959
- Peterson TC, Easterling DR (1994) Creation of homogeneous composite climatological reference series. *Int J Climatol* 14:671–679
- Peterson TC, Owen TW (2005) Urban heat island assessment: metadata are important. *J Clim* 18:2637–2646
- Peterson TC, Vose RS (1997) An overview of the global historical climatology network temperature data base. *Bull Am Meteorol Soc* 78:2837–2849
- Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N et al (1998) Homogeneity adjustments of in situ atmospheric climate data: a review. *Int J Climatol* 18:1493–1517
- Ren G, Chu Z, Zhou Y et al (2005) Recent progresses in studies of regional temperature changes in China. *Clim Environ Res* 10(4):701–716
- Rennie JJ, Lawrimore JH, Gleason BE, Thorne PW, Morice CP, Menne MJ, Williams CN, Gambi de Almeida W, Christy JR, Flannery M, Ishihara M, Kamiguchi K, Klein-Tank AMG, Mhanda A, Lister DH, Razuvaev V, Renom M, Rusticucci M, Tandy J, Worley SJ, Venema V, Angel W, Brunet M, Dattore B, Diamond H, Lazzara MA, Le Blancq F, Luterbacher J, Mächel H, Revadekar J, Vose RS, Yin X (2014) The international surface temperature initiative global land surface databank: monthly temperature data release description and methods. *Geosci Data J* 1(2):75–102
- Rohde R, Muller RA, Jacobsen R et al (2013) A new estimate of the average earth surface land temperature spanning 1753 to 2011. *Geoinform Geostat*. doi:[10.4172/gigs.1000101](https://doi.org/10.4172/gigs.1000101)
- Simmons AJ, Berrisford P, Dee DP, Hersbach H, Hirahara S, Thépaut J-N (2017) A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Q J R Meteorol Soc* 143:101–119. doi:[10.1002/qj.2949](https://doi.org/10.1002/qj.2949)
- Slingo J et al (2013) The recent pause in global warming parts 1–3 Rep., The met office, FitzRoy Road, Exeter, UK
- Smith TM, Reynolds RW, Lawrimore J (2008) Improvements to NOAA’s historical merged land–ocean surface temperature analysis (1880–2006). *J Clim* 21:2283–2296
- Sun Q, Miao C, Duan Q et al (2014) Would the ‘real’ observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. *Environ Res Lett* 9(1):015001. doi:[10.1088/1748-9326/9/1/015001](https://doi.org/10.1088/1748-9326/9/1/015001) (1–16)
- Trewin BC (2004) Effects of changes in algorithms used for the calculation of Australian mean temperature. *Aust Meteorol Mag* 53:1–11
- Trewin BC (2010) Exposure, instrumentation and observing practice effects on land temperature measurements. *Wiley Interdisciplinary reviews. Clim Change* 1:409–506. doi:[10.1002/wcc.46](https://doi.org/10.1002/wcc.46)
- Trewin BC (2013) A daily homogenized temperature data set for Australia. *Int J Climatol* 33:1510–1529
- Venema VKC, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Stepanek P, Zahradnicek P, Viarre J, Müller-Westermeier G, Lakatos M, Williams CN, Menne MJ, Lindau R, Rasol D, Rustemeier E, Kolokythas K, Marinova T, Andresen L, Acquafotta F, Fratianni S, Cheval S, Klancar M, Brunetti M, Gruber C, Prohom Duran M, Likso T, Esteban P, Brandsma T (2012) Benchmarking homogenization algorithms for monthly data. *Clim Past* 8:89–115. doi:[10.5194/cp-8-89-2012](https://doi.org/10.5194/cp-8-89-2012)
- Vincent LA, Wang XL, Milewska EJ et al (2012) A second generation of homogenized Canadian monthly surface air temperature for climate trend analysis. *J Geophys Res* 117:D18110. doi:[10.1029/2012JD017859](https://doi.org/10.1029/2012JD017859)
- Vincent LA, Zhang X, Brown RD, Feng Y, Mekis E, Milewska EJ, Wan H, Wang XL (2015) Observed trends in Canada’s climate and influence of low-frequency variability modes. *J Clim* 28:4545–4560 (00697.1)
- Vose RS, Schmoyer RL, Steurer PM, Peterson TC, Heim R, Karl TR, Eischeid J (1992) The Global Historical Climatology Network: Long-term monthly temperature, precipitation, sea level pressure, and station pressure data. ORNL/CDIAC-53, NDP-041, 325 pp. [Available from Carbon Dioxide Information Analysis

- Center, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831]
- Wan H, Wang XL, Swail VR (2010) Homogenization and trend analysis of Canadian near-surface wind speeds. *J Clim* 23:1209–1225
- Wang XL (2008a) Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. *J App Meteor Climatol* 47:2423–2444. doi:[10.1175/2008JAMC1741.1](https://doi.org/10.1175/2008JAMC1741.1)
- Wang XL (2008b) Penalized maximal F test for detecting undocumented mean shift without trend change. *J Atmos Ocean Technol* 19:368–384
- Wang XL, Feng Y (2010) RHtestsV4 user manual, climate research division, science and technology branch, environment Canada, Toronto, Ontario, Canada. <http://etccdi.pacificclimate.org/software.shtml>
- Wang KC, Zhou CL (2015) Regional contrasts of the warming rate over land significantly depend on the calculation methods of mean air temperature. *Sci Rep* 5:12324. doi:[10.1038/srep12324](https://doi.org/10.1038/srep12324)
- Wang XL, Wen QH, Wu Y (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. *J Appl Meteorol Climatol* 46:916–931
- Wang XL, Chen H, Wu Y, Feng Y, Pu Q (2010) New techniques for detection and adjustment of shifts in daily precipitation data series. *J Appl Meteor Climatol* 49(12):2416–2436. doi:[10.1175/2010JAMC2376.1](https://doi.org/10.1175/2010JAMC2376.1)
- Wang XL, Feng Y, Vincent LA (2013) Observed changes in one-in-20 year extremes of Canadian surface air temperatures. *Atmos Ocean* 52(3):222–231. doi:[10.1080/07055900.2013.818526](https://doi.org/10.1080/07055900.2013.818526)
- Wang J, Xu C, Hu M et al (2014) A new estimate of the China temperature anomaly series and uncertainty assessment in 1900–2006. *J Geophys Res Atmos*. doi:[10.1002/2013JD020542](https://doi.org/10.1002/2013JD020542)
- Wang F, Ge Q, Wang S, Li Q, Jones P (2015) A new estimation of urbanization contribution to the warming trend in China. *J Clim* 28:8923–8938. doi:[10.1175/JCLI-D-14-00427.1](https://doi.org/10.1175/JCLI-D-14-00427.1)
- Wickham C et al (2013) Influence of urban heating on the global temperature land average using rural sites identified from MODIS classifications. *Geoinfor Geostat*. doi:[10.4172/2327-4581.1000104](https://doi.org/10.4172/2327-4581.1000104)
- Wijngaard JB, Klein Tank AMG, Konnen GP et al (2003) Homogeneity of 20th century European daily temperature and precipitation series. *Int J Climatol* 23:679
- Williams CN, Menne MJ, Thorne PW (2012) Benchmarking the performance of pairwise homogenization of surface temperatures in the United States. *J Geophys Res* 117:D05116. doi:[10.1029/2011JD016761](https://doi.org/10.1029/2011JD016761)
- Willmott CJ (1981) On the validation of models. *Phys Geogr* 2:184–194
- Willmott CJ, Ackleson SG, Davis RE, Feddema JJ, Klink KM, Legates DR, O'Donnell J, Rowe CM (1985) Statistics for the evaluation and comparisons of models. *J Geophys Res* 90:8995–9005. doi:[10.1029/JC090iC05p08995](https://doi.org/10.1029/JC090iC05p08995)
- Xu W, Li Q, Wang X et al (2013) Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. *J Geophys Res Atmos* 118(17):9708–9720
- Yan Z, Li Z, Li Q et al (2009) Effects of site-change and urbanisation in the Beijing temperature series 1977–2006. *Int J Climatol*. doi:[10.1002/joc.1971](https://doi.org/10.1002/joc.1971)
- Yin H, Donat MG, Alexander LV et al (2015) Multi-dataset comparison of gridded observed temperature and precipitation extremes over China. *Int J Climatol* 35(10):2809–2827. doi:[10.1002/joc.4174](https://doi.org/10.1002/joc.4174)
- You Q, Kang S, Anguilar E et al (2011) Changes in daily climate extremes in China and its connection to the large scale atmospheric circulation during 1961–2003. *Clim Dyn* 36:2399–2417. doi:[10.1007/s00382-009-0735-0](https://doi.org/10.1007/s00382-009-0735-0)
- Zhai P, Chao Q, Zou X (2004) Progress in China's climate change study in the 20th century. *J Geogr Sci* 14:3–11. doi:[10.1007/BF02841101](https://doi.org/10.1007/BF02841101)
- Zhai P, Yu R, Guo Y, Li Q, Ren X, Wang Y, Xu W, Liu Y, Ding Y (2016) The strong El Niño in 2015/2016 and its dominant impacts on global and China's climate. *J Meteorol Res* 74(3):309–321. doi:[10.11676/qxxb2016.049](https://doi.org/10.11676/qxxb2016.049)
- Zhang X, Hegerl G, Zwiers FW, Kenyon J (2005) Avoiding inhomogeneity in percentile-based indices of temperature extremes. *J Clim* 18:1641–1651. doi:[10.1175/JCLI3366.1](https://doi.org/10.1175/JCLI3366.1)
- Zhou L, Dickinson RE, Tian Y, Fang J, Li Q, Kaufmann RK (2004) Evidence for a significant urbanization effect on climate in China. *Proc Natl Acad Sci USA* 101:9540–9544