JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD039914

Special Collection:
Land-atmosphere coupling: measurement, modelling and analysis

Sampling Error of Mean and Trend of Nighttime Air Temperature at Weather Stations Over China Using Satellite Data as Proxy

Linghong Chen[1] and Kaicun Wang[2]

[1]State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, [2]Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China

Key Points:

- Sampling error of mean and trend of temperature collected at weather stations at nighttime over China were investigated
- The sampling error in national mean temperature were found to be more than 3°C due to uneven distribution of weather stations
- The sampling error in temperature trend account for one thirds of the national mean temperature trend due to local urbanization

Correspondence to:
K. Wang,
kcwang@pku.edu.cn

Citation:
Chen, L., & Wang, K. (2024). Sampling error of mean and trend of nighttime air temperature at weather stations over China using satellite data as proxy. *Journal of Geophysical Research: Atmospheres*, *129*, e2023JD039914. https://doi.org/10.1029/2023JD039914

Received 31 AUG 2023
Accepted 7 JUL 2024

Abstract Meteorological observations of surface air temperature have provided fundamental data for climate change detection and attribution. However, the weather stations are unevenly distributed, and are still very sparse in remote regions. The possible sampling error is well known, but not well quantified because we are lack of the adequate and regularly distributed measurements. The high resolution of satellite land surface temperature retrieval during night time provide a nice proxy for near surface temperature as both temperatures controlled by surface longwave radiative cooling and the nocturnal temperature inversion depress land-atmosphere turbulent exchange. The sampling error of mean value and trend were assessed by comparing station point measurements (pixel of ~0.01°) with grid (1°) mean and national mean from 2001 to 2021. This method permits us to make the first assessment of under-sampling error and spatial representative error on both national mean and trend of air temperature during nighttime collected at ~2,400 weather stations over China. The sampling error in national mean temperature is more than 3°C. The under-sampling error due to lack of observation explains two thirds and the spatial representative error due to the difference between station and grid/regional mean elevation contribute the other one third. The sampling error in trend account for one third of the national mean trend. The urban heat island effect associated with urbanization around the weather stations (spatial representative error) can explain four fifths of the sampling error in trend, which is consistent with existing studies based on air temperature collected at paired weather station.

Plain Language Summary Meteorological observations provide fundamental measurements of surface air temperature. These station-based temperature records are usually processed to grid data sets considering the uneven distribution of meteorological stations. The gridding procedure represents a non-uniformly sampling. In this paper, we quantified the sampling error over China using remote sensed temperature record as proxy of near surface temperature during nighttime. The causes of sampling error both in mean value and trend were both quantified.

1. Introduction

The surface temperature over China increased at a rate higher than the global average and the United States located in the similar mid-latitudes (Li et al., 2020; Li, Long, et al., 2022; Sun et al., 2022). Rapid warming directly affects both earth system (Liu, Mi, et al., 2018; Yang et al., 2014) and human society (Yang et al., 2021) with a national population of 1.4 billion. Quantification of means and trends of temperature is necessary to detect and attribute climate change over China. These values may depend on possible different spatial averaging definitions and available measurement techniques (Lovejoy, 2017; Richardson et al., 2018).

Meteorological observations of surface air temperature provide fundamental evidences for regional and global warming. However, the weather stations are unevenly distributed and still very sparsely distributed in remote regions, such as the Tibetan Plateau and Taklimakan Desert (Zhou & Wang, 2016). Considering the uneven distribution of weather stations, the weather stations data is usually interpolated to grid data (Ge et al., 2013; Shen & Zhao, 2021; Sun et al., 2016), that is, dividing the region to lat-lon square grid boxes and interpolating the station data into grid mean.

In an ideal scenario, grid box should be minimized to ensure the precision of the data set. However, it is required adequate size to maintain a sufficient number of stations within each grid box. In the early years, the grid size of 5° × 5° was used due to the poor spatial distribution of available meteorological data (Jones & Moberg, 2003).

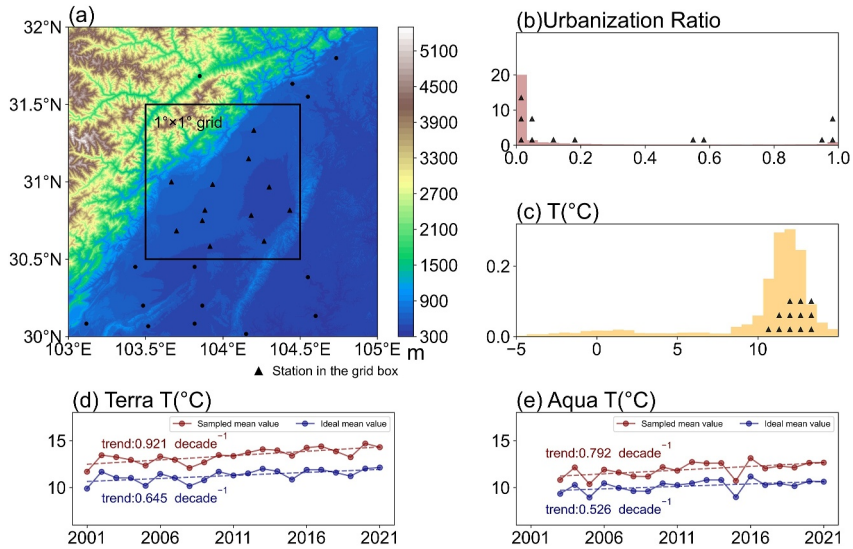

Figure 1. Schematic diagram of sampling error at grid scale in this study. (a) The map of elevation. In (a) the black box is a 1° × 1° grid scale, with stations within being represented by black triangles (stations outside the grid are represented by black points). In (b) and (c), the urban ratio and temperature at the grid scale are shown by histogram bars, while the values of the stations are represented by black triangles. The time series of the sampled mean result (the mean value of MODIS LST pixels located at stations) and the ideal mean value (the mean values of all MODIS LST pixels in the grid box) is shown in (d) and (e).

Recently, the grid size was usually set to 1° × 1° or smaller, for the increasing number of meteorological stations (Ma et al., 2017; Shen & Zhao, 2021). This gridding procedure introduces a biased spatial sampling approach, in which temperatures from station observations are averaged to estimate the mean temperature across all points within a grid.

If the weather stations have sufficient spatial representativeness and evenly distribution, producing sampling error tends to approach zero. But the representativeness of meteorological stations has been proven to be deficient (Bhowmik & Costa, 2015; Orlowsky & Seneviratne, 2014). The sampling error of the derived regional/national mean temperature which could induce considerable unreliability in both multi-years mean temperature and warming rate, are still not well quantified.

With higher spatial resolution data like remote sensing retrieved temperature, one can quantify the sampling error occurs in this interpolation process. Moderate Resolution Imaging Spectroradiometer land surface temperature (MODIS LST) data set was used to demonstrate the definition of both sampled result, true value and sampling error. The ideal value of grid temperature mean can be considered as the mean value of all pixels within a grid box, while we use a sampled result, the mean value of all pixels which are nearest to stations within the gird box to replace it. This sampling error may impact the accuracy of grid temperature data set.

We take a 1° × 1° grid in the Southwest China as an example (See Figure 1). Weather stations are frequently situated in close proximity to towns for instrument maintenance therefore deviating from the random sampling. Observations in uninhabited places, such as mountains, are lacking (Figure 1a). Meanwhile, distribution of impervious surface (the percentage of impervious surface is defined as urban ratio) around stations has a great influence on temperature sampling (Figure 1b). Excessive stations located in cities or towns in this grid. The sampled temperature distribution shows higher value than that of all pixels (Figure 1c). The series obtained by sampling result (the mean value of MODIS LST pixels located at stations) different from the ideal mean value series (the mean values of all MODIS LST pixels in the grid box), showing different warming trends. The differences both in mean values and trends of temperature is defined as spatial representative error, for this kind of sampling error due to the lack of spatial representativeness of existing station.

Using conventional data sets instead of directly averaging station data to grid data may reduce the sampling error in mean value. Previous studies have shown that the bias of re-analyses data and observational data was related to elevation difference between model grid and station (Ma et al., 2008; Yan et al., 2020; Zhou et al., 2018). For

example, the Berkeley Earth Land/Ocean Temperature Record (Rohde & Hausfather, 2020) used elevation as a factor in the kriging process over land and produced monthly absolute temperature fields and anomalies with 1° × 1° resolution. Data sets focus on the Chinese climate, such as CN05 (Xu et al., 2009) and CN05.1 (Wu & Gao, 2013), which also used elevation correction in the interpolation process.

In addition, some commonly used land near surface air temperature data sets ignored the correction of elevation by using anomaly rather absolute mean values. CRUTEM5 (Osborn et al., 2021) generated gridded fields of temperature anomalies (5° × 5° resolution) with an uncertainty model, the grid-mean equaled the arithmetic mean of any station anomalies that lie within each grid cell, without interpolation. GHCN (Menne et al., 2018) calculated the monthly average anomaly (5° × 5° resolution) as the arithmetic mean of all available station-based anomalies within the bounding box as well. As demonstrated in this study, the elevation based correction method can effectively reduce the impact on the mean values but can't reduce the spatial representative error on the trends.

Traditional homogenization methods were defective in detecting and adjusting the impact of the gradual urbanization around the weather stations, which amplified the warming trends of daily minimum and mean air temperature (Jiang et al., 2020). Most weather stations in China have undergone urbanization (Jiang et al., 2020; Zhang & Wang, 2021). It implied spatial representative error in the trend of temperature, which still need to be resolved. Studies have used methods, such as matchmaking between urban and rural stations, to quantify the effect of urbanization on station temperature observations (Karl et al., 1988; Wen et al., 2019; X. Yang et al., 2011). For example, Zhang et al. (2021) used 43 reference stations of 763 observation stations to adjust the urbanization bias. Ren et al. (2008) used 282 weather stations to estimate the contribution of urban warming to total annual mean surface air temperature change as 37.9%. Tysa et al. (2019) classified weather stations into six urbanization levels and quantified the urbanization contributions in China as 14.6%. These studies provided meaningful assessment of influence of urbanization at station scale. However, the matching of stations had high requirements, so the distribution of station pairs is sparse across the country. Therefore, the findings of previous study cannot be directly used to represent the sampling error of the temperature trend, although sampling error of trend was indeed very relevant to urbanization.

When we obtain the large-scale (global or national) average of temperature through existing observations, gridded data shows null value in a few places in western China, even if the grid box reaches a course size of 5° × 5°. Therefore, temperature of areas without weather stations were excluded in large-scale average and trend. This under-sampling error was first discussed by Cowtan and Way (2014), and HadCRUT4 was taken as an example. The study focused on the grids that missing observation data caused a negative error in the global temperature trend due to rapid warming in Africa and the polar regions. Huang et al. (2017) filled observational gaps an Arctic based on the data interpolating empirical orthogonal functions method and found that the amplified Arctic warming over the past decade had significantly contributed to a continual global warming trend, rather than a hiatus or slowdown from 1998 to 2012. Kadow et al. (2020) filled observational gaps in HadCRUT4 using combined with numerical climate model data. This under-sampling error also needs to be discussed in China when we need a national mean value, especially at a smaller grid size, not limited to 5° × 5° commonly used grid scale in global data sets.

Within a very small time and space scale, both spatial representative error and under-sampling error can be calculated by intensive observation network data sets. For example, Heihe Integrated Observatory Network (Liu, Li, et al., 2018) has provided air temperature data from 22 automatic weather stations at an ∼1° × 1° grid scale over the last 10 years. This observation amount is not enough for sampling error in national and climatological scale, and far less than the data set by remote sensing. MODIS LST data set is widely used in researches of global and regional warming (Li et al., 2021; Zhao et al., 2021). It provides LST data for a sufficiently long-time range for both daytime and nighttime (Duan et al., 2019).

We focus on the sampling error at nighttime. At daytime, solar radiation heats land surface and air successively. Phase difference between LST and air temperature exists until a period of time after sunset. Besides, proportion changes of latent heat and sensible heat at daytime also influences the insolation of land surface temperature and air temperature (Panwar et al., 2019; Vancutsem et al., 2010). At nighttime, turbulent energy fluxes are suppressed and are explosive, leading to the stable planet boundary layer stratification under longwave radiative cooling effect. LST couples with air temperature better because near surface temperature inversion limits the land–atmosphere interaction (Good, 2016; Steeneveld et al., 2006). The previous studies indicate that LST and air temperature showed similar variability at nighttime in spite of the temperature inversion (Anniballe et al., 2014;

Krishnan et al., 2015; Li et al., 2018; Sun et al., 2015). Besides, the strong correlation of MODIS LST and station observation minimum air temperature at nighttime is demonstrated (Mira et al., 2017; Zhu et al., 2013). It is true that air temperature is not exactly equal to land surface temperature during nighttime because of the temperature inversion. As explained in the method section, the study based on the difference of the temperatures between point and grid/national mean rather than temperature itself, which substantially reduced the impact of the assumption made here.

In this study, we first investigated the value of ∼2,400 weather stations' spatial representative error both in mean value and trend of temperature at the grid scale (1° × 1°) and national scale. The magnitude of the spatial representative error in temperature trend was investigated as five times large as the under-sampling error. Moreover, the possible influencing factors of sampling error were explored, using elevation and urban ratio data sets.

2. Data and Methods

2.1. Data

The longitude and latitude information of weather stations were obtained from the China Meteorological Data Service Center (CMDC) in 2018, and data from ∼2,400 stations were applied. We used the LST product from the MODIS onboard Terra and Aqua satellites (Index of/MOLT/MOD11A1.006 (usgs.gov)). The Terra satellite overpasses twice daily (∼10:30 AM and PM local time) since December 1999, and the Aqua satellite overpasses (∼1:30 PM and AM local time) since May 2002. The Collection 6 (C6) daily LST data set, including MOD11A1 (Terra) from 2001 to 2021 and MYD11A1(Aqua) from 2003 to 2021, was processed to monthly data. Considering the representation of land surface temperature to near surface air temperature, only LST data at night was selected as mentioned above.

The high-resolution (30 m) data set of global artificial impervious area (GAIA) (http://data.ess.tsinghua.edu.cn.) from 1985 to 2018 were developed by Gong et al. (2020). The GAIA data set has been widely used in urban climate research to present the process of urbanization (Jiang et al., 2020; Q. Ren et al., 2022; Zhang & Wang, 2021).

The SRTM (https://srtm.csi.cgiar.org/) digital elevation data is produced by CGIAR-CSI with a resolution of 90 m for over 80% of the globe. We used this product to represent the elevation of China.

We used additional data sets to validate our results of MODIS LST. The ground-based daily meteorological observations of near surface air temperature collected at 2,400 stations in China from 2001 to 2018 were obtained from the China Meteorological Administration as mentioned above. The ground-based air temperatures were used to prove the feasibility of using MODIS LST to calculate the sampling error of surface temperature. The MCD43A2 albedo data set produced daily using Terra and Aqua MODIS data at 500-m resolution, and was used to provide information about higher warming rate of Northeast China.

2.2. Method

2.2.1. Sampling Error at the Grid Scale

We defined the grid mean temperature sampling by the locations of weather stations as:

$$T_{station,i} = \frac{\sum_{j=1}^{n_1} T_{station,j}}{n_1}, \quad (1)$$

where $T_{station,j}$ is the temperature of the pixel with shortest distance to weather station j, and n_1 is the number of weather stations in grid i. That is, $T_{station,i}$ was calculated by averaging temperature values at the observational stations. Note that we used temperature data from MODIS LST instead of station observational minimum air temperature, to ensure consistence between the temperature records of station and whole grid. Meanwhile, similar systematical error of remote sensing observation would offset in subsequent subtracting.

Similarly, the grid mean temperature averaged by all pixels is written as:

$$T_{pixel,i} = \frac{\sum_{j=1}^{n_2} T_{pixel,j}}{n_2}, \quad (2)$$

where $T_{pixcel,j}$ is the temperature of the pixel and n_2 is the number of all pixels in grid i. That is, $T_{station,i}$ is calculated by averaging all temperature values. We regard this temperature as the ideal value; thus, the mean spatial representative sampling error of temperature is written as:

$$T_{ME,i} = \overline{T}_{station,i} - \overline{T}_{pixel,i}, \quad (3)$$

where $\overline{T}_{station,i}$ is the mean value of $T_{station,i}$ along the time axis and $\overline{T}_{pixel,i}$ is the mean value of $T_{pixel,i}$ along the time axis. The trend of temperature was obtained from the linear regression Equation 4:

$$\begin{cases} T_i = a + b \times time \\ T_{trend,i} = b \end{cases}, \quad (4)$$

where time is the year in sequence, from 2001 to 2021 for Terra and from 2003 to 2021 for Aqua. Similarly, the spatial representative sampling error of temperature trend is written as:

$$T_{TE,i} = T_{trend,station,i} - T_{trend,pixel,i}, \quad (5)$$

2.2.2. Sampling Error at the National Scale

While the locations of weather stations cause sampling error in a grid, there are grids that have no weather stations identified, leaving null values in nationwide calculations. We defined three kinds of temperature and used their difference to represent the sampling error mentioned above:

$$T_{SS} = \frac{\sum_{i=0}^{n_3} T_{station,i}\varphi_i}{n_3}, \quad (6)$$

$$T_{SG} = \frac{\sum_{i=0}^{n_3} T_{pixel,i}\varphi_i}{n_3}, \quad (7)$$

$$T_{AG} = \frac{\sum_{i=0}^{n_4} T_{pixel,i}\varphi_i}{n_4}, \quad (8)$$

where T_{SS} is the nationwide mean temperature of grids with weather stations while the grid value is calculated by the station sampling mean. T_{SG} is the nationwide mean temperature of grids with weather stations, while the grid value is calculated by the mean of all pixels. T_{AG} is the nationwide mean temperature of all grids while the grid mean value is calculated by all pixels. φ_i is the latitude weight. n_3 is the number of grids with weather stations, and n_4 is the number of all grids. Then, the three kinds of temperature mean sampling error can be written as:

$$T_{ME1} = \overline{T_{SS}} - \overline{T_{SG}}, \quad (9)$$

$$T_{ME2} = \overline{T_{SG}} - \overline{T_{AG}}, \quad (10)$$

$$T_{ME3} = \overline{T_{SS}} - \overline{T_{AG}}, \quad (11)$$

The three kinds of temperature mean relative sampling error can be written as:

$$T_{TRE1} = \frac{T_{trend,SS} - T_{trend,SG}}{T_{trend,SG}}, \quad (12)$$

$$T_{TRE2} = \frac{T_{trend,SG} - T_{trend,AG}}{T_{trend,AG}}, \quad (13)$$

$$T_{TRE3} = \frac{T_{trend,SS} - T_{trend,AG}}{T_{trend,AG}}, \quad (14)$$

2.2.3. Elevation

Previous studies have demonstrated that the difference of elevation between meteorological stations and reanalysis data is a dominant factor of sampling error in temperature mean value. This difference of elevation discrepancy can be quantified by the sampling error of elevation as well. The mentioned DEM data were resampled to 1 km resolution to match the MODIS LST data set, using bilinear interpolation method. Similarly, the sampling error of DEM of a grid i is written as:

$$DEM_{AE,i} = DEM_{station,i} - DEM_{pixel,i}, \quad (15)$$

where $DEM_{station,i}$ is the mean value of DEM pixels nearest with the locations of weather stations in grid i, $DEM_{pixel,i}$ is the mean value of all DEM pixels in grid i.

2.2.4. Urbanization Impact

We defined the urbanization ratio (UR) as the percentage of impervious surfaces in 1 km grid to match the resolution of the MODIS LST data set. The trend of the urbanization ratio is calculated by

$$\begin{cases} UR_i = a + b \times time \\ UR_{trend,i} = b \end{cases}, \quad (16)$$

where time is the year sequence, from 2001 to 2018 for the GAIA data set. Then the trend sampling error of UR of grid i is written as:

$$UR_{TE,i} = UR_{trend,station,i} - UR_{trend,pixel,i}, \quad (17)$$

where $UR_{trend,station,i}$ is the trend of the mean value of UR pixels sampled by the locations of weather stations in grid i, $UR_{trend,pixel,i}$ is the trend of the mean value of urban ratio of all pixels in grid i.

3. Results

3.1. Mean and Trend of Nighttime Surface Temperature

Figure 2 presents the mean and trend of land surface temperatures over China. The temperature mean value of at about 1:30 AM (Aqua) is slightly smaller than that at 10:30 PM (Terra). Scatter points of the Qinghai-Tibet Plateau show a greater warming trend, and the pixels with enhanced warming rate is associated with lake. While waterbodies show higher temperatures at night, the expansion area of lakes explains the issue (Li, Long, et al., 2022; Zhang et al., 2019; Zhao et al., 2022). Such pixels account for a very small proportion, thus making little impact in a 1° × 1° grid map (Figure 3). The area north of 45°N shows a greater warming trend in both Figures 2 and 3. The trend of snow water equivalent shows the same pattern with warming trend (Jiang et al., 2022). This suggests that the accumulation of snow during winter and its subsequent melting during spring will both increase. The resulting variations in albedo stemming from these processes exert influences on temperature. The trend of albedo in this area is positive in winter and turns to be negative in spring (Figure 4). Results of albedo trends and their difference substantiates our hypothesis. Increasing snow water in spring would enhance the warming trend of temperatures.

3.2. Spatial Representative Error at the Grid Scale

Figure 5 presents the representative error in mean and trend of temperature. For representative error of mean value (T_{ME}), high values are found in the Hengduan Mountains and Southeast Tibet, probably because weather stations tend to be built in valleys instead of peaks in mountain areas (Zhou et al., 2018). The edge of the Tarim Basin also shows high values, probably because the desert has a greater heat capacity, tends to be colder at night than the weather stations which located in oasis (Hu et al., 2014; Yao et al., 2022). The rest of the country shows mostly

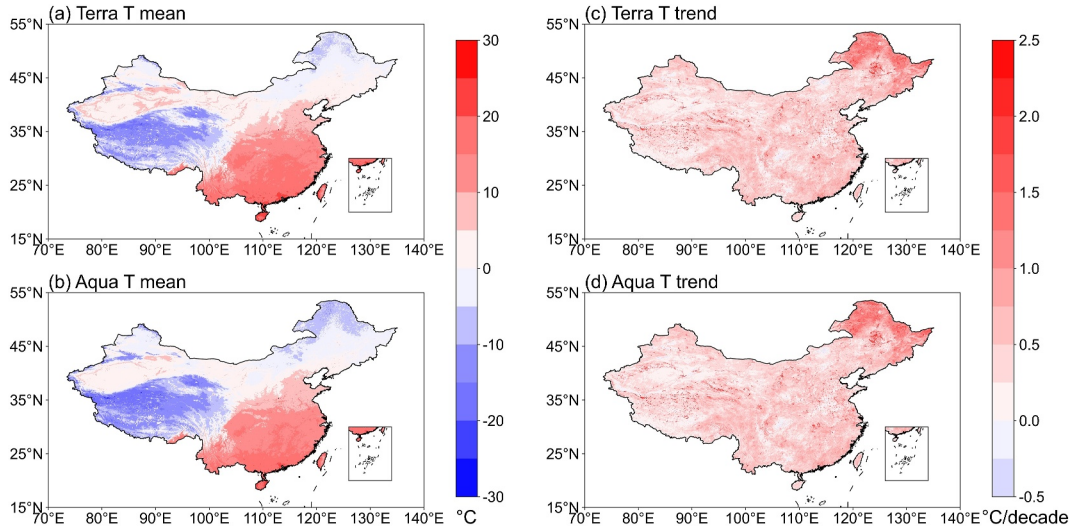

Figure 2. Multiyear average of MODIS land surface temperature (~1 km spatial resolution) at approximately nighttime (10:30 PM from Terra and 1:30 am from Aqua) from 2001 to 2021 (left) and their trend from 2001 to 2021 (right). For Aqua, time series are from 2003 to 2021.

positive values. The T_{ME} shows a right-skewed distribution (Figure 6), with national means of 1.11 ± 1. 9°C (Terra) and 1.12 ± 1.68°C (Aqua).

For spatial representative error of trend (T_{TE}), most grids of the country show positive values, except for the northeastern area. Many previous studies have illustrated the impact of urbanization on weather stations (Jiang et al., 2020; Wen et al., 2019), and the positive T_{TAE} has been proved to be caused by urbanization. As mentioned above, areas with low human activity are clearly affected by spring snowmelt, but stations surrounding towns tend to immediately remove snow after snowfall. The T_{TE} shows right skewed distribution as well, with national mean is 0.11 ± 0.22°C/decade (Terra) and 0.11 ± 0.26°C/decade (Aqua).

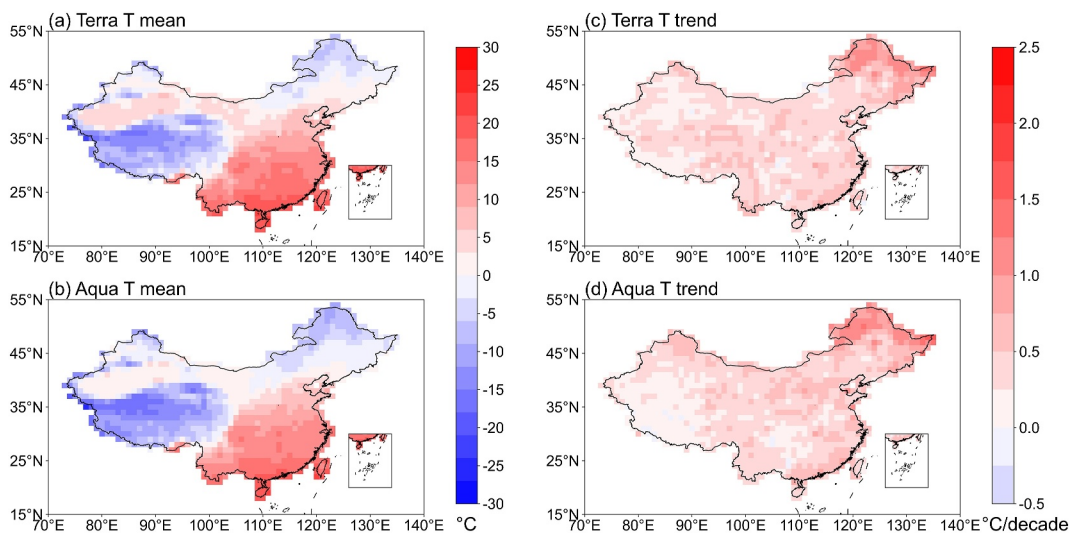

Figure 3. Multiyear average of MODIS land surface temperature (1° × 1° grid scale) at approximately nighttime (10:30 PM from Terra, and 1:30 am from Aqua) from 2001 to 2021 (left) and their trend from 2001 to 2021 (right). For Aqua, time series are from 2003 to 2021.

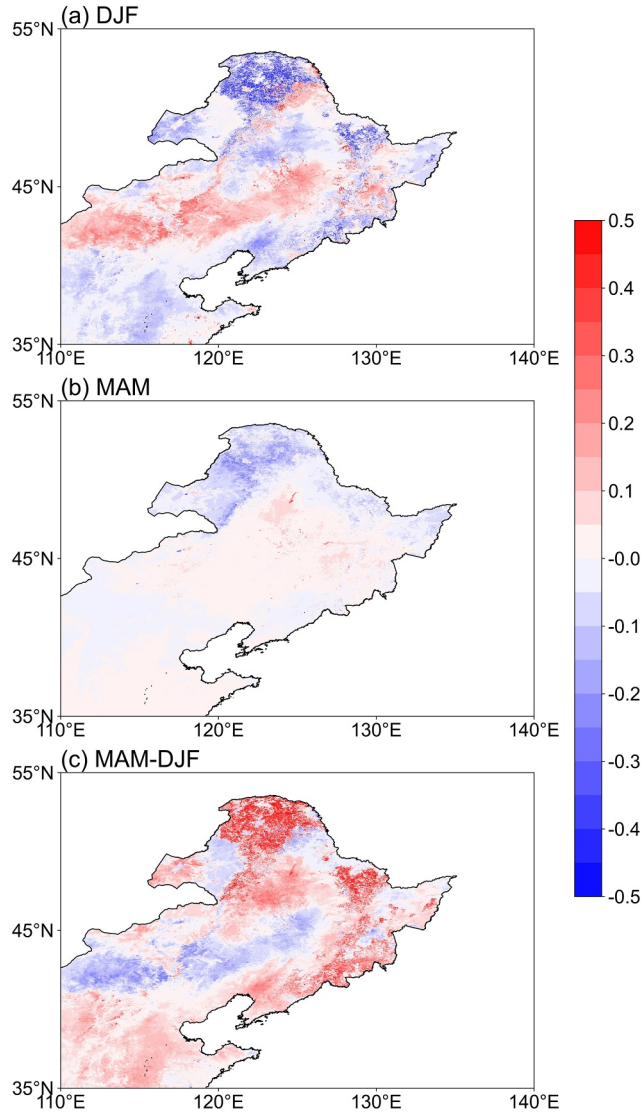

Figure 4. Multi-year trend of MODIS land surface albedo (time series from 2001 to 2021) in winter (a) and spring (b). The difference of albedo trend in winter and spring shows in (c).

3.3. Sampling Error of Mean and Trend of National Mean Temperature

The time series of three kinds of national mean temperatures from 2001 (for Aqua 2003) to 2021 shows in Figure 7, including mean values that contain spatial representative error (brown lines), under-sampling error (green lines) and true value (gray lines). The true value is smallest, with a Terra mean value of 1.69°C (for Aqua 0.42°C), for many grids in the Qinghai-Tibet Plateau have no station observation. The under-sampling error dominantly contributed to the mean value of temperature for 2.61°C (Terra) and 2.72°C (Aqua), and spatial representative error contributed for 1.09°C (Terra) and 0.92°C (Aqua). The total sampling error reaches 3.70°C (Terra) and 3.64°C (Aqua). Three kinds of sampling error for the whole year and all four seasons are shown in Table 1. Spatial representative error of temperature mean values (T_{ME1}) shows a larger amplitude of variation in seasons than under-sampling error of temperature mean values (T_{ME2}). There was little variation between seasons for T_{ME1} and T_{ME2}.

The temperature (T_{SS}) trend based on satellite land surface temperature retrieval is very similar to previous studies of observational minimum temperature trend during 1980 and 2015, 0.43°C/decade for Terra versus 0.40°C/decade based on air temperature (Shen & Zhao, 2021). T_{SS} shows most rapidly warming in three kinds of national

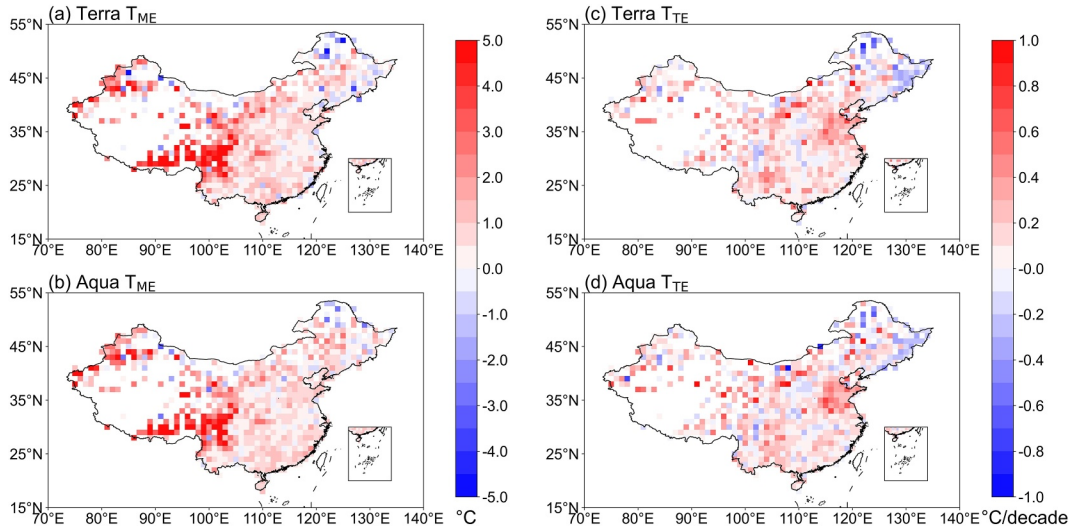

Figure 5. The spatial representative error in mean (left) and trend (right) of temperature from 2001 to 2021 (For Aqua, from 2003 to 2021). The temperature mean sampling error shows the difference of two kinds of temperature grid mean values over the time series, and is equal to mean value of pixels located at weather stations minus the mean value of all pixels in a grid. The temperature trend sampling error shows the difference in two kinds of temperature grid mean value's trend over the time series.

mean temperature with a spatial representative error of 0.09°C/decade for Terra (0.09°C/decade for Aqua), and an under-sampling error of 0.02°C/decade for Terra and Aqua. Three kinds of sampling relative error across whole year and in four seasons shows in Table 2.

While the under-sampling error of the trend over the world was emphasized by Cowtan and Way (2014), we found that in China, the spatial representative error contributed much more. The relatively spatial representative error of temperature trend (T_{TRE1}) of whole year is 26.1% (Terra) and 25.4% (Aqua), approximately five times as the relatively under-sampling error of temperature mean T_{TRE2}. The T_{TRE1} shows little variation between seasons. T_{TRE2} varies more from season to season, and reaches the same magnitude as T_{TRE1} in winter. Satellite data represents more inaccurate in winter because of snow cover, the result of winter leaves higher uncertainty (Wang et al., 2004; Westermann et al., 2012).

3.4. Determining Factor of Sampling Error

The spatial representative error caused by different elevation of station and grid is the dominant determining factor of the mean sampling error of temperature, and shows an opposite spatial pattern to that of temperature (See Figures 3a and 3b; Figure 8a), especially at the Hengduan mountains and southeast Tibet. The relationship of mean value sampling error is shown in Figure 8b. The regression coefficients of Figure 8c are −4.51°C/km (Terra, $p < 0.01$) and −4.19°C/km (Aqua, $p < 0.01$), slightly lower than the temperature lapse rate, as shown in He and Wang (2020). And the results support previous studies in error of mean value of temperature. Mean differences in surface air temperature between the reanalyzes and the in-situ observations have been attributed to the differences in elevation between station and grid (Du et al., 2018; Ma et al., 2008; Zhou et al., 2018). Wang et al. (2015) calibrated elevation error between reanalysis and observation in arid northwestern China, and the slopes of linear regression on annual basis are −4.8°C/km and −5.6°C/km for NCEP R2 and ERA-Interim.

The rapid urbanization process around stations affects the trend difference of different temperature series (S. Jiang et al., 2020; Wen et al., 2019). Figure 8c shows the trend error of the urbanization ratio. The weather stations have undergone faster urbanization process than the 1° × 1° grid. Except for the northeast region affected by melting snow in spring, the trend error of the urbanization ratio shows same pattern as the trend error of temperature (See Figures 3c and 3d), especially in the middle region of Shaanxi and Sichuan. The relationship of trend sampling error is showed in Figure 8d. The regression coefficients of Figure 8d are 0.85°C/decade (Terra, $p < 0.01$) and 0.80°C/decade (Aqua, $p < 0.01$).

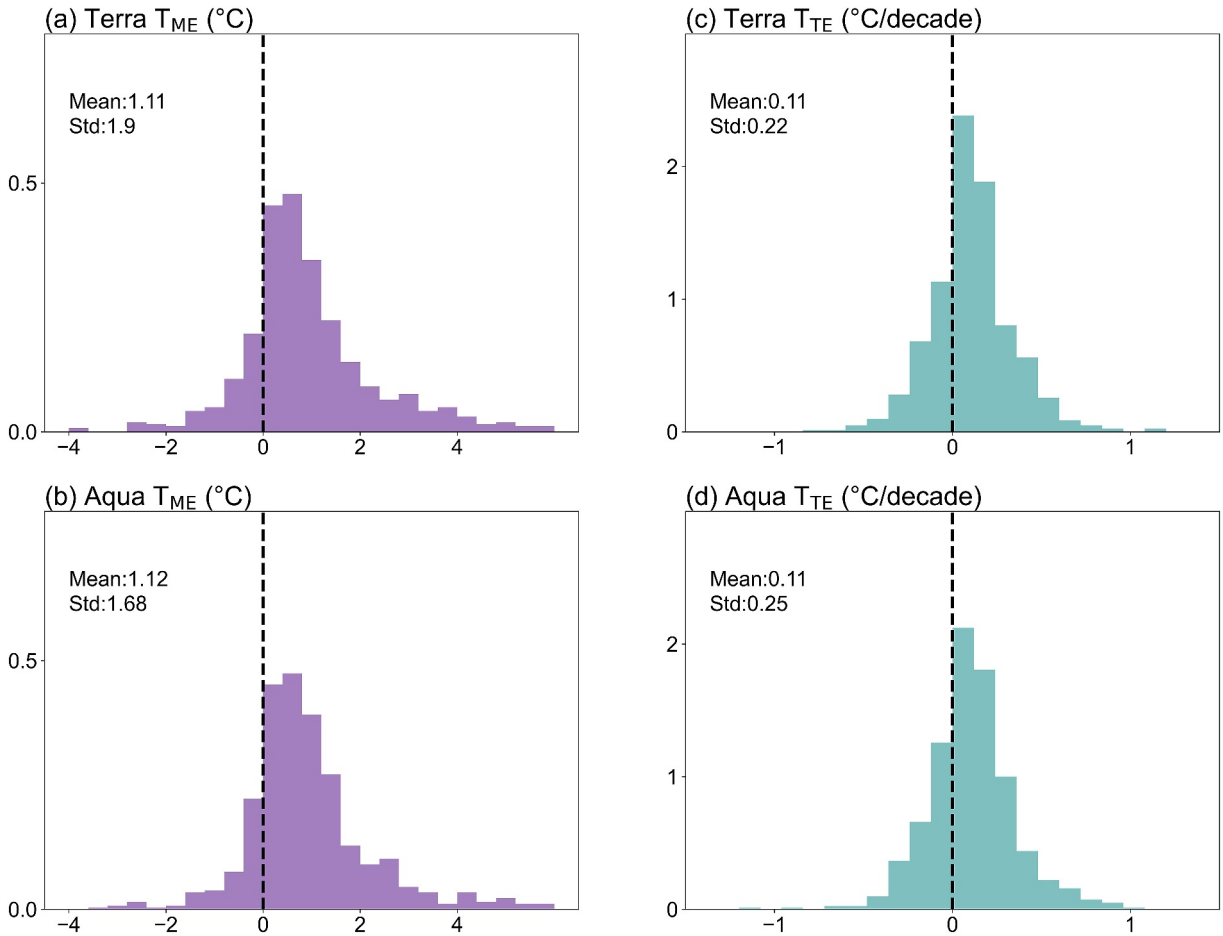

Figure 6. Histogram of spatial representative error in mean (left) and trend (right) of temperature from 2001 to 2021 (For Aqua, from 2003 to 2021). The histogram bars base on 1° × 1° grid data sets of sampling error in whole China.

Our relatively spatial representative error was similar to the urbanization impact contribution to the national warming trend by Jiang et al. (2020) using urban–rural station pairs and by Wang and Ge (2012) dividing stations into three urbanization categories, but smaller than that by Sun et al. (2016) using fingerprint method with climate models. Similarity with observational facts supports the reliability using land surface temperature as the surface temperature to calculate sampling error.

4. Discussion

4.1. Sensitivity Test of Size of Lon-Lat Grid Box

Since global conventional temperature data sets recommend a lon-lat grid box size of 5° × 5°, faithful quantification of under-sampling error was figured out at this scale (Cowtan & Way, 2014; Kadow et al., 2020). Besides, smaller grid box than 1°was also used in gridding procedure of observational temperature (Shen & Zhao, 2021). It is essential to downsize and upsize the grid box as sensitivity test of sampling error.

We change the size of grid boxes to 0.5°, 2° and 5° (Figure 9, as examples), the value of both under-sampling, spatial representative error and their sum exhibits minimal variation with grid size. The proportion of under-sampling and spatial representative error remains nearly constant over the grid sizes. This result meets the expected requirements essentially. Generally, the proportion of sampling error will change when the study area maintains similar area size with largest grid box sizes. This test demonstrated the validity of calculating sampling error in China.

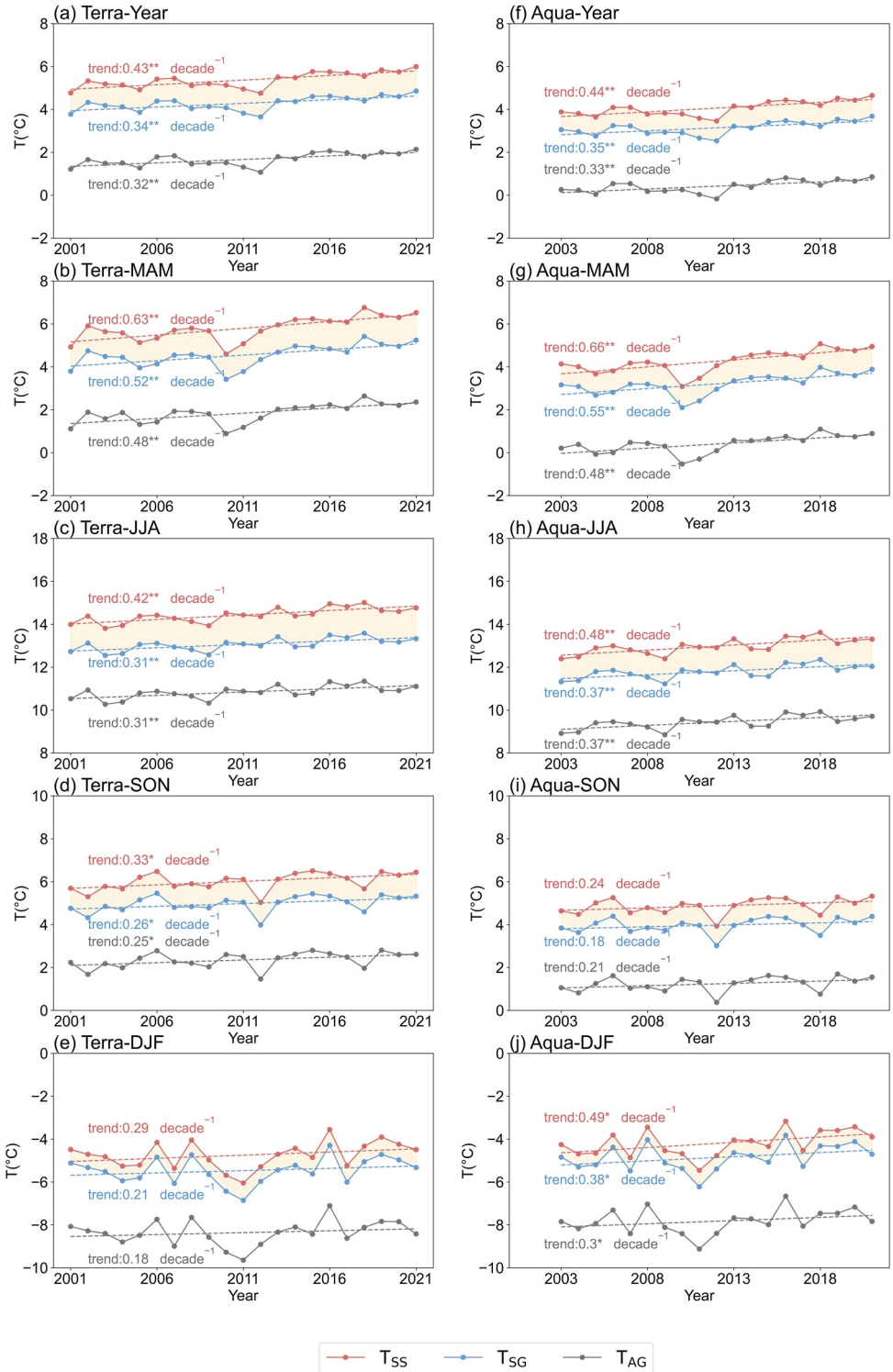

Figure 7. National annual and seasonal time series of the three kinds of the national mean temperature. The red lines (T_{SS}) show the time series of temperature spatial mean value of station sampling, that is, mean value of grids value calculated as the mean value of points locating stations, and grids without stations were dropped. The blue lines (T_{SG}) show the time series of temperature spatial the mean value of all pixels in the grids, and grids without stations were dropped. The dark gray lines (T_{AG}) show the time series of the temperature spatial the mean of all pixels in the grids, and no grid was dropped. Double asterisk indicates the results of trend with a significance level of 0.01, while single asterisk with a significance level of 0.05.

4.2. Feasibility of Using LST as Proxy for Air Temperature

We elaborated physical process of land-atmosphere temperature coupling in introduction. It should be noted that the scarce hourly temperature records restricted researches at satellite's overpass time. Krishnan et al. (2015) demonstrated the Tower-based air temperature explained 98% of the variance in LST during nighttime. More studies focused on the minimum air temperature, which is crucial for extreme temperature and warming research. Still, the similarity of LST and minimum air temperature were discussed in previous researches. Li et al. (2018) depicted the similar seasonal trend and daily variations between Aqua LST and Global Historical Climatology Network-Daily (GHCND) observational minimum air temperature. For urbanization issues, similar amplitude of urban island temperature at the surface and near the surface were found (Anniballe et al., 2014; Sun et al., 2015). For the weather stations we discussed in this work, the temporal series of both MODIS LST and observational minimum air temperature were depicted in Figure 10. The correlation coefficient between MODIS LST anomaly and minimum air temperature anomaly is 0.927 (Terra) and 0.925 (Aqua).

4.3. Uncertainty of Nighttime Temperature Assumption

Our assumption of nighttime temperature based on the strong coupling of air temperature and LST. At nighttime, the surface and air temperatures gradually decrease due to cooling effect of negative surface longwave radiation. And the similarity of the temperature values was discussed in 4.2, proved the feasibility of the strong coupling assumption. Under this assumption, both air temperature and LST exhibit negligible temporal fluctuations. Then the difference in sampling error of Terra and Aqua can be considered as the uncertainty of the method.

The biases of MODIS LST products would also introduce uncertainties in this study, including sensor calibration, geolocation error, and cloud masking and other issues such as atmospheric water vapor absorption and surface emissivity issues (Duan et al., 2019; K. Wang & Liang, 2009). Cloud contamination caused large number of invalid values, and the erroneous rates of MODIS nighttime cloud detection are obviously higher than those achieved in daytime (Long et al., 2020; Tan et al., 2021; Zeng et al., 2018; Østby et al., 2014). In our method, we average the MODIS LST images to monthly mean value in advance, to avoid the inconsistent contribution of monthly temperature to the annual average. Moreover, the study based on the comparison between point (pixel) and its surrounding pixels, therefore the conclusions derived are not very sensitive to absolute accuracy of MODIS LST. However, the comparison between two independent data sets (such as air temperature measured at weather station) and (MODIS LST) is highly dependent on their own absolute accuracy. This is the reason that we only use the MODIS LST in this study. The uncertainties from satellite imaging in subsequent sampling error is considered to be smaller than those in MODIS LST. However, further attention and discussion in biases of MODIS products are needed.

While the uncertainty of assumption in method is demonstrated, sampling error of mean and trend of national mean temperature can be simplified to the average of that from Terra and Aqua. Recent researches proved the feasibility of combining LST data from Terra and Aqua to derive nighttime LST with fine spatiotemporal resolution (Shiff et al., 2021; Yu et al., 2022; Zhang et al., 2022). Then the sampling error of national mean value and trend with uncertainty is shown in Tables 3 and 4. National mean sampling error shows minor uncertainty in both mean value and trend.

Similarly, the dominant determining factor of sampling error of temperature can be expressed as regression coefficient between the averaged MODIS LST and the sampling error of elevation and urbanization ratio from 2003 to 2021. The regression coefficient of sampling error of elevation and temperature mean value is -4.37 ± 0.36°C/km ($p < 0.01$). The regression coefficient of trend sampling error of urbanization ratio and temperature is 0.88 ± 0.17°C/decade ($p < 0.01$).

4.4. Elevation Based Density Requirements of Weather Stations

Elevation is the decisive static factor of sampling error in temperature mean value. As it hardly changes over the time, distribution of elevations in a grid could provide the information about locations of weather stations to reduce sampling error. Generally, the number of stations required increases as the distribution of DEM exhibits a higher degree of dispersion. We investigated the frequency density of each $1° \times 1°$ grid box in China. The interval of each histogram was set to 200 m, at size which elevation difference could introduce a sampling error of ~1°C. The histogram bin with largest area was considered to be the most valuable location to set a weather station, and

Table 1
The National Mean Value Sampling Error: Spatial Representative Error (T_{ME1}), Under-Sampling Error (T_{ME2}) and all Sampling Error (T_{ME3}) of Temperature

		T_{ME1} (°C)	T_{ME2} (°C)	T_{ME3} (°C)
Terra	Year	1.09	2.61	3.70
	MAM	1.26	2.71	3.96
	JJA	1.37	2.22	3.59
	SON	1.04	2.63	3.67
	DJF	0.72	2.89	3.61
Aqua	Year	0.92	2.72	3.64
	MAM	1.06	2.80	3.86
	JJA	1.19	2.36	3.55
	SON	0.91	2.73	3.64
	DJF	0.66	2.97	3.63

the histogram area was counted as the spatial representativeness of elevation. As the area under the histogram sums up to 1, we choose 0.8 as the threshold of cumulative area. That is, we cumulatively sum these histogram areas in descending order until the sum reaches 0.8, and count the number of bins as required number of stations. We have taken a grid box at Tibet Plateau as example (the gray box in Figure 11c). Figure 11a shows the histogram of the elevations in the grid box, and the rank of bins by area under the histogram is shown. Figure 11b shows the ranked area of each bins, and the cumulative values. And the same calculation method was applied in the whole country.

Figure 11c shows the minimum amount of weather station in a grid to keep 80% of elevations are represented. The North China Plain, Northeast China Plain and middle and low region of Yangtze River of China requires only one or two weather station in a $1° \times 1°$ grid box, while the regions near the coast and at mountains in Tibet Plateau are expected to maintain more stations. This simple calculation only based on data set of elevation. It provides a simple but static suggestion of weather stations' development.

5. Conclusions

In this study, the temperature sampling error of meteorological stations was investigated based on MODIS LST data set. The total sampling error of temperature mean reaches 3.67 ± 0.07°C, with sampling error of under-sampling of 2.67 ± 0.11°C, and sampling error of station location of 1.01 ± 0.17°C. The sampling error of

Table 2
The Relative National Trend Sampling Error: Spatial Representative Sampling Error (T_{TRE1}), Under-Sampling Error (T_{TRE1}) and all Sampling Error (T_{TRE1}) of Temperature

		T_{TRE1} (%)	T_{TRE2} (%)	T_{TRE3} (%)
Terra	Year	26.1	5.2	32.7
	MAM	21.9	7.0	30.4
	JJA	32.8	3.0	36.7
	SON	28.1	0.4	28.6
	DJF	38.0	19.7	65.2
Aqua	Year	25.4	6.9	34.1
	MAM	19.7	14.4	36.9
	JJA	28.7	0.0	28.7
	SON	28.3	−12.8	11.8
	DJF	28.1	29.3	65.7

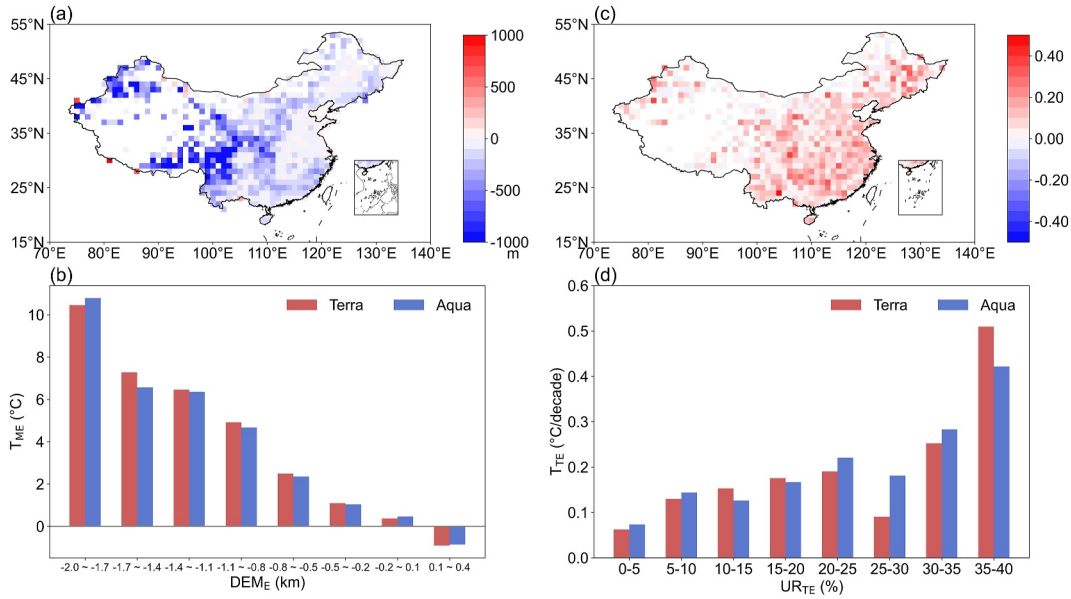

Figure 8. (a) The sampling error of the DEM for each grid, equals the mean value of pixels located at weather stations minus the mean value of all pixels in the grid. (b) The mean sampling error of temperature under different sampling error of the DEM. (c) The Trend sampling error of urbanization ratio, for each grid, equals the trend of series which mean value of pixels located at weather stations minus the trend of series which mean value of all pixels in the grid. (d) The trend sampling error of temperature under different trend sampling error of urbanization ratio.

temperature mean value barely varies between seasons. For the first time, we have subdivided the sampling error for under-sampling and spatial representation of the mean value of temperature in China in the past 20 years. For representative problems, it requires attention when using temperature data sets that have not undergone elevation correction.

The total relatively sampling error of trend reaches 32.9% ± 2.7%, with under-sampling error of 6.9% ± 0.2% with greater variation between seasons, and spatial representative error of 24.3% ± 2.3% with little variation between seasons. Notably, our spatial representative error is an integrated result of 2,422 stations, expanding the number of station samples limited by station pairing. The spatial representative error in temperature is approximately five times as large as the under-sampling error. For China's warming trend, the spatial representative error is assignable.

The sampling error of the station elevation is an important reason for the sampling error of the mean temperature, with regression coefficients of −4.37 ± 0.36°C/km. While the urbanization around the station is an important reason for the sampling error of the temperature trend, with regression coefficients of 0.85°C/decade (Terra) and 0.80°C/decade (Aqua). These results were not obtained by early studies, showing that rapid local urbanization around weather stations would impact the sampling error of temperature or other meteorological factors to a great extent.

In summary, this study proposed the spatial representative error by station sampling, and first calculated it with the under-sampling error in China. 1° × 1° grid boxes were used in the study. And it was demonstrated that the value of sampling error hardly varies with grid size. Sampling error in both the temperature mean and trend were analyzed, and considerable positive error was represented. We first proposed that the spatial representative error in the temperature trend was much greater than the under-sampling error. The elevation and impervious area data sets were used and successfully explained the sampling error in the mean value and trend of temperature.

More work can be done in the future to explore the sampling error of weather stations. The conclusions derived here are limited to nighttime because changes in air temperature and land surface temperature are controlled by different mechanisms and have different representative scales. The sampling error during the daytime are expected to be different and need further study. Other data set with sufficient resolution, such as near surface air temperature data produced by satellites or high-resolution regional model can be applied to calculate the sampling

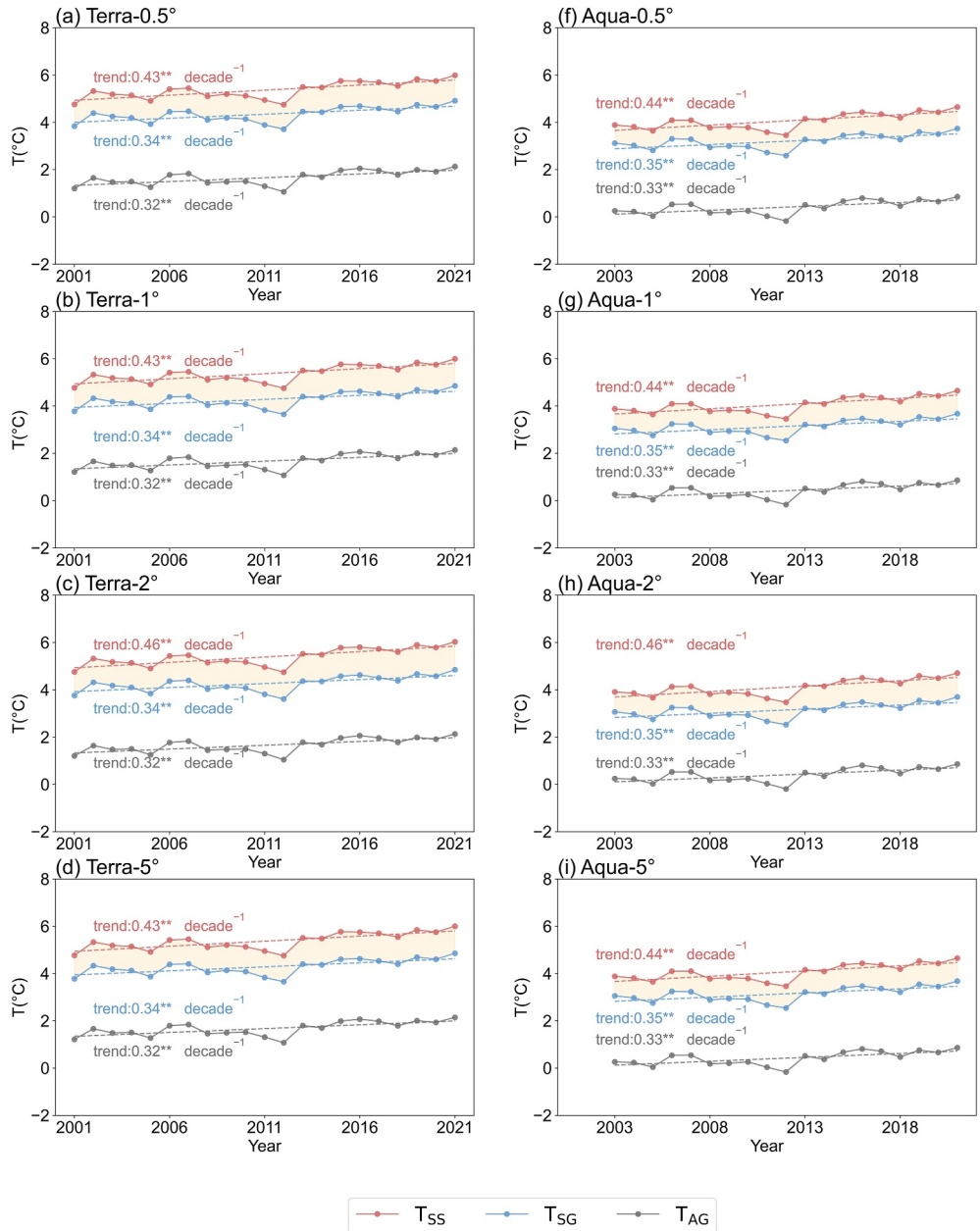

Figure 9. National annual and seasonal time series of the three kinds of the national mean temperature with different grid size. The red lines (T_{SS}) show the time series of temperature spatial mean value of station sampling, that is, mean value of grids value calculated as the mean value of points locating stations, and grids without stations were dropped. The blue lines (T_{SG}) show the time series of temperature spatial the mean value of all pixels in the grids, and grids without stations were dropped. The dark gray lines (T_{AG}) show the time series of the temperature spatial the mean of all pixels in the grids, and no grid was dropped. Double asterisk indicates the results of trend with a significance level of 0.01, while single asterisk with a significance level of 0.05.

error (Chen et al., 2016; Long et al., 2020; Pu & Bonafoni, 2023; Sekertekin & Bonafoni, 2020; Yin & Sun, 2023; Zhang et al., 2022). In particular regions, other effects of sampling error, for example, snow cover and vegetation growth, still need to be addressed (Piao et al., 2015; Yao et al., 2022). We also expect the long-term data of automatic meteorological networks to provide a more substantial data set of near-surface air temperature. The location of automatic stations could help researchers to predict future sampling error using downscaled data from future scenarios.

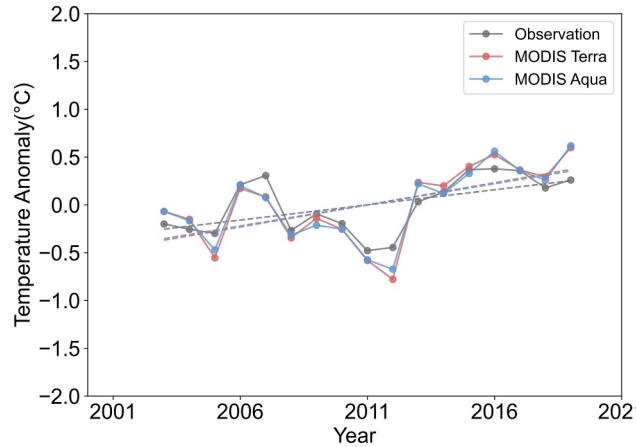

Figure 10. Arithmetic mean of ~2,400 metrological stations' temperature anomalies over China. The gray line bases on observational minimum near surface air temperature. The orange line bases on MODIS Terra land surface temperature. The green line bases on MODIS Aqua land surface temperature.

Table 3
The National Mean Value Sampling Error With Uncertainty: Spatial Representative Error (T_{ME1}), Under-Sampling Error (T_{ME2}) and all Sampling Error (T_{ME3}) of Temperature

	T_{ME1} (°C)	T_{ME2} (°C)	T_{ME3} (°C)
Year	1.01 ± 0.17	2.67 ± 0.11	3.67 ± 0.07
MAM	1.16 ± 0.21	2.75 ± 0.10	3.91 ± 0.10
JJA	1.28 ± 0.19	2.29 ± 0.14	3.58 ± 0.06
SON	0.98 ± 0.14	2.68 ± 0.10	3.66 ± 0.04
DJF	0.70 ± 0.07	2.93 ± 0.08	3.62 ± 0.01

Note. The value of sampling error is calculated by averaging sampling error from Terra and Aqua, and the uncertainty is the difference of that from Terra and Aqua from 2003 to 2021.

Table 4
The Relative National Trend Sampling Error With Uncertainty: Spatial Representative Sampling Error (T_{TRE1}), Under-Sampling Error (T_{TRE1}) and all Sampling Error (T_{TRE1}) of Temperature

	T_{TRE1} (%)	T_{TRE2} (%)	T_{TRE3} (%)
Year	24.3 ± 2.3	6.9 ± 0.2	32.9 ± 2.7
MAM	19.5 ± 0.1	13.5 ± 1.5	35.7 ± 1.9
JJA	27.8 ± 1.6	0.2 ± 0.1	28.2 ± 1.5
SON	31.6 ± 6.2	−10.1 ± 5.9	18.3 ± 13.3
DJF	23.6 ± 8.7	25.7 ± 6.8	55.4 ± 19.4

Note. The value of sampling error is calculated by averaging sampling error from Terra and Aqua, and the uncertainty is the difference of that from Terra and Aqua from 2003 to 2021.

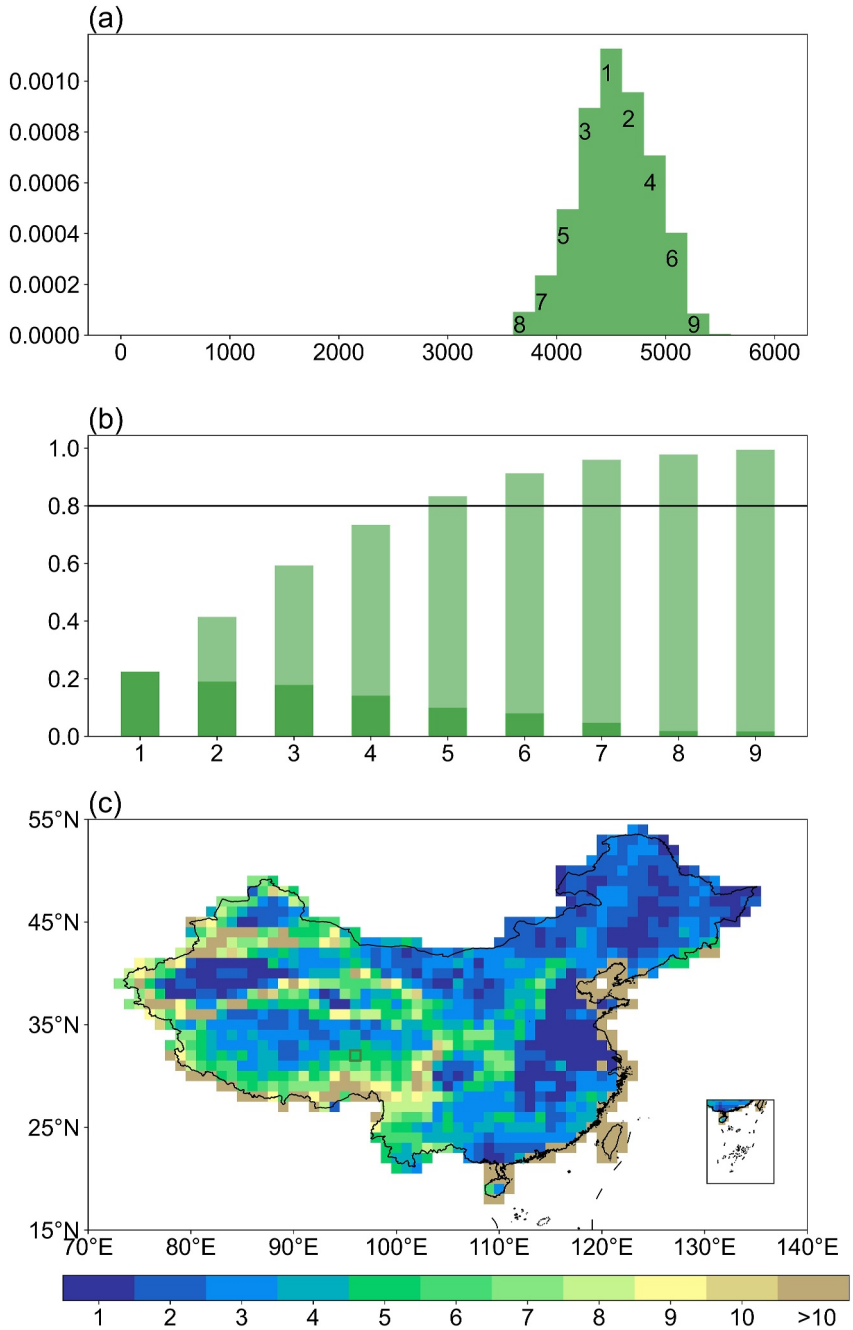

Figure 11. (a) The histogram of the elevations in an example grid box at Tibet Plateau. (b) The ranked area of each bins (deeper green), and the cumulative values (lighter green) of the example grid box. (c)The minimum amount of weather station in a grid to keep 80% of -elevations are represented.

Data Availability Statement

All land surface temperature data used during this study are openly available from NASA EOSDIS Land Processes Distributed Active Archive Center at Wan et al. (2015). The daily air temperature and location information (longitude and latitude) of weather stations were obtained from the China Meteorological Data Service Center (CMDC) in 2018. The high-resolution (30 m) data set of global artificial impervious area (GAIA) are included at (Gong et al., 2020). The SRTM digital elevation data are openly available from CGIAR-CSI at (Jarvis et al., 2008).

Acknowledgments
This work was supported by the National Key Research and Development Program of China (2022YFF0801302) and the National Natural Science Foundation of China (41930970). We thank the CMA, the NASA and the CGIAR for providing the data sets.

References

Anniballe, R., Bonafoni, S., & Pichierri, M. (2014). Spatial and temporal trends of the surface and air heat island over Milan using MODIS data. *Remote Sensing of Environment*, *150*, 163–171. https://doi.org/10.1016/j.rse.2014.05.005

Bhowmik, A. K., & Costa, A. C. (2015). Representativeness impacts on accuracy and precision of climate spatial interpolation in data-scarce regions. *Meteorological Applications*, *22*(3), 368–377. https://doi.org/10.1002/met.1463

Chen, Y., Sun, H., & Li, J. (2016). Estimating daily maximum air temperature with MODIS data and a daytime temperature variation model in Beijing urban area. *Remote Sensing Letters*, *7*(9), 865–874. https://doi.org/10.1080/2150704X.2016.1193792

Cowtan, K., & Way, R. (2014). Coverage bias in the HADCRUT4 temperature series and its impact on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1935–1944. https://doi.org/10.1002/qj.2297

Du, J., Wang, K., Wang, J., Jiang, S., & Zhou, C. (2018). Diurnal cycle of surface air temperature within China in current reanalyzes: Evaluation and diagnostics. *Journal of Climate*, *31*(11), 4585–4603. https://doi.org/10.1175/JCLI-D-17-0773.1

Duan, S., Li, Z., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., et al. (2019). Validation of Collection 6 MODIS land surface temperature product using in situ measurements. *Remote Sensing of Environment*, *225*, 16–29. https://doi.org/10.1016/j.rse.2019.02.020

Ge, Q., Wang, F., & Luterbacher, J. (2013). Improved estimation of average warming trend of China from 1951–2010 based on satellite observed land-use data. *Climatic Change*, *121*(2), 365–379. https://doi.org/10.1007/s10584-013-0867-4

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., et al. (2020). Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sensing of Environment*, *236*, 111510. https://doi.org/10.1016/j.rse.2019.111510

Good, E. J. (2016). An in situ-based analysis of the relationship between land surface "skin" and screen-level air temperatures. *Journal of Geophysical Research: Atmospheres*, *121*(15), 8801–8819. https://doi.org/10.1002/2016JD025318

He, Y., & Wang, K. (2020). Contrast patterns and trends of lapse rates calculated from near-surface air and land surface temperatures in China from 1961 to 2014. *Science Bulletin*, *65*(14), 1217–1224. https://doi.org/10.1016/j.scib.2020.04.001

Hu, Z., Zhang, C., Hu, Q., & Tian, H. (2014). Temperature changes in central Asia from 1979 to 2011 based on multiple datasets. *Journal of Climate*, *27*(3), 1143–1167. https://doi.org/10.1175/JCLI-D-13-00064.1

Huang, J., Zhang, X., Zhang, Q., Lin, Y., Hao, M., Luo, Y., et al. (2017). Recently amplified arctic warming has contributed to a continual global warming trend. *Nature Climate Change*, *7*(12), 875–879. https://doi.org/10.1038/s41558-017-0009-5

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe version 4, available from the CGIAR-CSI SRTM 90m Database. http://srtm.csi.cgiar.org

Jiang, L., Yang, J., Zhang, C., Wu, S., Li, Z., Dai, L., et al. (2022). Daily snow water equivalent product with SMMR, SSM/I and SSMIS from 1980 to 2020 over China. *Big Earth Data*, *6*(4), 420–434. https://doi.org/10.1080/20964471.2022.2032998

Jiang, S., Wang, K., & Mao, Y. (2020). Rapid local urbanization around most meteorological stations explains the observed daily asymmetric warming rates across China from 1985 to 2017. *Journal of Climate*, *33*(20), 9045–9061. https://doi.org/10.1175/JCLI-D-20-0118.1

Jones, P. D., & Moberg, A. (2003). Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *Journal of Climate*, *16*(2), 206–223. https://doi.org/10.1175/1520-0442(2003)016<0206:HALSSA>2.0.CO;2

Kadow, C., Hall, D. M., & Ulbrich, U. (2020). Artificial intelligence reconstructs missing climate information. *Nature Geoscience*, *13*(6), 408–413. https://doi.org/10.1038/s41561-020-0582-5

Karl, T. R., Diaz, H. F., & Kukla, G. (1988). Urbanization: Its detection and effect in the United States climate record. *Journal of Climate*, *1*(11), 1099–1123. https://doi.org/10.1175/1520-0442(1988)001<1099:UIDAEI>2.0.CO;2

Krishnan, P., Kochendorfer, J., Dumas, E. J., Guillevic, P. C., Bruce Baker, C., Meyers, T. P., & Martos, B. (2015). Comparison of in-situ, aircraft, and satellite land surface temperature measurements over a NOAA Climate Reference Network site. *Remote Sensing of Environment*, *165*, 249–264. https://doi.org/10.1016/j.rse.2015.05.011

Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., et al. (2021). Temperature-based and radiance-based validation of the collection 6 MYD11 and MYD21 land surface temperature products over barren surfaces in northwestern China. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(2), 1794–1807. https://doi.org/10.1109/TGRS.2020.2998945

Li, Q., Dong, W., & Jones, P. (2020). Continental scale surface air temperature variations: Experience derived from the Chinese region. *Earth-Science Reviews*, *200*, 102998. https://doi.org/10.1016/j.earscirev.2019.102998

Li, Q., Sheng, B., Huang, J., Li, C., Song, Z., Chao, L., et al. (2022a). Different climate response persistence causes warming trend unevenness at continental scales. *Nature Climate Change*, *12*(4), 343–349. https://doi.org/10.1038/s41558-022-01313-9

Li, X., Long, D., Scanlon, B. R., Mann, M. E., Li, X., Tian, F., et al. (2022b). Climate change threatens terrestrial water storage over the Tibetan Plateau. *Nature Climate Change*, *12*(9), 801–807. https://doi.org/10.1038/s41558-022-01443-0

Li, X., Zhou, Y., Asrar, G. R., & Zhu, Z. (2018). Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous United States. *Remote Sensing of Environment*, *206*, 84–97. https://doi.org/10.1016/j.rse.2017.12.010

Liu, H., Mi, Z., Lin, L., Wang, Y., Zhang, Z., Zhang, F., et al. (2018a). Shifting plant species composition in response to climate change stabilizes grassland primary production. *Proceedings of the National Academy of Sciences*, *115*(16), 4051–4056. https://doi.org/10.1073/pnas.1700299114

Liu, S., Li, X., Xu, Z., Che, T., Xiao, Q., Ma, M., et al. (2018b). The Heihe integrated observatory network: A basin-scale land surface processes observatory in China. *Vadose Zone Journal*, *17*(1), 1–21. https://doi.org/10.2136/vzj2018.04.0072

Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., et al. (2020). Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach. *Remote Sensing of Environment*, *246*, 111863. https://doi.org/10.1016/j.rse.2020.111863

Lovejoy, S. (2017). How accurately do we know the temperature of the surface of the earth? *Climate Dynamics*, *49*(11), 4089–4106. https://doi.org/10.1007/s00382-017-3561-9

Ma, L., Zhang, T., Li, Q., Frauenfeld, O. W., & Qin, D. (2008). Evaluation of ERA-40, NCEP-1, and NCEP-2 reanalysis air temperatures with ground-based measurements in China. *Journal of Geophysical Research*, *113*(D15), D15115. https://doi.org/10.1029/2007JD009549

Ma, S., Zhou, T., Stone, D. A., Angélil, O., & Shiogama, H. (2017). Attribution of the July–August 2013 heat event in Central and Eastern China to anthropogenic greenhouse gas emissions. *Environmental Research Letters*, *12*(5), 054020. https://doi.org/10.1088/1748-9326/aa69d2

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., & Lawrimore, J. H. (2018). The global historical Climatology network monthly temperature dataset, version 4. *Journal of Climate*, *31*(24), 9835–9854. https://doi.org/10.1175/JCLI-D-18-0094.1

Mira, M., Ninyerola, M., Batalla, M., Pesquer, L., & Pons, X. (2017). Improving mean minimum and maximum month-to-month air temperature surfaces using satellite-derived land surface temperature. *Remote Sensing*, *9*(12), 1313. https://doi.org/10.3390/rs9121313

Orlowsky, B., & Seneviratne, S. I. (2014). On the spatial representativeness of temporal dynamics at European weather stations. *International Journal of Climatology*, *34*(10), 3154–3160. https://doi.org/10.1002/joc.3903

Osborn, T. J., Jones, P. D., Lister, D. H., Morice, C. P., Simpson, I. R., Winn, J. P., et al. (2021). Land surface air temperature variations across the globe updated to 2019: The CRUTEM5 data set. *Journal of Geophysical Research: Atmospheres*, *126*(2), e2019JD032352. https://doi.org/10.1029/2019JD032352

Østby, T. I., Schuler, T. V., & Westermann, S. (2014). Severe cloud contamination of MODIS land surface temperatures over an Arctic ice cap, Svalbard. *Remote Sensing of Environment*, *142*, 95–102. https://doi.org/10.1016/j.rse.2013.11.005

Panwar, A., Kleidon, A., & Renner, M. (2019). Do surface and air temperatures contain similar imprints of evaporative conditions? *Geophysical Research Letters*, *46*(7), 3802–3809. https://doi.org/10.1029/2019GL082248

Piao, S., Yin, G., Tan, J., Cheng, L., Huang, M., Li, Y., et al. (2015). Detection and attribution of vegetation greening trend in China over the last 30 years. *Global Change Biology*, *21*(4), 1601–1609. https://doi.org/10.1111/gcb.12795

Pu, R., & Bonafoni, S. (2023). Thermal infrared remote sensing data downscaling investigations: An overview on current status and perspectives. *Remote Sensing Applications: Society and Environment*, *29*, 100921. https://doi.org/10.1016/j.rsase.2023.100921

Ren, G., Zhou, Y., Chu, Z., Zhou, J., Zhang, A., Guo, J., & Liu, X. (2008). Urbanization effects on observed surface air temperature trends in north China. *Journal of Climate*, *21*(6), 1333–1348. https://doi.org/10.1175/2007JCLI1348.1

Ren, Q., He, C., Huang, Q., Shi, P., Zhang, D., & Güneralp, B. (2022). Impacts of urban expansion on natural habitats in global dryland. *Nature Sustainability*, *5*(10), 869–878. https://doi.org/10.1038/s41893-022-00930-8

Richardson, M., Cowtan, K., & Millar, R. J. (2018). Global temperature definition affects achievement of long-term climate goals. *Environmental Research Letters*, *13*(5), 054004. https://doi.org/10.1088/1748-9326/aab305

Rohde, R. A., & Hausfather, Z. (2020). The Berkeley earth Land/Ocean Temperature record. *Earth System Science Data*, *12*(4), 3469–3479. https://doi.org/10.5194/essd-12-3469-2020

Sekertekin, A., & Bonafoni, S. (2020). Sensitivity analysis and validation of daytime and nighttime land surface temperature retrievals from Landsat 8 using different algorithms and emissivity models. *Remote Sensing*, *12*(17), 2776. https://doi.org/10.3390/rs12172776

Shen, P., & Zhao, S. (2021). 1/4 to 1/3 of observed warming trends in China from 1980 to 2015 are attributed to land use changes. *Climatic Change*, *164*(3–4), 59. https://doi.org/10.1007/s10584-021-03045-9

Shiff, S., Helman, D., & Lensky, I. M. (2021). Worldwide continuous gap-filled MODIS land surface temperature dataset. *Scientific Data*, *8*(1), 74. https://doi.org/10.1038/s41597-021-00861-7

Steeneveld, G. J., van de Wiel, B. J. H., & Holtslag, A. A. M. (2006). Modeling the evolution of the atmospheric boundary layer coupled to the land surface for three contrasting nights in CASES-99. *Journal of the Atmospheric Sciences*, *63*(3), 920–935. https://doi.org/10.1175/JAS3654.1

Sun, H., Chen, Y., & Zhan, W. (2015). Comparing surface- and canopy-layer urban heat islands over Beijing using MODIS data. *International Journal of Remote Sensing*, *36*(21), 5448–5465. https://doi.org/10.1080/01431161.2015.1101504

Sun, Y., Zhang, X., Ding, Y., Chen, D., Qin, D., & Zhai, P. (2022). Understanding human influence on climate change in China. *National Science Review*, *9*(3), nwab113. https://doi.org/10.1093/nsr/nwab113

Sun, Y., Zhang, X., Ren, G., Zwiers, F. W., & Hu, T. (2016). Contribution of urbanization to warming in China. *Nature Climate Change*, *6*(7), 706–709. https://doi.org/10.1038/nclimate2956

Tan, J., Che, T., Wang, J., Liang, J., Zhang, Y., & Ren, Z. (2021). Reconstruction of the daily MODIS land surface temperature product using the two-step improved similar pixels method. *Remote Sensing*, *13*(9), 1671. https://doi.org/10.3390/rs13091671

Tysa, S. K., Ren, G., Qin, Y., Zhang, P., Ren, Y., Jia, W., & Wen, K. (2019). Urbanization effect in regional temperature series based on a remote sensing classification scheme of stations. *Journal of Geophysical Research: Atmospheres*, *124*(20), 10646–10661. https://doi.org/10.1029/2019JD030948

Vancutsem, C., Ceccato, P., Dinku, T., & Connor, S. J. (2010). Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment*, *114*(2), 449–465. https://doi.org/10.1016/j.rse.2009.10.002

Wan, Z., Hook, S., & Hulley, G. (2015). MOD11A1 MODIS/Terra land surface temperature/emissivity daily L3 global 1km SIN grid V006 [Dataset]. *NASA EOSDIS Land Processes Distributed Active Archive Center*. https://doi.org/10.5067/MODIS/MOD11A1.006

Wang, F., & Ge, Q. (2012). Estimation of urbanization bias in observed surface temperature change in China from 1980 to 2009 using satellite land-use data. *Chinese Science Bulletin*, *57*(14), 1708–1715. https://doi.org/10.1007/s11434-012-4999-0

Wang, K., & Liang, S. (2009). Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sensing of Environment*, *113*(7), 1556–1565. https://doi.org/10.1016/j.rse.2009.03.009

Wang, K., Liu, J., Zhou, X., Sparrow, M., Ma, M., Sun, Z., & Jiang, W. (2004). Validation of MODIS global land surface albedo product using ground measurements in a semidesert region on the Tibetan Plateau. *Journal of Geophysical Research*, *109*(D5), D05107. https://doi.org/10.1029/2003JD004229

Wang, S., Zhang, M., Sun, M., Wang, B., Huang, X., Wang, Q., & Feng, F. (2015). Comparison of surface air temperature derived from NCEP/DOE R2, ERA-Interim, and observations in the arid northwestern China: A consideration of altitude errors. *Theoretical and Applied Climatology*, *119*(1), 99–111. https://doi.org/10.1007/s00704-014-1107-1

Wen, K., Ren, G., Li, J., Zhang, A., Ren, Y., Sun, X., & Zhou, Y. (2019). Recent surface air temperature change over Mainland China based on an urbanization-bias adjusted dataset. *Journal of Climate*, *32*(10), 2691–2705. https://doi.org/10.1175/JCLI-D-18-0395.1

Westermann, S., Langer, M., & Boike, J. (2012). Systematic bias of average winter-time land surface temperatures inferred from MODIS at a site on Svalbard, Norway. *Remote Sensing of Environment*, *118*, 162–167. https://doi.org/10.1016/j.rse.2011.10.025

Wu, J., & Gao, X. (2013). A gridded daily observation dataset over China region and comparison with the other datasets (in Chinese). *Chinese Journal of Geophysics- Chinese Edition*, *56*, 1102–1111. https://doi.org/10.6038/cjg20130406

Xu, Y., Gao, X., Shen, Y., Xu, C., Shi, Y., & Giorgi, F. (2009). A daily temperature dataset over China and its application in validating a RCM simulation. *Advances in Atmospheric Sciences*, *26*(4), 763–772. https://doi.org/10.1007/s00376-009-9029-z

Yan, Y., You, Q., Wu, F., Pepin, N., & Kang, S. (2020). Surface mean temperature from the observational stations and multiple reanalyses over the Tibetan Plateau. *Climate Dynamics*, *55*(9–10), 2405–2419. https://doi.org/10.1007/s00382-020-05386-0

Yang, J., Zhou, M., Ren, Z., Li, M., Wang, B., Liu, D. L., et al. (2021). Projecting heat-related excess mortality under climate change scenarios in China. *Nature Communications*, *12*(1), 1039. https://doi.org/10.1038/s41467-021-21305-1

Yang, K., Wu, H., Qin, J., Lin, C., Tang, W., & Chen, Y. (2014). Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global and Planetary Change*, *112*, 79–91. https://doi.org/10.1016/j.gloplacha.2013.12.001

Yang, X., Hou, Y., & Chen, B. (2011). Observed surface warming induced by urbanization in east China. *Journal of Geophysical Research*, *116*(D14), D14113. https://doi.org/10.1029/2010JD015452

Yao, J., Chen, Y., Guan, X., Zhao, Y., Chen, J., & Mao, W. (2022). Recent climate and hydrological changes in a mountain–basin system in Xinjiang, China. *Earth-Science Reviews*, *226*, 103957. https://doi.org/10.1016/j.earscirev.2022.103957

Yin, H., & Sun, Y. (2023). Anthropogenic influence on temperature change in China over the period 1901–2018. *Journal of Climate*, *36*(7), 2131–2146. https://doi.org/10.1175/JCLI-D-22-0122.1

Yu, P., Zhao, T., Shi, J., Ran, Y., Jia, L., Ji, D., & Xue, H. (2022). Global spatiotemporally continuous MODIS land surface temperature dataset. *Scientific Data*, *9*(1), 143. https://doi.org/10.1038/s41597-022-01214-8

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., & Hong, Y. (2018). A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. *ISPRS Journal of Photogrammetry and Remote Sensing*, *141*, 30–45. https://doi.org/10.1016/j.isprsjprs.2018.04.005

Zhang, G., Luo, W., Chen, W., & Zheng, G. (2019). A robust but variable lake expansion on the Tibetan Plateau. *Science Bulletin*, *64*(18), 1306–1309. https://doi.org/10.1016/j.scib.2019.07.018

Zhang, S.-Q., Ren, G.-Y., Ren, Y.-Y., Zhang, Y.-X., & Xue, X.-Y. (2021). Comprehensive evaluation of surface air temperature reanalysis over China against urbanization-bias-adjusted observations. *Advances in Climate Change Research*, *12*(6), 783–794. https://doi.org/10.1016/j.accre.2021.09.010

Zhang, T., Zhou, Y., Zhao, K., Zhu, Z., Chen, G., Hu, J., & Wang, L. (2022). A global dataset of daily maximum and minimum near-surface air temperature at 1 km resolution over land (2003–2020). *Earth System Science Data*, *14*(12), 5637–5649. https://doi.org/10.5194/essd-14-5637-2022

Zhang, Z., & Wang, K. (2021). Quantifying and adjusting the impact of urbanization on the observed surface wind speed over China from 1985 to 2017. *Fundamental Research*, *1*(1), 785–791. https://doi.org/10.1016/j.fmre.2021.09.006

Zhao, R., Fu, P., Zhou, Y., Xiao, X., Grebby, S., Zhang, G., & Dong, J. (2022). Annual 30-m big lake maps of the Tibetan plateau in 1991–2018. *Scientific Data*, *9*(1), 164. https://doi.org/10.1038/s41597-022-01275-9

Zhao, X., Xia, H., Pan, L., Song, H., Niu, W., Wang, R., et al. (2021). Drought monitoring over yellow river basin from 2003–2019 using reconstructed MODIS land surface temperature in google earth engine. *Remote Sensing*, *13*(18), 3748. https://doi.org/10.3390/rs13183748

Zhou, C., He, Y., & Wang, K. (2018). On the suitability of current atmospheric reanalyzes for regional warming studies over China. *Atmospheric Chemistry and Physics*, *18*(11), 8113–8136. https://doi.org/10.5194/acp-18-8113-2018

Zhou, C., & Wang, K. (2016). Land surface temperature over global deserts: Means, variability, and trends. *Journal of Geophysical Research: Atmospheres*, *121*(24), 14344–14357. https://doi.org/10.1002/2016JD025410

Zhu, W., Lű, A., & Jia, S. (2013). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment*, *130*, 62–73. https://doi.org/10.1016/j.rse.2012.10.034