Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime MODIS instantaneous observations

Zefeng Xing, Zhao-Liang Li, Si-Bo Duan, Xiangyang Liu, Xiaopo Zheng, Pei Leng, Maofang Gao, Xia Zhang, Guofei Shang

Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime MODIS instantaneous observations

Zefeng Xing[1,2,3], Zhao-Liang Li[1,3,*], Si-Bo Duan[1], Xiangyang Liu[1], Xiaopo Zheng[3], Pei Leng[1], Maofang Gao[1], Xia Zhang[4], Guofei Shang[4]

1 Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China

2 Key Laboratory of Cultivated Land Use, Ministry of Agriculture, Chinese Academy of Agricultural Engineering, Beijing 100121, China

3 ICube Laboratory, UMR 7357, CNRS-University of Strasbourg, 300 bd Sébastien Brant, CS 10413, F-67412 Illkirch Cedex, France

4 School of Land Science and Space Planning, Hebei GEO University, 050031, Hebei, China

***** Correspondence: lizhaoliang@caas.cn

Abstract

Accurate estimations of daily mean land surface temperature (LST) are important for investigating the urban heat island effect, land-atmosphere energy exchanges, and global climate change. Moderate Resolution Imaging Spectroradiometer (MODIS) sensors can provide up to four instantaneous LSTs of a single day across the world. However, numerous studies, such as those on climate change and hydrology, require the input of daily mean LSTs rather than instantaneous value. In this paper, we propose a practical method to estimate the daily mean LST using instantaneous LST products derived from MODIS. Based on the in situ LST measurements collected from 235 sites distributed globally, multiple linear regressions of two to four valid instantaneous LSTs at different MODIS observations moments (at least one daytime and one nighttime observations) can provide reliable estimates of daily mean LSTs under all-weather conditions with a root mean square error (RMSE) of less than 1.60 K. In addition, the conditions of clouds would affect the estimation accuracy of daily mean LST to a certain extent. Subsequently, an algorithm is proposed to produce the most complete coverage of daily mean LSTs from instantaneous LST products derived from MODIS.

Validation results with in situ measurements show that the daily mean LSTs estimated from the MOD11A1 and MYD11A1 products are similar to the daily mean of the in situ LST, with an RMSE of 2.17 K. Furthermore, the daily mean LST derived from MODIS data is successfully applied to calculate the global annual cycle parameters (ACPs) in the annual temperature cycle (ATC) model. The results of this study show that the daily mean LST can be retrieved accurately from combinations of daytime and nighttime LSTs derived from MODIS. We expect that our findings will be useful for various applications involving global LST trend analysis and climate change.

Keywords: Land surface temperature; daily mean temperature; MODIS

1. Introduction

Land surface temperature (LST) is an important variable within the Earth climate system. LST is related to the balance of energy and water between the land surface and atmosphere (Anderson et al., 2008; Becker and Li 1990; Li et al., 2013), influencing the growth of vegetation and land use and land cover (Amiri et al., 2009; Tran et al., 2017). LSTs retrieved from satellite measurements are widely used in drought monitoring (Huang et al., 2008), surface soil moisture and evapotranspiration estimations (Jiang and Islam 1999; Leng et al., 2017; Sandholt et al., 2002; Wang et al., 2020), near-surface air temperature retrieval (Vancutsem et al., 2010; Zhu et al., 2013), and urban heat island and climate change studies (Li et al., 2015; Weng et al., 2004). High-accuracy LST values can document secular trends in global temperature, such that the International Geosphere and Biosphere Programme (IGBP) listed LST as one of the priority parameters (Townshend et al., 1994). Moreover, LST is regarded as an essential climate variable by the Global Climate Observing System (GCOS) for evaluating land surface and land-atmosphere exchange processes, as well as providing observations of the changes in surface temperature at global and regional scales (Trenberth et al., 2013).

However, there are still some problems to be solved urgently in the research on the LST applications. For example, the satellite-derived LST is an instantaneous value under a clear sky condition, and the instantaneous LST at a certain time of the day cannot represent the

59 “average state” of the surface temperature of the day. Thus, the “annual mean LST” derived
60 from the annual temperature cycle (ATC) models with instantaneous LSTs is not the actual
61 annual mean LST, and it is difficult to apply to current climate change and modeling studies
62 (Bechtel 2015; Bechtel and Sismanidis 2018; Fu and Weng 2018; Liu et al., 2019b; Weng and
63 Fu 2014a). If the instantaneous LSTs observed by satellites can be converted into daily mean
64 surface temperatures, we can obtain the ATC parameters with a clear physical meaning. In
65 research on climate change, hydrology, etc., daily, monthly or annual mean LST is more
66 valuable than instantaneous LST as they are key indicators when monitoring global surface
67 temperatures over a long time series (Lawrimore et al., 2011; Semenov and Stratonovitch
68 2010; Vincent et al., 2012; Warren 2006). Therefore, it is necessary to carry out research to
69 convert the limited instantaneous LST derived from satellites into daily mean LST.
70 Thermal infrared sensors onboard polar-orbiting satellites can provide global observations
71 with high spatial and temporal resolution. For example, LST data-sets from the Moderate
72 Resolution Imaging Spectroradiometer (MODIS) onboard the Terra (2000 -present) and Aqua
73 (2002 -present) platforms (Duan et al., 2019; Wan 2014), as well as from the Sea and Land
74 Surface Temperature Radiometer (SLSTR) onboard the Sentinel-3A (2016 -present) and 3B
75 (2018 -present) platforms, have been widely used in many fields (Hu et al., 2019; Qin et al.,
76 2013; Sobrino and Irakulis 2020; Son et al., 2012; Wooster et al., 2012). However, these
77 sensors can only provide a few daily observations in same location; cloud contamination also
78 affects these observations. Thus, these problems inhibit the applications of the LST products
79 retrieved by the polar-orbiting satellites. To better understand how rising LSTs affect glaciers,
80 ice sheets, vegetation in Earth’s ecosystems, and climate change, estimating daily mean LSTs
81 from limited daily observations is especially important, which is the foundation of estimating
82 monthly or annual mean LSTs. Compared with the NCEP/NCAR Reanalysis dataset and
83 GISS Surface Temperature Analysis dataset produced from ground-based stations using
84 interpolation and average methods (Kalnay et al., 1996; Lenssen et al., 2019), satellite
85 observations can provide real measurement data with global coverage and higher spatial
86 resolution. For this reason, the daily mean temperature estimated from instantaneous satellite
87 observations can provide effective supplementary data for global climate change studies.

The daytime and nighttime LSTs from MODIS provide the possibility of estimating the daily mean LST. Recently, several studies have attempted to estimate the daily mean LST using daytime and nighttime LSTs derived from MODIS. Under clear sky conditions, variations in the LST can be modeled by diurnal temperature cycle (DTC) models (Duan et al., 2012; Göttsche and Olesen 2001; Weng and Fu 2014b), allowing the calculation of daily mean LST from the DTC models (Ouyang et al., 2012). However, successful application of DTC models requires a nearly cloud-free diurnal cycle and at least four valid observations (Duan et al., 2013), which limits the application of DTC models in estimating the daily mean LST. Williamson et al. (2014) proposed the Interpolated Curve Mean Daily Surface Temperature (ICM) method by interpolating MODIS daytime LSTs to the daily mean LST using coincident diurnal air temperature curves (Williamson et al., 2014). However, the ICM method is not practical because it relies on air temperature data; moreover, the air temperature can only be monitored on site, such that the data is not continuous in space. Williamson et al. (2014) also proposed a simple method to calculate the daily mean LST by averaging the daily maximum and minimum LSTs obtained using daytime and nighttime LSTs derived from MODIS. This simple average method exhibits a stronger correlation with the daily mean air temperature than the ICM method. However, the daily maximum and minimum LSTs are difficult to estimate from MODIS LSTs using the simple average method. In the context of estimated daily mean air temperature, when continuous observations cannot be applied, scientists usually used the combination of the limited observations to calculate the daily mean LST. For example, Nordli et al. (1996) employ a linear combination of measurements taken at different times of the day (Nordli et al., 1996). Some other researchers build linear or multiple regression models based on the maximum, and minimum temperature for the MODIS daytime and nighttime LSTs to estimate daily mean air temperature (Dall'Amico and Hornsteiner 2006; Hu et al., 2020; Huang et al., 2015; Rao et al., 2019). Chen et al. (2017) used the average values of daytime and nighttime MODIS products, i.e., MOD11C3 and MYD11C3, to estimate the monthly mean LST (Chen et al., 2017). They used ground-based measurements to validate the monthly mean Aqua day/night LSTs, Terra day/night LSTs, and Aqua/Terra day/night LSTs. Their results showed that the mean values at Aqua/Terra monthly day/night

observation times provide an accurate estimate of the monthly mean LST, with a root mean square error (RMSE) of 2.65 K. However, their method only considers a limited number of combinations. In addition, their investigation focused on estimating the monthly mean LST, and not on calculations of the daily mean LST. Therefore, estimating the daily mean LST from a few instantaneous observations requires a practical method.

Inspired by the work of Chen et al. (2017), in this study, we consider using more linear combinations of LST values observed at different observation times for Terra and Aqua to estimate daily mean LST values with ground-based measurements, providing that there are at least one daytime and one nighttime observation. In addition, we validate the accuracy associated with estimating the daily mean LST for the different combinations using in situ measurements. This paper is organized as follows. Section 2 introduces materials and methods. Section 3 provides the results of estimating daily mean LST. Discussions and conclusions are provided in Sections 4 and 5, respectively.

2. Materials and Methods

2.1. In situ measurements

To calculate the in situ LST, ground-based measurements of upwelling and downwelling longwave radiation at 235 sites were collected from different geographical coordinates and land cover categories. The in situ observations were used to calculate the actual daily mean LST and the in situ LST at the MODIS observation times. These in situ sites were mainly derived from six observation networks: Surface Radiation Budget Network (SURFRAD), AmeriFLux network, OzFlux Network (Australia), European Fluxes Database cluster, Asia flux network, and China automatic weather stations.

The SURFRAD (https://www.esrl.noaa.gov/gmd/grad/surfrad/) network was created in the mid- 1990s to support satellite retrieval validation, modeling, and various scientific studies by providing measurements of surface radiation over the United States (Augustine et al., 2000). All seven SURFRAD sites were selected in this study. The primary measurements from SURFRAD are the downwelling and upwelling components of broadband solar and thermal infrared irradiance every 3min (before 2009) or every minute (after 2009). The longwave

radiation is measured by two pyrgeometers (ventilated Eppley pyrgeometer, spectral range from 3.0 to 50.0 μm) deployed at a 10-m high tower. The longwave radiation can be used to calculate high-quality LSTs and have been widely used to evaluate MODIS LST products (Duan et al., 2019; Wang and Liang 2009).

AmeriFlux (https://ameriflux.lbl.gov/) is a network of ~170 long-term research stations measuring ecosystem CO_2, water, and energy fluxes in North, Central, and South America. AmeriFlux was launched in 1996 and aims to connect field sites representing major climate and ecological biomes, including tundra; grasslands; crops; and conifer, deciduous, and tropical forests (Boden et al., 2013). AmeriFlux data have improved our understanding of flux variations and modeling in climate science (Novick et al., 2018). Processed level 2 (without gap filling) data for upwelling and downwelling longwave radiation at 120 sites were used in this study.

OzFlux (http://data.ozflux.org.au/) is the regional Australian and New Zealand flux tower network that aims to provide a continental-scale national research facility for monitoring and assessing trends and improve predictions of Australasia's terrestrial biosphere and climate (Beringer et al., 2016). This network consists of nearly 30 flux towers in Australia and New Zealand; 25 OzFlux sites with measurements of longwave radiation were selected in this study. The longwave radiation is measured by CNR1 or CNR4 radiometers with a temporal resolution of half an hour.

The European Fluxes Database Cluster (http://www.europe-fluxdata.eu/) was created with the aim to host a single infrastructure flux measurement between ecosystems and atmosphere and provide standard and high-quality data processing (Aubinet et al., 1999). The database receives and distributes fluxes of different greenhouse gases, meteorological variables, ancillary data and meta-information acquired at sites from Europe, Africa, and other continents. In this study, 43 sites with half-hourly downwelling and upwelling longwave radiation were selected.

AsiaFlux (http://asiaflux.net/) is a regional research network bringing together scientists from universities and institutions in Asia to study the exchanges of carbon dioxide, water vapor, and energy between terrestrial ecosystems and the atmosphere across daily to

inter-annual time scales. The incoming and outgoing longwave radiation of each site were measured by CNR-1, KIPP&ZONEN. In this study, we collected 22 sites with half-hourly longwave radiation from the AsiaFlux.

In addition, 18 China automatic weather sites were used in this study, including 15 sites in the Heihe Basin (Li et al., 2009) and three sites in the Haihe Basin (Liu et al., 2013). These sites are equipped with a Kipp & Zonen CNR1 net radiometer, which measures downwelling and upwelling longwave radiance (spectral range from 5 to 50 μm) from a height of 10 m. The temporal resolution of the recorded longwave radiance is10 min.

These ground-based stations (sites) are located in different regions of the world and cover different land-use types and topographies. Fig. 1 shows the distribution of these sites. The sites marked in blue color will be used in this study for modeling and temporal validation and the sites marked in red color will be used for spatial validation.

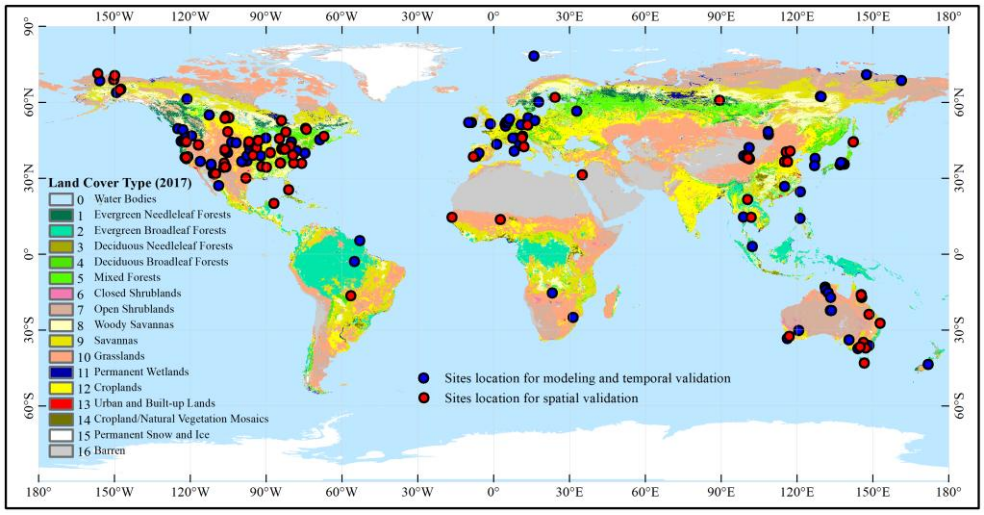


Fig. 1. Location of the in situ observation sites used in this study for modeling, temporal and spatial validation. The base map shows the land cover types which come from the MCD12C1 product in 2017.

Fig. 2 shows the distribution of the in situ sites in different land cover types and climate zones. One can see that there are at least 10 in situ sites in each surface type, except for water bodies (ID=0), evergreen needleleaf forests (ID=1), open shrublands (ID=7), croplands (ID=12), cropland/natural vegetation mosaics (ID=14), permanent snow and ice (ID=15) and Barren (ID=16). In addition, these sites are also distributed in all five climatic zones, and most sites are located in dry, temperate, and continental climate zones, with relatively few

sites in tropical and polar climate zones.

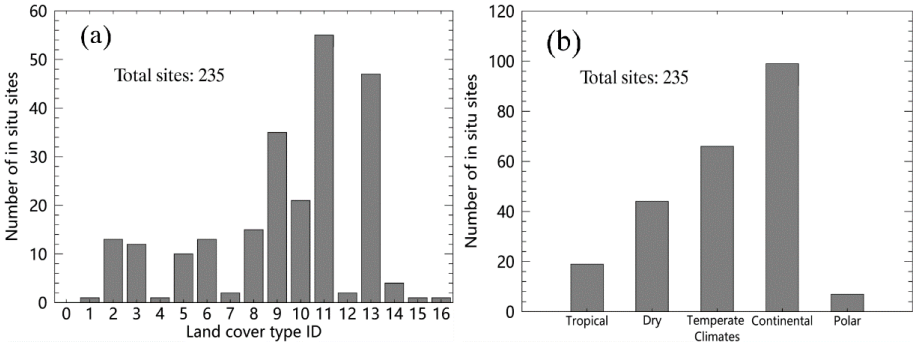


Fig. 2. The distribution of in situ sites in different (a) land cover types shown in Fig. 1, and (b) climate zones.

2.2. Daily mean LST calculated from in situ LST measurements

2.2.1 In situ LST measurements

For the pyrgeometers at the SURFRAD sites and the CNR1 net radiometer at the AmeriFlux, OzFlux, EuropeanFluxes, and China automatic weather sites, in situ LST was calculated from downwelling and upwelling broadband hemispherical radiation using Stefan-Boltzmann's law as follows:

$$LST_g = \sqrt[4]{\frac{R_\uparrow - (1-\varepsilon_b) R_\downarrow}{\sigma \varepsilon_b}} \quad , \tag{1}$$

where LST_g is the in situ LST (K), $R_\uparrow$ is the upwelling longwave radiation (W/m^2), $R_\downarrow$ is the downwelling longwave radiation (W/m^2), σ is the Stephan-Boltzmann constant ($5.67 \times 10^{-8} W/m^2/K^4$), and ε_b is the surface broadband emissivity.

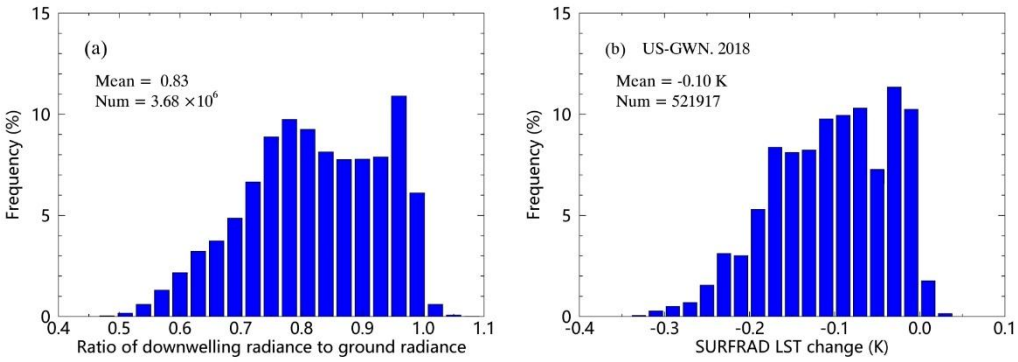


Fig. 3. Frequency histogram of (a) the ratio of downwelling radiance to ground radiance (i.e., σLST_g^4) at the seven SURFRAD sites and (b) change in the estimated LSTs caused by a variation of the broadband emissivity from 0.97 to 0.98. The data at US-GWN (SURFRAD) in 2018 under

all-conditions were used.

According to Eq. (1), we can obtain the expression of $\Delta LST_g = -\left(\frac{LST_g}{4}\right) \cdot \Delta\varepsilon \cdot (1 - \frac{R_\downarrow}{\sigma LST_g^4})$. Assuming that LST_g is equal to approximately 300 K and the ratio of downwelling radiance ($R_\downarrow$) to ground radiance (σLST_g^4) is greater than 0.5, then $\Delta LST_g < -37.5 \cdot \Delta\varepsilon$. This indicates that ε increases by 0.01 ($\Delta\varepsilon$) and LST does not decrease by more than 0.37 K (ΔLST_g). Fig. 3(a) shows the frequency histogram of the ratio of downwelling radiance to ground radiance at the seven SURFRAD sites. The result shows that this ratio ranges from 0.5 to 1.0 with a mean value of 0.83. This result indicates that the emissivity changes by 0.01, and the change in LST_g will not exceed 0.37 K. Fig. 3(b) shows the impact that changing the broadband emissivity from 0.97 to 0.98 has on the estimated LST at the SURFRAD GWN site for all data in 2018 using Eq. (1). The results indicate that an error of 0.01 in the broadband emissivity will cause a LST error of -0.3 to 0.0 K. According to previous studies, the broadband emissivity (3-30 μm) of most soil and vegetation surfaces ranges from 0.955 to 0.982 (Ogawa et al., 2002; Ogawa et al., 2008). In other studies, the broadband emissivity of the seven SURFRAD sites can be assumed as a fixed value of 0.97 when calculating the in situ LST (Heidinger et al., 2013). The land cover types of the 235 sites in this study are soil or partly covered with vegetation. Therefore, the broadband longwave emissivity values of all 235 sites were assumed to be 0.97 in this study.

2.2.2. Daily mean in situ LST

The in situ measurements are usually used to validate the satellite LST products (Duan et al., 2019; Li et al., 2014; Wang et al., 2008). Therefore, the in situ measurements can be regarded as the "true" LST values. In this study, once the instantaneous LSTs of one day are calculated, the daily mean LST can also be calculated. The daily mean in situ LST can be calculated as follows:

$$dmLST_g = \frac{1}{n}\sum_{i=1}^{n} LST_g(i), \quad (2)$$

where $dmLST_g$ is the daily mean in situ LST, $LST_g(i)$ is the in situ LST values derived from Eq. (1), and n is the daily total number of in situ measurements. We note that the recorded time of the observations (i.e., UTC time or local standard time) was transformed to the local solar

time. If the observations are incomplete on a certain day, the daily mean in situ LST for that day will not be calculated.

2.3. MODIS data

The MODIS Terra/Aqua Land Surface Temperature and Emissivity L3 Global 0.05-degree latitude/longitude Climate Modeling Grid products (MOD11C1 and MYD11C1, Collection 6) and 1-km SIN Grid products (MOD11A1 and MYD11A1, Collection 6) were downloaded from the National Aerodynamics and Space Administration (NASA) website (https://search.earthdata.nasa.gov) and will be used in this study. The MOD11A1 and MYD11A1 products provide daily LST and emissivity values using the generalized split-window algorithm (Wan and Dozier 1996). They are used to determine MODIS observation time as well as to determine whether pixels corresponding to ground observation stations (sites) were affected by clouds. In addition, the MOD11A1 and MYD11A1 products are used to validate the accuracy of the proposed method in estimating the daily mean LST. The MOD11C1 and MYD11C1 products are generated mostly by a physics-based day/night algorithm (Wan and Li 1997). The generalized split-window algorithm is used to supplement the LSTs retrieved by the day/night LST algorithm at grids where there is no valid pair of day and night observations (Becker and Li 1990; Wan and Dozier 1996). In this study, the MOD11C1 and MYD11C1 products will be used to calculate the global MODIS daily mean LST and then used to simulate annual temperature cycles. The day/night LSTs, quality control (QC), and observation time layers from the MOD11C1/MYD11C1 and MOD11A1/MYD11A1 products are extracted in this study.

2.4. Estimation of daily mean LST from available combinations of in situ observations

2.4.1. Rationale

According to previous DTC models, the diurnal cycle of the clear-sky LST can be approximated by a constant plus a sine or cosine function (Duan et al., 2012; Göttsche and Olesen 2001; Van den Bergh et al., 2006). The mean LST value of the diurnal cycle (i.e.,

within 24 h of a cycle) is mathematically equivalent to this constant (Ouyang et al., 2012). In practice, the average of any two points that differ by 12 h on this diurnal cycle is also equivalent to this constant. Therefore, at least two observations that differ by 12 h can be used to estimate the daily mean LST based on this principle. The two observation times of a polar-orbiting satellite (i.e., Terra or Aqua) in the same area is approximately 12 h (the overpass times of Terra and Aqua are approximately 10:30 AM/PM and 1:30 AM/PM, respectively) (Wan and Dozier 1996), such that we can use the combination of these two to four instantaneous observations to estimate the daily mean LST at any weather conditions. Considering that the presence of clouds will reduce the LST value of a day, or that the change in the LST does not follow the cosine or sine function, the multiple linear regression model rather than the simple average method will be used in this study to describe how the daytime and nighttime LSTs respond to the actual daily mean LST.

Taking into account the valid instantaneous LSTs obtained by the MODIS sensor in a day, we use nine combinations of daytime and nighttime LSTs to estimate the daily mean LST, including four combinations with two valid LSTs (one daytime and one nighttime LSTs), four combinations with three valid LSTs (two daytime and one nighttime LSTs, or two nighttime and one daytime LSTs), and a combination with four valid LSTs (two daytime and two nighttime LSTs).

2.4.2. Determination of the in situ LST at MODIS observation times

In order to develop the relationship between the daily mean LST and day/night instantaneous LSTs, the in situ LST (LST_g) at the Terra/Aqua observation time must be determined. As the time of the ground-based LST measurement may not match the Terra/Aqua observation time, the in situ LST at the Terra/Aqua observation time cannot be obtained directly. A Science Data Set (SDS) observation time layer corresponding to the 235 sites was extracted from the MOD11A1 and MYD11A1 products. Assuming that the LST varies linearly around the satellite observation time, the linear interpolation method (Eq. (3)) can be used to calculate the in situ LST (LST_g) at the Terra/Aqua observation time of cloud-free pixels (Liu et al., 2019a). Then, the in situ LSTs at the Terra day/night observation time and Aqua day/night observation time can be obtained as follows:

$$LST_g(t) = LST_g(t_1) + (t - t_1) \cdot \frac{\left[LST_g(t_2) - LST_g(t_1) \right]}{(t_2 - t_1)}, \quad (3)$$

where $LST_g(t_1)$ and $LST_g(t_2)$ are the in situ LSTs of two adjacent observation times ($t_1$and t_2, respectively) and $LST_g(t)$ is the in situ LST at time t ($t_1 \le t \le t_2$).

2.4.3. Estimation of daily mean LST via multiple linear regression model

As the traditional average method of direct averaging tends to overestimate the daily average surface temperature (Chen et al., 2017; Ouyang et al., 2012), we perform multiple linear regression of the nine day/night combinations with 158 sites (a two-thirds of sites in each climate zones shown in Fig. 2(b)) from 2003 to 2012 to determine the model coefficients. For the combinations of two valid LSTs (one daytime and one nighttime LST), the regression models can be written as follows:

$$dmLST_g = k_1 \times LST_g(T^d) + k_2 \times LST_g(T^n) + b, \quad (4)$$

$$dmLST_g = k_1 \times LST_g(T^d) + k_2 \times LST_g(A^n) + b, \quad (5)$$

$$dmLST_g = k_1 \times LST_g(A^d) + k_2 \times LST_g(A^n) + b, \text{ and} \quad (6)$$

$$dmLST_g = k_1 \times LST_g(A^d) + k_2 \times LST_g(T^n) + b \ . \quad (7)$$

For the combinations of three valid LSTs (two daytime LSTs and one nighttime LST, or two nighttime LSTs and one daytime LST), the regression models can be described as follows:

$$dmLST_g = k_1 \times LST_g(T^d) + k_2 \times LST_g(A^d) + k_3 \times LST_g(T^n) + b, \quad (8)$$

$$dmLST_g = k_1 \times LST_g(T^d) + k_2 \times LST_g(A^d) + k_3 \times LST_g(A^n) + b, \quad (9)$$

$$dmLST_g = k_1 \times LST_g(T^n) + k_2 \times LST_g(A^n) + k_3 \times LST_g(T^d) + b, \text{ and} \quad (10)$$

$$dmLST_g = k_1 \times LST_g(T^n) + k_2 \times LST_g(A^n) + k_3 \times LST_g(A^d) + b. \quad (11)$$

For the combinations of four valid LSTs (two daytime LSTs and two nighttime LSTs), the regression model can be described as follows:

$$dmLST_g = k_1 \times LST_g(T^d) + k_2 \times LST_g(T^n) + k_3 \times LST_g(A^d) + k_4 \times LST_g(A^n) + b. \quad (12)$$

In Eqs. (4)-(12), $dmLST_g$ is the ground-based daily mean LST, k_i and b are the fitting coefficients, "*T*" and "*A*" represent the Terra and Aqua satellites, respectively, and

superscripts "*d*" and "*n*" represent the daytime and nighttime, respectively. For example, $LST_g(T^d)$ is the in situ LST at the Terra daytime observation.

2.4.4. Traditional average method

According to previous studies, daily mean LST can be calculated by averaging the MODIS Aqua daytime (10:30 am) and nighttime (10:30 pm) LSTs (Chen et al., 2017; Ouyang et al., 2012; Williamson et al., 2014). The simple average method can be expressed as follows:

$$dmLST_g = 0.5 \times LST_g(A^d) + 0.5 \times LST_g(A^n) \quad (13)$$

2.4.5. Estimation daily mean LST from MODIS data

Because the accuracy of estimating daily average surface temperature varies with different combinations, we propose a framework for calculating MODIS daily mean LST from available LSTs. Fig. 4 illustrates the framework for estimating the daily mean LST from available MODIS observations. Firstly, we obtained valid observations of daytime and nighttime LSTs from MODIS instantaneous LST products. Then, we calculated the daily mean LST derived from MODIS using a combination of Eqs. (4) -(12) and the fitting coefficients listed in Table1 based on the number of valid observations. Based on Fig. 4, if there are four valid MODIS observations, Eq. (12) can be used to estimate the daily mean LST; for three valid observations, Eqs. (8)-(11) can be used to calculate the daily mean LST; if there are two observations (one daytime and one nighttime LST), then Eqs. (4)-(7) can be used to estimate the daily mean LST. However, if the two observations are both daytime or nighttime LSTs, or if the number of valid observations less than two, the daily mean LST of this pixel is set to a null value.

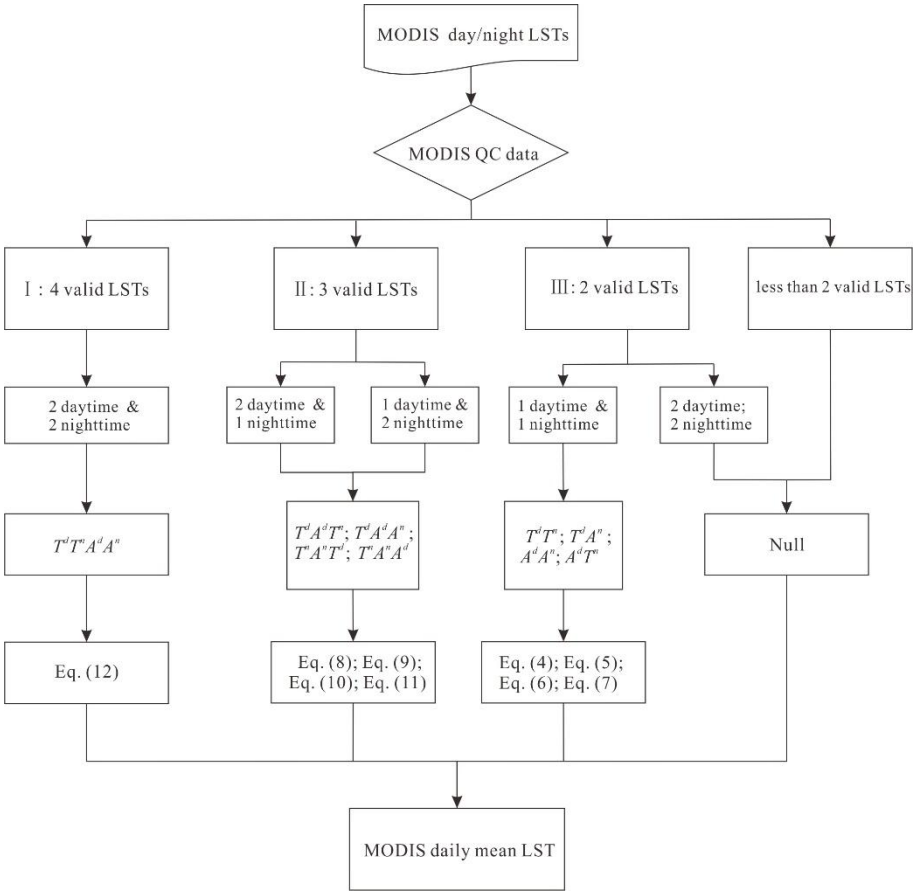


Fig. 4. Framework for estimating the daily mean LST from available MODIS day/night LSTs. T^dT^n, T^dA^n, A^dA^n, A^dT^n, $T^dA^dT^n$, $T^dA^dA^n$, $T^nA^nT^d$, $T^nA^nA^d$, and $T^dT^nA^dA^n$ are the nine combinations of MODIS valid LSTs. "*T*" and "*A*" represent Terra and Aqua satellites, respectively. Superscript "*d*" and "*n*" represent daytime and nighttime, respectively. "Ⅰ", "Ⅱ", and "Ⅲ" represent the priority of the calculation process.

In situ measurements at the seven SURFRAD sites (BND, DRA, FPK, GWN, PSU, SXF, and TBL sites) were used to validate the daily mean LSTs estimated from the instantaneous LST products derived from MODIS. Table 1 provides specific details of the seven SURFRAD sites. The 1-km MOD11A1 and MYD11A1 products in 2018 were selected to calculate the daily mean LST of the seven SURFRAD sites using the above-mentioned framework. To minimize the impact of cloud contamination on the validation results, only high-quality LST values (i.e., QC = 0) were extracted for the pixel closest to each site based on the longitude and latitude. As the error of MODIS LST would affect the accuracy of estimating the daily mean LST, we validated the accuracy of the day/night LST derived from MODIS at the seven SURFRAD sites.

Table 1. Details of the seven SURFRAD sites used in this study.

Site code*	Latitude	Longitude	Elevation	Land cover type
BND	40.0516° N	88.3733° W	230 m	Croplands
TBL	40.1256° N	105.2378° W	1689 m	Grasslands
DRA	36.6232° N	116.0196° W	1007 m	Open shrublands
FPK	48.3080° N	105.1018° W	634 m	Grasslands
PSU	40.7203° N	77.9310° W	376 m	Cropland
SXF	43.7343° N	96.6233° W	473 m	Croplands

* BND: Bondville, TBL: Table Mountain, DRA: Desert Rock, FPK: Fort Peck, PSU: Pennsylvania State University, SXF: Sioux Falls.

2.5. Validation strategy

To validate the generalization of the multiple regression method proposed in this study, we divide the ground measurements of 235 sites into three parts. The first part (158 sites) is two-thirds of sites in each climate zones from 2003 to 2012, which is used to construct a multiple linear regression method to obtain the corresponding coefficients. The second part (158 sites) is the two-thirds of the sites in each climate zone from 2013 to 2018, which is used to validate our approach on a time scale. The third part (77 sites) is one-third of sites in each climatic zone from 2003 to 2018, which is used to validate our method on a spatial scale.

For the comparison between various combinations, the sample sizes of the above three data sets are equal to the last combination, in which there are four valid observations. Additionally, we use all 235 sites from 2003 to 2018 to assess the proposed method on the situation in which there are only two or three valid observations at MODIS view times.

2.6. Application of daily mean LST

The MODIS daily mean surface temperature of MODIS has great potential applications. In this study, we used the MYD11C1 and MOD11C1 products to calculate global daily mean LST. Then an ATC model was used to fit the annual variations of daily mean LST to obtain the annual mean LST, amplitude, and phase. The ATC model is usually applied to modeling seasonal or interannual variations in air temperature or land surface temperature (Bechtel and

Sismanidis 2018; Fu and Weng 2018; Huang et al., 2016; Weng and Fu 2014a; Zou et al., 2018).We used the ATC model with three free parameters to describe the annual variation in the daily mean LST. The ATC model consists of a constant and sine function with the reference day as the spring equinox (Bechtel 2011). To better explain the physical meaning of the phase in the model, we used the cosine function instead of the sine function in this study and the first day of the year as the reference day (Xing et al., 2020):

$$dmLST(x) = a + b\cos\left[\frac{2\pi}{365}(x-c)\right] \tag{14}$$

where *dmLST* is the daily mean LST, *x* is the day of year, and *a*, *b* and *c* are three annual cycle parameters (ACPs); *a* is the annual mean LST, *b* is the annual amplitude, and *c* is the annual phase, which is the date when *dmLST* reaches its maximum values in a year.

3. Results

3.1. Daily mean in situ LST versus daily mean LST estimated via the combinations of day/night observations

Fig. 5 shows the density scatter plots for daily mean in situ LST estimated using Eq. (2) against the daily mean LST estimated with the multiple linear regression method with Eqs. (4)-(12). Table 2 lists the statistical parameters for estimating the daily mean LST using the linear regression method. The results show that the nine linear combinations closely correspond to the actual daily mean LST, with an R^2 value greater than 0.99 and an RMSE value of less than 1.60 K. The multiple linear regression of the two valid LSTs (i.e., T^dT^n, T^dA^n, A^dA^n, and A^dT^n, see Figs. 5(a)-5(d), respectively) showed a similar performance with RMSE values of 1.58, 1.55, 1.50, and 1.60 K, respectively. The RMSE value of the three valid LSTs (i.e., $T^dA^dT^n$, $T^dA^dA^n$, $T^nA^nT^d$, and $T^nA^nA^d$, see Figs. 5(e)-5(h), respectively) are 1.41, 1.43, 0.93, and 0.91 K, respectively. The regression of the four valid LSTs ($T^dT^nA^dA^n$, see Fig. 5(i)) is the most accurate for estimating the daily mean LST, with an RMSE of 0.80 K. This result indicates that the accuracy of the multiple regression model using Eqs. (4)-(12), with the fitting coefficients listed in Table 2, achieved a good accuracy of daily mean LST with RMSE less than 1.62 K. In addition, an increase in the number of valid LSTs used for regression

leads to higher accuracies when estimating the daily average LST. Fig. 6 shows histograms of the RMSE values for the 158 sites used to estimate the daily mean LSTs through the multiple linear regression model with the nine combinations. The RMSE values of most sites are within a reasonable range, from 0.5 to 2.5 K, indicating that the linear regression model can provide reliable estimates of the daily mean LST.

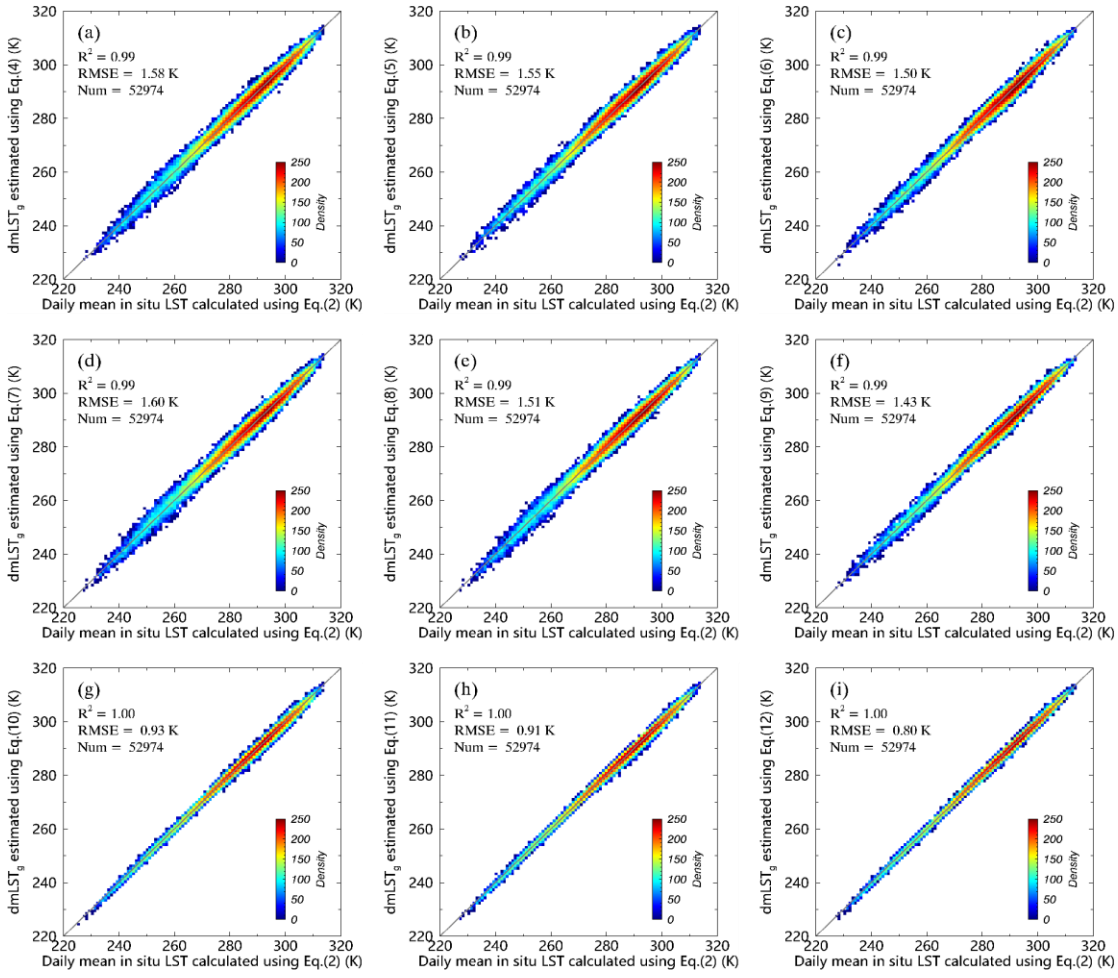


Fig. 5. Density scatter plots of daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with multiple linear regression method using Eqs. (4)-(12) at 158 sites (a two-thirds of sites in each climate zones shown in Fig. 2b) from 2003 to 2012. The straight grey line is a 1:1 line.

Table 2. Statistics for comparing the relationship between the regressions of the nine combinations and actual daily mean LST.

Case	Combinations	Fitting coefficients					R^2	RMSE (K)
		k_1	k_2	k_3	k_4	b		
1	T^dT^n	0.3925	0.5993	-	-	1.40	0.99	1.58
2	T^dA^n	0.4354	0.5630	-	-	0.64	0.99	1.55
3	A^dA^n	0.4244	0.5637	-	-	2.75	0.99	1.50
4	A^dT^n	0.3821	0.5992	-	-	3.64	0.99	1.60
5	$T^dA^dT^n$	0.2172	0.1802	0.5875	-	2.88	0.99	1.51
6	$T^dA^dA^n$	0.1942	0.2437	0.5528	-	2.19	0.99	1.43
7	$T^nA^nT^d$	0.3354	0.3216	0.3665	-	-6.26	1.00	0.93
8	$T^nA^nA^d$	0.3243	0.3318	0.3582	-	-4.31	1.00	0.91
9	$T^dT^nA^dA^n$	0.1807	0.3210	0.1907	0.3241	-4.75	1.00	0.80

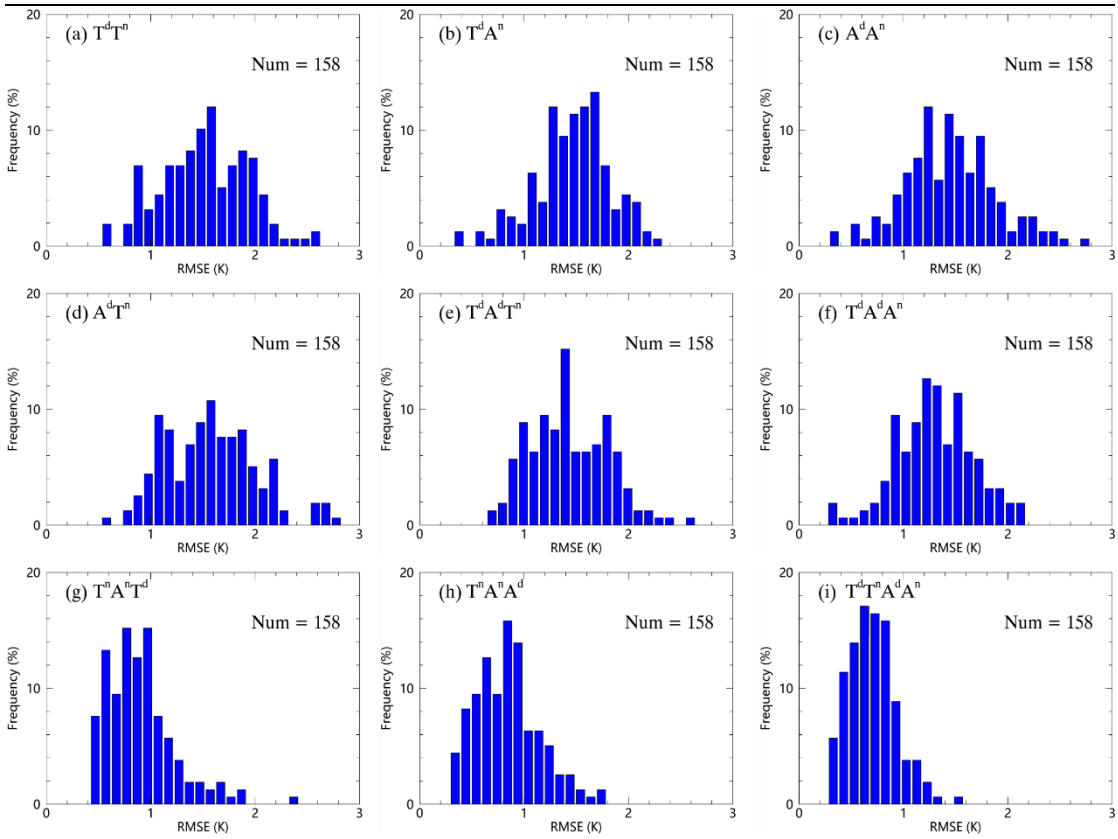


Fig. 6. Histograms of the RMSE values for the 158 sites to estimate the daily mean in situ LST calculated via the multiple linear regression method using Eqs. (4)-(12): (a)T^dT^n, (b)T^dA^n, (c)A^dA^n, (d)A^dT^n, (e)$T^dA^dT^n$, (f)$T^dA^dA^n$, (g)$T^nA^nT^d$, (h)$T^nA^nA^d$, and (i)$T^dT^nA^dA^n$. "T" and "A" represent the Terra and Aqua satellites, respectively and Superscripts "d" and "n" represent daytime and nighttime, respectively.

Fig. 7 shows the density scatter plot of actual daily mean LST versus the estimated daily mean LST with the traditional average method at 158 sites from 2003 to 2012. The RMSE and bias values of the traditional average method are 2.59 K and 1.97 K, respectively. The result indicates that the accuracy of the traditional average method is significantly lower than that of the multiple linear regression method.

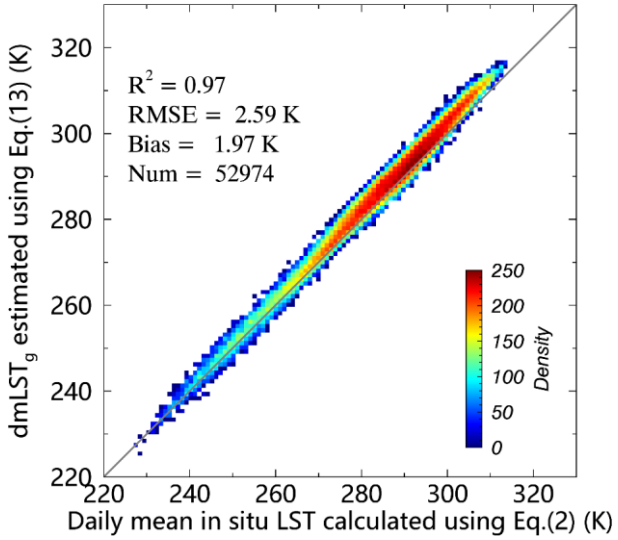


Fig. 7. Density scatter plot of daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with the traditional average method using Eq. (13) at 158 sites from 2003 to 2012. The straight grey line is a 1:1 line.

3.2. Error analysis

The total uncertainty of the proposed method for estimating the daily mean LST from MODIS products was associated with the error obtained for the instantaneous LSTs derived using MODIS and the algorithm fitting error for each combination of the day/night LSTs. Thus, the uncertainty of the proposed method for each combination was evaluated as follows:

$$\delta T_{m,total} = \sqrt{(\delta dmLST_{m,\,fitting})^2 + (\delta LST)^2}\,, \tag{15}$$

where $\delta T_{m,total}$ is the total daily mean LST uncertainty using different combinations of MODIS daytime/nighttime LSTs; $\delta dmLST_{m,fitting}$ is the algorithm fitting error for Eqs. (4)-(12); and δLST is the error from MODIS day/night LSTs for Eqs. (4)-(7) as follows:

$$\delta LST = \sqrt{[k_1 \cdot \delta LST(i^d)]^2 + [k_2 \cdot \delta LST(j^n)]^2}\,, \text{ and} \tag{16}$$

for Eqs. (8)-(11):

$$\delta LST = \sqrt{[k_1 \cdot \delta LST(i^d)]^2 + [k_2 \cdot \delta LST(j^d)]^2 + [k_3 \cdot \delta LST(j^n)]^2}\,, \text{ and} \tag{17}$$

$$\delta LST=\sqrt{[k_1\cdot\delta LST(i^n)]^2+[k_2\cdot\delta LST(j^n)]^2+[k_3\cdot\delta LST(j^d)]^2}\ . \quad (18)$$

for Eq. (12):

$$\delta LST=\sqrt{[k_1\cdot\delta LST(T^d)]^2+[k_2\cdot\delta LST(T^n)]^2+[k_3\cdot\delta LST(A^d)]^2+[k_4\cdot\delta LST(A^n)]^2} \quad (19)$$

where $\delta LST(i^d)$ and $\delta LST(j^n)$ are the accuracies of the Terra/Aqua daytime (d) and nighttime (n) LST, respectively; i and j represent the Terra (T) or Aqua (A) satellites, respectively; and k_1, k_2, k_3, and k_4 are the slopes in Eqs. (4)-(12).

According to Eq. (14), the total uncertainty of the calculated daily mean LST depends on the algorithm fitting error ($\delta dmLST_{m,fitting}$) and MODIS day/night LST error (δLST). The algorithm fitting error ($\delta dmLST_{m,fitting}$) ranged from 0.8 to 1.6 K as displayed in Fig. 5(a)-(i). According to the validation results of a previous study, the accuracy of the C6 MODIS LST product ($LST(i^d)$ or $LST(j^n)$) at the SURFRAD sites is between 1.0 and 2.0 K (Duan et al., 2019). Therefore, the total uncertainty of the daily mean LST was estimated. For Eqs. (4)-(7), $\delta T_{m,total}$ ranged from 1.7 to 2.2 K; for Eqs. (8)-(11), $\delta T_{m,total}$ ranged from 1.1 to 2.0 K; and for Eq. (12), $\delta T_{m,total}$ ranged from 1.0 to 1.3 K. The results show that more combinations of valid LSTs could produce a better accuracy.

3.3. Validation

The in situ measurements of 158 sites, same as those used for determining the model coefficients, but from 2013 to 2018 were used to validate the multiple linear method at a temporal scale. Fig. 8 illustrates that the density scatter plots of true daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with multiple linear regression method using Eqs. (4)-(12) (model coefficients are given in Table 2) at 158 sites from 2013 to 2018. The multiple linear regression of four valid LSTs (see Fig. 8(i)) shows a best performance in estimating daily mean LST, with RMSE of 0.79 K and bias of -0.05 K. The combinations of three valid LSTs (i.e., $T^dA^dT^n$, $T^dA^dA^n$, $T^nA^nT^d$, and $T^nA^nA^d$, see Fig. 8(e)-8(h), respectively) also have good performance in estimating daily mean LST, with RMSE of 1.45, 1.31, 1.06, and 0.91 K, respectively, and bias of -0.24, 0.00, -0.19, and 0.06 K, respectively. However, the combinations of two valid LSTs (i.e., T^dT^n, T^dA^n, A^dA^n, and A^dT^n,

see Fig. 8(a)-8(d), respectively) show relatively low performance, with RMSE of 1.63, 1.78, 1.42, and 1.51 K, respectively, and bias of -0.35, 0.87, 0.13, and -0.12 K, respectively. Overall, the nine multiple linear regression equations demonstrate a good performance in estimating daily mean LST, with RMSE ranging from 0.79 to 1.78 K, and bias ranging from -0.35 to 0.87 K. The validation results show that more combinations of valid LSTs could produce higher performance.

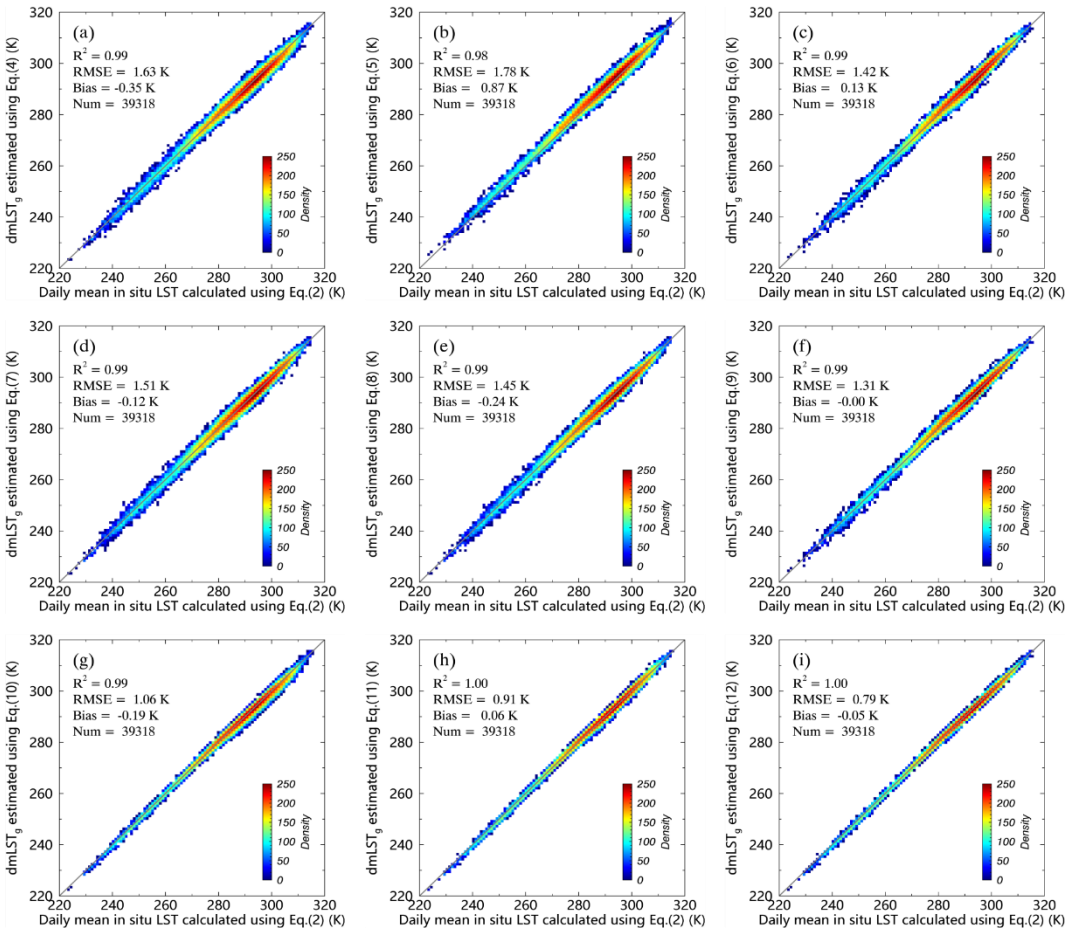


Fig. 8. Same as Fig. 5, but from year 2013 to year 2018.

Fig. 9 shows the validation results from the spatial scale. Figs. 9(a)-9(i) are the density scatter plots of true daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with multiple linear regression method using Eqs. (4)-(12) at 77 sites from 2003 to 2018. Overall, the validation results are similar to the previous validation results from the temporal scale in that the combinations with more valid LSTs show a better performance, with RMSE values ranging from 0.82 to 1.86 K, and bias values ranging from -0.42 to 1.07 K.

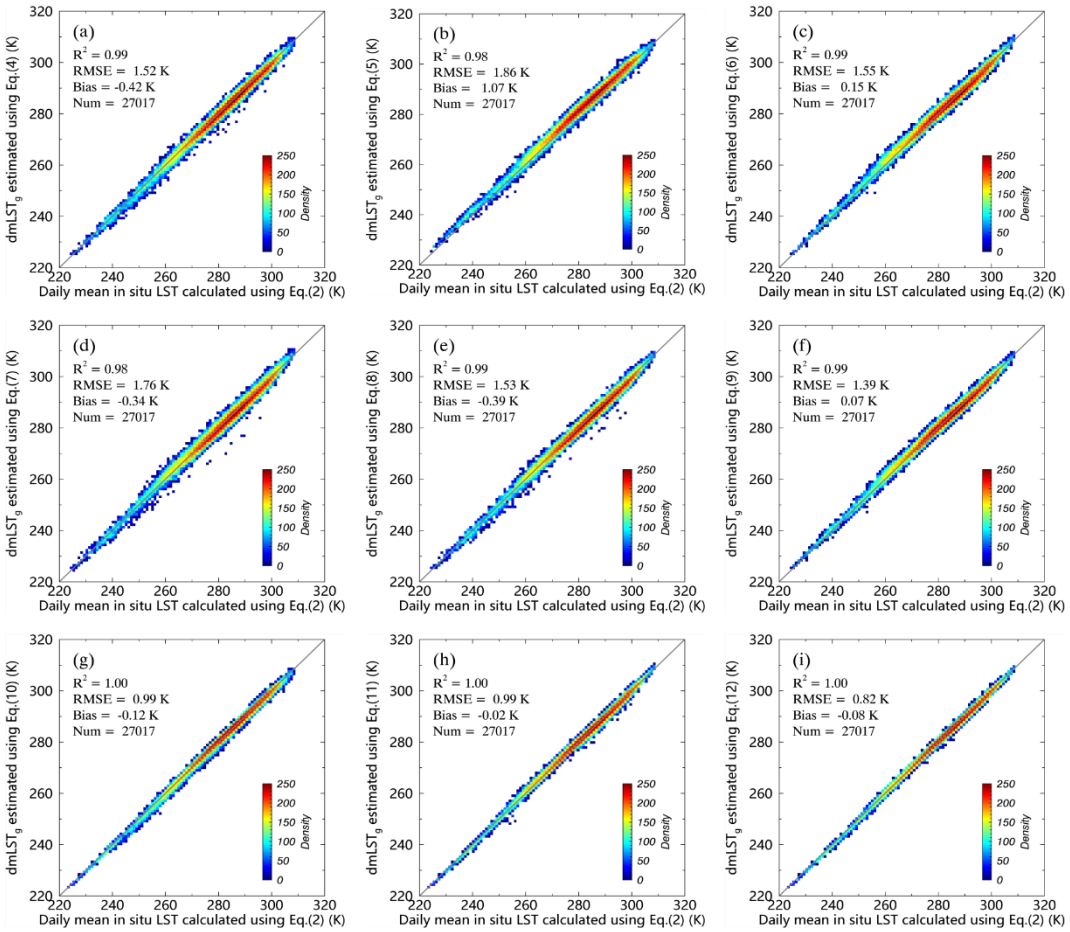


Fig. 9. Density scatter plots of daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with multiple linear regression method using Eqs. (4)-(12) at 77 sites from 2003 to 2018. The straight grey line is a 1:1 line.

We also validate the generation of the multiple linear regression method on the situation in which there are only two or three valid observations at MODIS observation times. Fig. 10 shows the in situ daily mean LST versus daily mean LST estimated by the multiple linear regression method for the situation in which there are only two or three valid observations of the all 235 sites from 2003 to 2018. The RMSE ranging from 1.50 to 2.16 K is found for the four combinations of only two valid LSTs. The combinations of only three valid LSTs show a good performance in estimating daily mean LST, with RMSE of 1.75, 1.47, 1.11, and 0.92 K, respectively, and bias of -0.71, -0.10, 0.23, and 0.01 K, respectively. Overall, the results are slightly poorer than the previous validation results at the spatial and temporal scale. The reason may be the influence of the cloud. The estimation accuracy is slightly higher when the clear sky conditions are better.

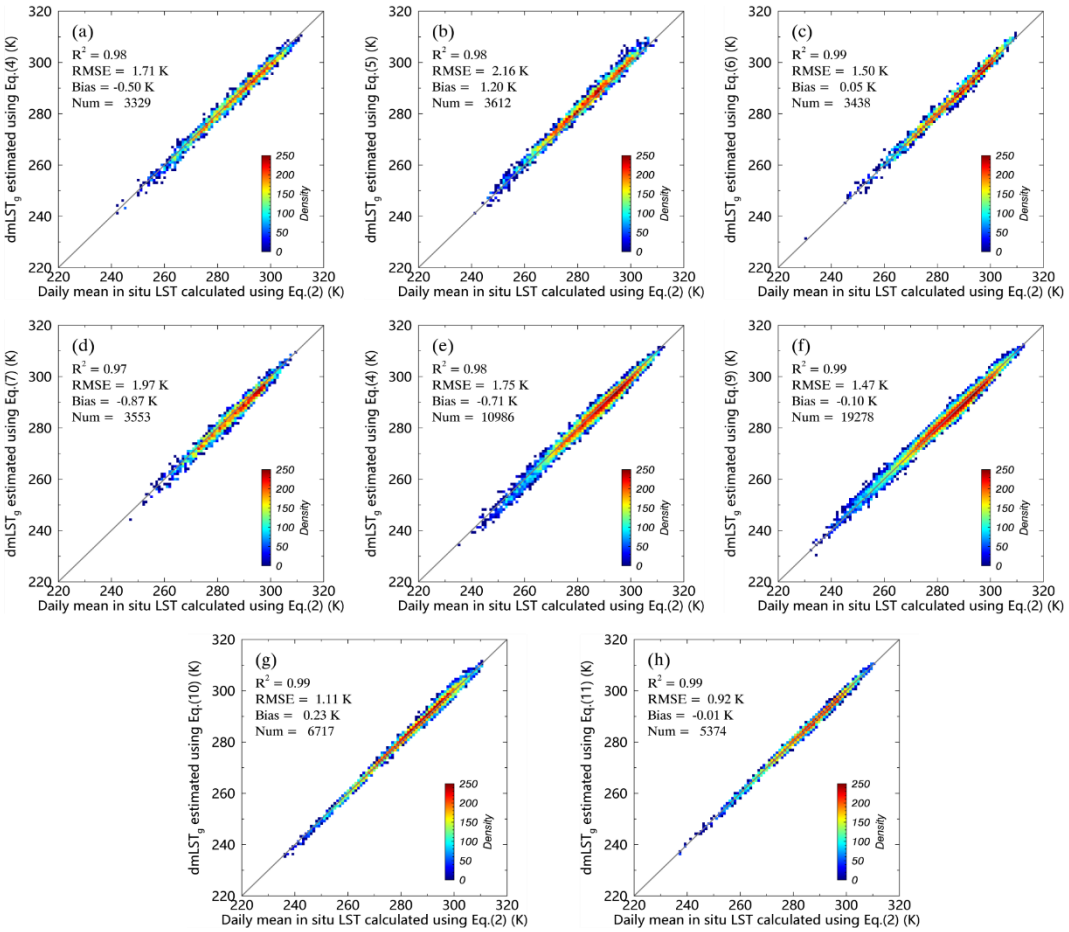


Fig. 10. Density scatter plots of daily mean in situ LST calculated using Eq. (2) versus daily mean in situ LST estimated with multiple linear regression method using Eqs. (4)-(11) at all 235 sites from 2003 to 2018. The straight grey line is a 1:1 line.

Fig. 11 illustrates the RMSEs of the nine linear combinations for calculating daily mean LST for different land cover types and climate zones using all temporal and spatial validation sites. Due to lack of in situ measurements, the deciduous coniferous forests (ID =3), Urban and build-up lands (ID=13), and farmland/natural vegetation mosaics (ID=14) are not included in Fig. 11(a). Overall, the performance of nine combinations varies with the different land cover types. For the combinations of T^dT^n and $T^nA^nT^d$, they perform poorly in the closed and opened shrublands (ID=6, 7). The all nine combinations present a good performance in evergreen broadleaf forests (ID=2), mixed forests (ID=5), permanent wetlands (ID=11), permanent snow and ice (ID=15), and they perform relatively poor in the barren (ID=16). As shown in Fig. 11(b), the nine combinations perform best in the tropical climate and show relatively a worse performance in the dry climate.

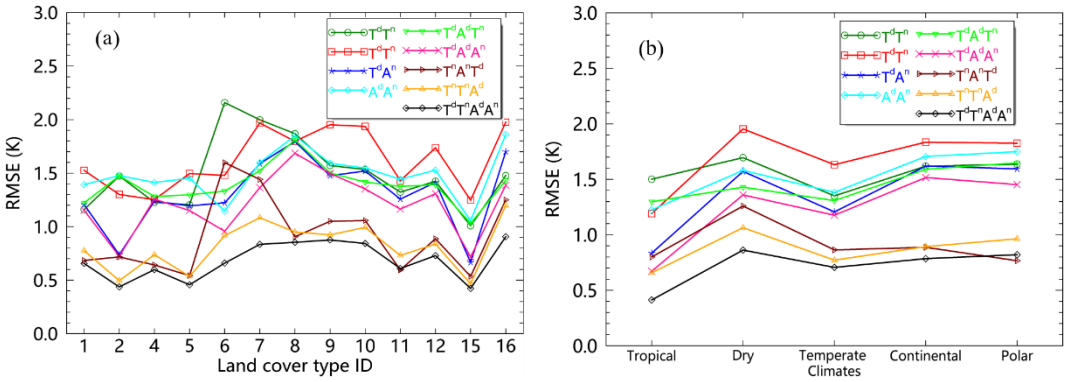


Fig. 11. RMSEs of the nine combinations for calculating daily mean LST for (a) different land cover types, and (b) different climates.

3.4. Estimation of daily mean LST from MODIS data

Fig. 12 shows the MODIS instantaneous LST versus the in situ LST at the seven SURFRAD sites in 2018. According to the validation results, the accuracy of the Aqua daytime LST is similar to that of the Terra daytime LST, with RMSEs of 3.88 and 3.52 K, respectively, and biases of 0.94 and -0.68 K, respectively. For the Terra and Aqua nighttime LST, the RMSEs are 3.16 and 2.95 K, respectively; the biases are 0.35 and 0.58 K, respectively. These errors and biases may due to the scale issue and the LST retrieval error. The ground-based (point) measurements can only represent a space of 1 -100 m^2 on the ground while the spatial resolution of MODIS pixels is about 1 km, which can represent a ground area of approximately 1 km^2 (Guillevic et al., 2012). Moreover, the validation process does not consider land surface uniformity. The validation results indicate that using ground-based measurement to validate the 1-km pixel has an uncertainty between 3.0 and 4.0 K.

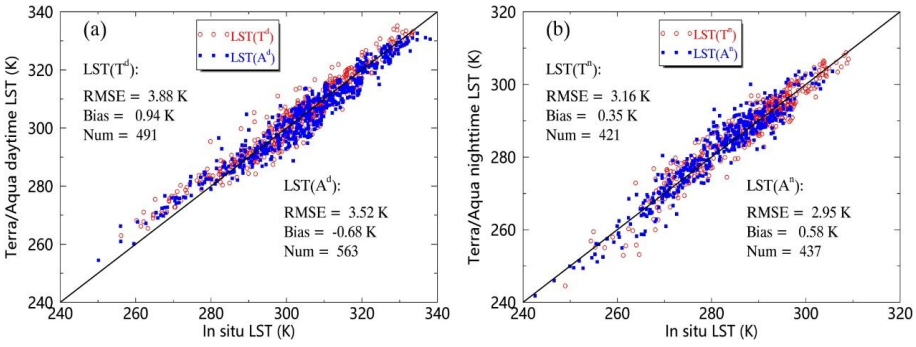


Fig. 12. Terra/Aqua (a) daytime and (b) nighttime LST versus in situ LST at seven SURFRAD sites in 2018. "*T*" and "*A*" represent Terra and Aqua satellites, respectively. Superscript "*d*" and "*n*" represent

daytime and nighttime, respectively. $LST(T^d)$ and $LST(A^d)$ are the Terra and Aqua daytime LST, respectively; $LST(T^n)$ and $LST(A^n)$ are the Terra and Aqua nighttime LST, respectively.

Fig. 13 shows the scatterplots of the daily mean in situ LST calculated using Eq. (2) versus the daily mean LST estimated from MODIS derived LSTs using Eqs. (4) -(12) in 2018 at the seven SURFRAD sites. The accuracies of the nine combinations for estimating the daily mean LST from MODIS derived LSTs are similar, with RMSE values ranging from 1.99 to 2.52 K, and a bias ranging from -1.43 to -0.52 K.

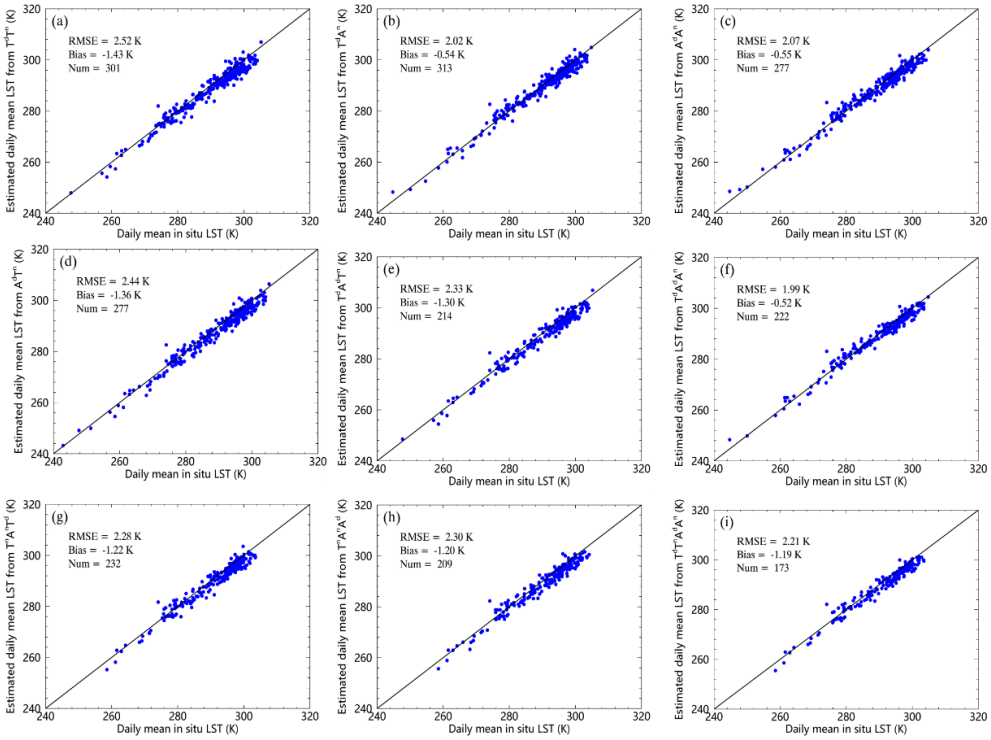
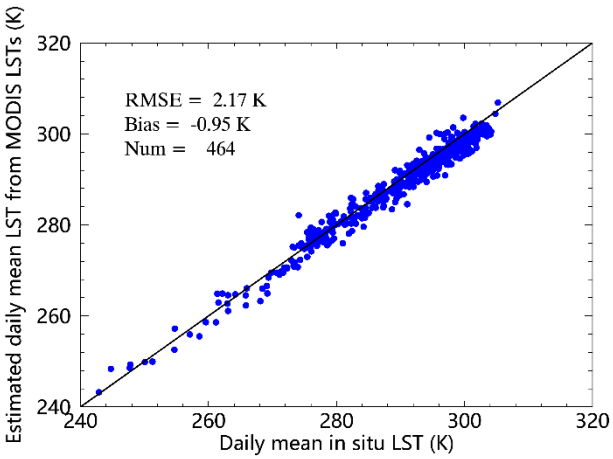


Fig. 13. Daily mean LST estimated using Eqs. (4)-(12) from the nine combinations of the MOD11A1 and MYD11A1 LSTs versus the daily mean in situ LST calculated with Eq. (2) in 2018 at the seven SURFRAD sites. "*T*" and "*A*" represent Terra and Aqua satellites, respectively. Superscript "*d*" and "*n*" represent daytime and nighttime, respectively.

Fig. 14 shows the scatterplots of the estimated daily mean LST using all available MODIS combinations with the diagram depicted in Fig. 4 versus the daily mean in situ LST calculated with Eq. (2) in 2018 at the seven SURFRAD sites. The results indicate that the daily mean LST estimated from the MOD11A1 and MYD11A1 products are consistent with the daily mean in situ LST, with an RMSE of 2.17 K and bias of -0.94 K. The large RMSE mainly comes from the discrepancy between the MODIS derived LSTs and the ground-based LST measurements as shown in Fig. 12.



565

566 **Fig. 14.** Daily mean LST estimated using the diagram depicted in Fig. 4 from the MOD11A1 and
567 MYD11A1 LSTs versus the daily in situ LST calculated using Eq. (2) in 2018 at the seven SURFRAD
568 sites.

569 ## 3.5. Application

570 Fig. 15 displays the spatial distribution of the daily mean LST at a global scale on the 1st of
571 January, April, July, and October 2018. The calculated daily mean LST reflected the broad
572 spatiotemporal variations in the LST. For example, the Sahara Desert and Middle East had the
573 highest daily mean LST for the selected days, whereas the daily mean LST in high-latitude
574 regions showed significant seasonal changes. Furthermore, in the northern hemisphere, the
575 daily mean LST in winter (i.e. Fig. 15(a)) and spring (i.e. Fig. 15(b)) is lower than that in
576 summer (i.e. Fig. 15(c)) and autumn (i.e. Fig. 15(d)). The results in the southern hemisphere
577 are similar. For example, the daily mean LST in Australia and Antarctica in winter (i.e. Fig.
578 15(c)) is lower than that observed other seasons. We note that there are abundant missing data
579 for the daily mean LST, especially in tropical areas due to cloud contamination and
580 polar-orbiting satellite configuration.

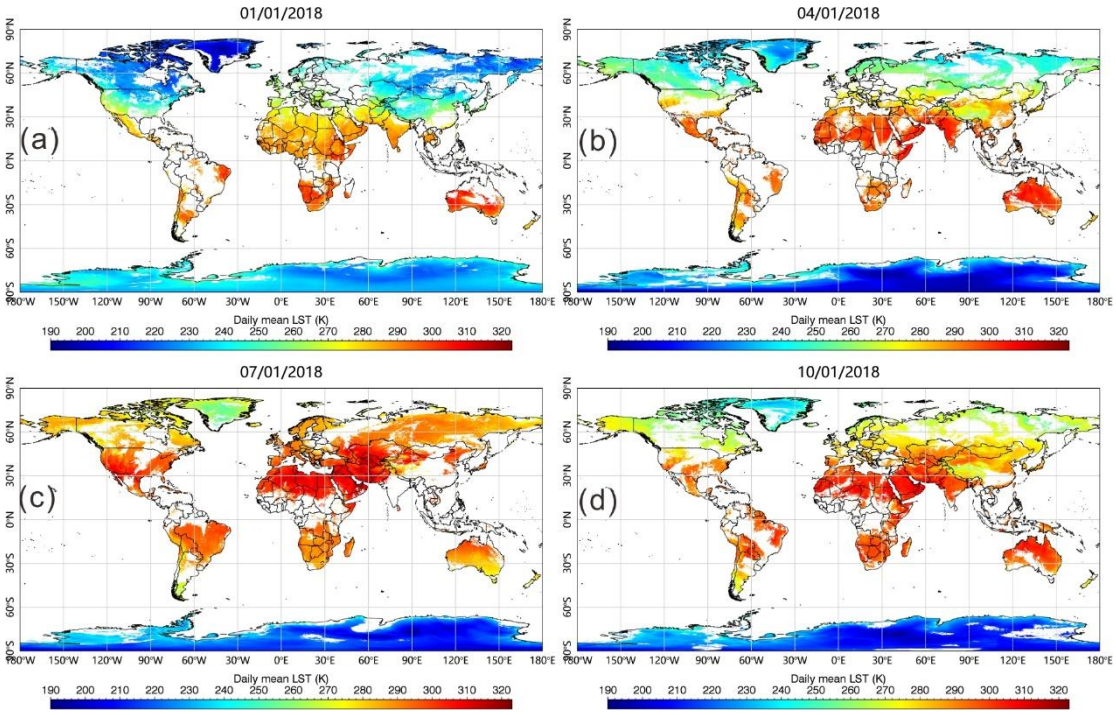


Fig. 15. Spatial distribution of the daily mean LST estimated using the framework proposed in Fig. 4 at the global scale on 1st of January, April, July, and October in 2018.

Fig. 16 displays the ATC fitting results of the daily mean LST at two randomly selected pixels in Northern and Southern hemispheres. The calculated ACPs (i.e., a, b, and c) represent the annual mean LST, annual amplitude, and annual phase, respectively. The daily mean LST reaches its maximum value in a year for the two sites in the Northern hemisphere (see Fig. 16(a)) and Southern hemispheres (see Fig. 16(b)) approximately on the 199th and 11th day of the year, respectively. Although numerous daily mean LST values were unknown throughout 2018 due to cloud contamination, the ATC model accurately describes the annual variation in the daily mean LST and yields representative and informative ACPs.

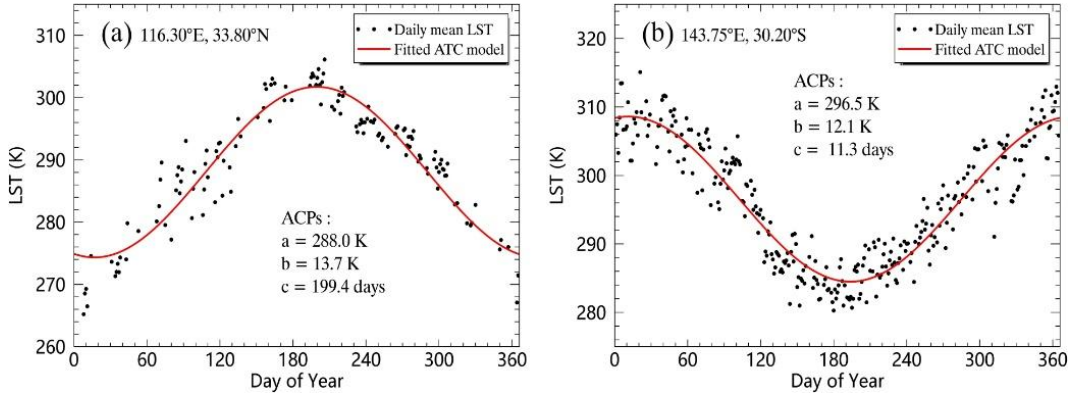


Fig. 16. ATC fitting of the daily mean LSTs in 2018 at two randomly selected sites in the (a) Northern

hemisphere and (b) Southern hemisphere.

We used the ATC model to globally calculate the ACPs from the daily mean LST in 2018. Fig. 17(a) -(c) displays the global annual mean LST, amplitude, and phase in 2018. Overall, the annual mean LST decreases with an increasing latitude, although it is highly related to the elevation patterns. For example, the annual mean LSTs in the Tibetan Plateau of China and Andes Mountains of South America are substantially lower than those in the same latitudes. As shown in Fig. 17(b), the annual amplitude values are the smallest in the tropics and increase with latitude. The highest values of the annual amplitude are in Siberian regions, which can reach 30.0 K. The annual phase occurs on approximately the 180th day of the year in the Northern hemisphere; meanwhile, in the Southern hemisphere, the annual phase values are observed at the beginning or end of the year. Previous studies showed that the annual phase is mainly related to the phenology of vegetation or the land cover types (Stine et al., 2009; Wang and Dillon 2014). Fig. 17(d) displays the RGB composite of the annual mean LST, amplitude, and phase. If the interpretation of the single ACP is more difficult, the RGB composite may give a good intuitive impression of the thermal characteristics of the landscapes (Bechtel 2015). However, the reasons for the spatial distribution of these ACPs are unknown at this stage, such that further analysis is required. Moreover, understanding the spatiotemporal changes in these ACPs can be used to analyze global climate change (Stine et al., 2009).

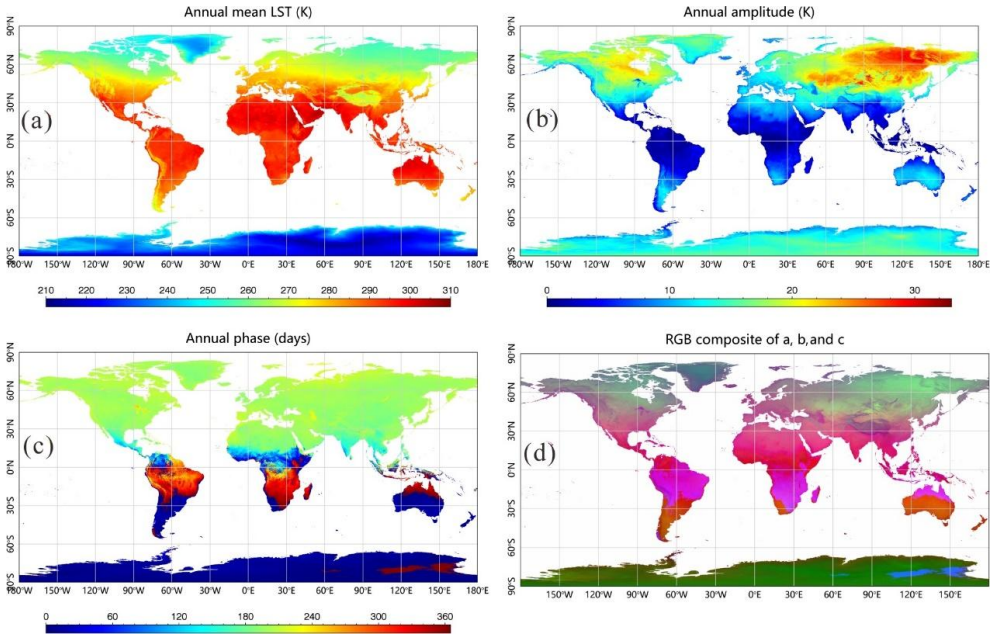


Fig. 17. Global ATC modeling with the daily mean MODIS LST in2018. (a) Annual mean LST (K), (b) annual amplitude (K), (c) annual phase (days), and (d) RGB composite of (a), (b), and (c).

4. Discussions

4.1. The effect of cloudy conditions

The multiple linear regression method proposed in this study only considers the cloud free conditions at four moments when the MODIS sensor passes to estimate the daily mean LST. The results show that it can achieve a good performance. The cloud conditions of the day would affect the accuracy of the daily average surface temperature estimation. Because the SURFRAD observations with finest temporal resolution (one or three minutes) are temporally continuous, they can be used to evaluate the influence of cloud conditions on estimating daily mean LST. In this study, we divide the in situ measurements of the seven SURFRAD sites (BND, DRA, FPK, GWN, PSU, SXF, and TBL sites) from 2003 to 2018 into two parts: one is the all clear-sky condition in a day, whereas the other is cloudy condition in a day. In order to compare the accuracy of estimated daily mean LST under the two conditions, only the data with four valid LSTs at MODIS view times of the day were selected.

To determine whether it is a clear sky or cloudy condition, we use a Clear-Sky Index (CSI) proposed by Marty and Philipona (Marty and Philipona 2000). CSI can be calculated by the follow equations:

$$CSI = \varepsilon_A / \varepsilon_{AC} \tag{20}$$

$$\varepsilon_A = R_{\downarrow} / (\sigma T_a^4) \tag{21}$$

$$\varepsilon_{AC} = \varepsilon_{AD} + k(e_a / T_a)^{1/8} \tag{22}$$

where σ is the Stephan-Boltzmann constant, T_a is the air temperature (K), ε_A indicates the apparent emittance of the sky, and ε_{AC} indicates an empirical apparent cloud-free emittance, $R_{\downarrow}$ is longwave radiation (W/m^2). ε_{AD} is an altitude-dependent emittance of a completely dry atmosphere, the value of ε_{AD} is taken as 0.22, e_a is water vapor pressure (Pa), and k is a constant location-dependent coefficient, which equals to 0.435 in this study. Water vapor pressure e_a is calculated from the relative humidity using the following equations:

$$e = \frac{RH}{100} e_s \tag{23}$$

$$e_s = 6.1121 \exp\left(\frac{17.502 t_a}{t_a + 240.97}\right) \tag{24}$$

The air temperature t_a has to be in ℃, RH is the relative humidity. CSI < 1 means clear-sky conditions, and CSI ≥ 1 means cloudy condition (Marty and Philipona 2000). Here, we defined that the clear-sky days are those CSI values of the day are less than 1, otherwise, it is cloudy condition. Fig. 18 shows the number of clear-sky and cloudy days at the seven SURFRAD sites from 2003 to 2018. Obviously, the number of cloudy days is greater than the clear-sky days at the seven sites.

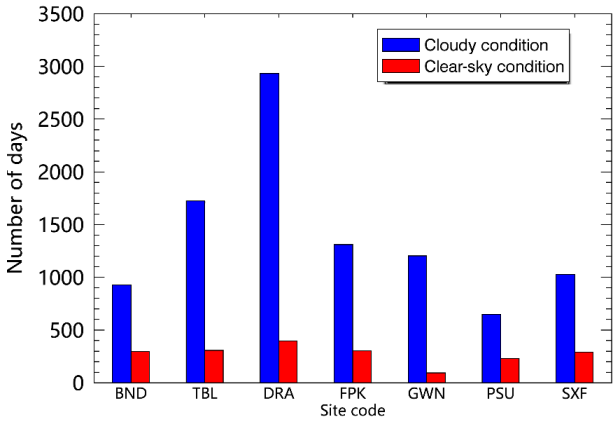


Fig. 18. Number of clear-sky and cloudy days at the seven SURFRAD sites from 2003 to 2018.

In situ measurements at BND, TBL, DRA, FPK, GWN, PSU, and SXF sites were used to evaluate the cloud effect on the estimation of daily mean LST. Fig. 19 shows the values of RMSE and bias for estimating daily mean LST under clear-sky and cloudy conditions at the seven SURFRAD sites from 2003 to 2018. RMSE of the cloudy condition is clearly larger than that of the clear-sky condition over DRA, FPK, GWN, PSU, and SXF sites. For BND and TBL sites, the RMSE of the cloudy condition is slightly lower than that of clear-sky condition. In terms of bias values, the bias values at BND, TBL, PSU, and SXF sites under clear-sky days are larger than that of cloudy days. For DRA, FPK, and GWN sites, the difference of the bias values is small. Overall, the multiple linear regression method proposed in this study performs slightly better under all clear-sky conditions than that of cloudy conditions.

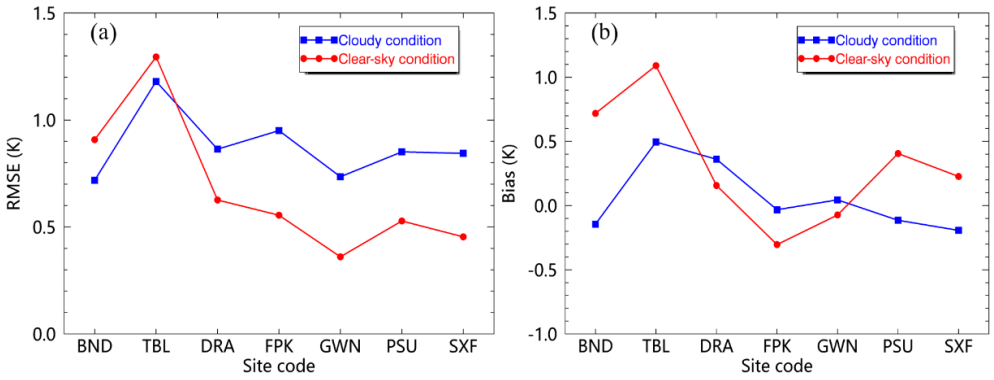


Fig. 19. (a) RMSE, and (b) bias values for estimating daily mean LST using Eq. (12) under clear-sky (red polyline) and cloudy conditions (blue polyline) at seven SURFRAD sites from 2003 to 2018.

4.2. Limitations

We have proposed and validated the linear combinations of the daytime and nighttime instantaneous LSTs to estimate the daily mean LST. Compared with the traditional average method for estimating daily mean temperature, the method proposed in this paper provides more combinations and higher accuracy. However, this study still has potential limitations. First, because there are fewer in situ sites in Africa and South America and some in situ sites cannot measure upwelling and downwelling long-wave radiation, the performance of model proposed in this study may be decreased in these regions. In the future, we will collect more ground measurements or the continuous observations of geostationary satellite to establish a more robust combination model and obtain more reliable model coefficients. Second, it is inevitable that the measurements in space by thermal infrared sensors are affected by cloud coverage, especially in tropical regions, there are many missing data when estimating the daily mean LST with MODIS LSTs based on the framework proposed in this paper. In contrast, passive microwaves can penetrate clouds and obtain the thermal information on the land surface under the clouds In contrast, passive microwaves can penetrate clouds and obtain the thermal information on the land surface under the clouds (Duan et al., 2020; Holmes et al., 2009; Huang et al., 2019; Prigent et al., 2016; Yoo et al., 2020). Therefore, either spatial downscaling of land surface temperatures derived from passive microwave data (Yoo et al., 2020), or combination of thermal infrared data and microwave data (Duan et al., 2017; Sun et al., 2019; Zhang et al., 2019; .Zhang et al., 2020) would help to produce the continuous daily mean LST products in time and space. Third, the model coefficients calculated in this study

are only applicable to MODIS data, for other polar orbiting satellite data, the model coefficients have to be recalculated. Overall, the linear combination method proposed in this paper provides more combinations for estimating the daily mean LST and can obtain reliable accuracy, but more in situ measurements needed to be collected to build a more robust model that can be applied globally in the future studies.

5. Conclusions

Accurate estimations of the daily mean LST is meaningful for a wide range of applications. In this study, the relationship between the daily mean LSTs estimated from different combinations at MODIS day/night observation times was evaluated using a large number of in situ LST measurements worldwide. A practical framework was proposed to retrieve the daily mean LST using daily LST products derived from MODIS and the potential application of satellited derived daily mean LST was also verified using an ATC model.

The results showed that the multiple linear regression of at least one daytime and one nighttime observation can provide reliable estimates of the daily mean LST under all-weather conditions, with an R^2 value greater than 0.99 and RMSE value less than 1.60 K. The multiple linear regression of four complete observations (i.e., $T^dT^nA^dA^n$) per day was the most accurate, with an RMSE of 0.80 K, which was followed by the multiple linear combination of three valid observations (i.e., combination of two daytime and one nighttime observations and combinations of one daytime and two nighttime observations) with RMSE values ranging from 0.91 to 1.51 K. The linear combination of one daytime and one nighttime observation (i.e., T^dT^n, T^dA^n, A^dA^n, and A^dT^n) performed slightly worse, with RMSE values ranging from 1.50 to 1.60 K. The validation results from the spatial and temporal scale show that more combinations of valid LSTs could produce more reliable accuracy. The accuracies obtained using nine combinations of observations with the MOD11A1 and MYD11A1 products to separately estimate the daily mean LST were similar, with RMSEs ranging from 1.99 to 2.52 K. Cloud conditions will affect the estimation accuracy of daily mean LST. Evaluation results showed that the multiple linear regression method proposed in this paper perform slightly better under all clear-sky conditions than that of cloudy conditions. Based on the multiple

linear regression method, a practical framework was proposed to produce the most complete coverage of the daily mean LSTs from the MODIS LST products. The validation results indicate that the daily mean LSTs estimated using the proposed framework at seven SURFRAD sites in 2018 were consistent with the daily mean in situ LST, with an RMSE of 2.17 K and bias of -0.95 K. Finally, based on the above estimation of the daily mean LSTs from the MOD11C1 and MYD11C1 LSTs in 2018, an ATC model was used to simulate the annual variation in the daily mean LST, which successfully yielded representative and informative ACPs. Due to cloud contamination, there are still numerous missing data when estimating the daily mean LST based on MODIS data. Combined with LST data from other thermal infrared sensors or microwave sensors, the daily mean LST could be obtained for a continuous time series. We expect our findings will play an important role in various applications involving climate monitoring and land-climate interactions.

Acknowledgements

This study was supported by the China Scholar Council (CSC) under Grant number 201800320098. The authors would like to thank the MODIS Science teams at NASA for providing the satellite observation data. We also acknowledge the six flux networks, i.e., SURFRAD, AmeriFlux, OzFlux, Euroflux, AsiaFlux, and ChinaFlux, for providing in situ measurements.

References

Amiri, R., Weng, Q., Alimohammadi, A., & Alavipanah, S.K. (2009). Spatial–temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. *Remote sensing of environment, 113*, 2606-2617

Anderson, M., Norman, J., Kustas, W., Houborg, R., Starks, P., & Agam, N. (2008). A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales. *Remote sensing of environment, 112*, 4227-4241

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., Moncrieff, J., Foken, T., Kowalski, A.S., Martin, P.H., Berbigier, P., & Bernhofer, C. (1999). Estimates of the annual net carbon and water exchange of

743 forests: the EUROFLUX methodology. *Advances in ecological research* (pp. 113-175): Elsevier
744 Augustine, J.A., DeLuisi, J.J., & Long, C.N. (2000). SURFRAD–A national surface radiation budget
745 network for atmospheric research. *Bulletin of the American Meteorological Society, 81*, 2341-2358
746 Bechtel, B. (2011). Multitemporal Landsat data for urban heat island assessment and classification of
747 local climate zones. In, *2011 Joint Urban Remote Sensing Event* (pp. 129-132): IEEE
748 Bechtel, B. (2015). A new global climatology of annual land surface temperature. *Remote Sensing, 7*,
749 2850-2870
750 Bechtel, B., & Sismanidis, P. (2018). Time series analysis of moderate resolution land surface
751 temperatures. *Remote Sensing Time Series Image Processing* (pp. 111-142): CRC Press
752 Becker, F., & Li, Z.-L. (1990). Towards a local split window method over land surfaces. *Remote*
753 *Sensing, 11*, 369-393
754 Beringer, J., Hutley, L.B., McHugh, I., Arndt, S.K., Campbell, D., Cleugh, H.A., Cleverly, J., De Dios,
755 V.R., Eamus, D., & Evans, B. (2016). An introduction to the Australian and New Zealand flux
756 tower network-OzFlux. *Biogeosciences*
757 Boden, T.A., Krassovski, M., & Yang, B. (2013). The AmeriFlux data activity and data system: an
758 evolving collection of data management techniques, tools, products and services. *Geoscience*
759 *Instrumentation, Methods, and Data Systems, 2*, 165-176
760 Chen, X., Su, Z., Ma, Y., Cleverly, J., & Liddell, M. (2017). An accurate estimate of monthly mean land
761 surface temperatures from MODIS clear-sky retrievals. *Journal of hydrometeorology, 18*,
762 2827-2847
763 Dall'Amico, M., & Hornsteiner, M. (2006). A simple method for estimating daily and monthly mean
764 temperatures from daily minima and maxima. *International Journal of Climatology: A Journal of*
765 *the Royal Meteorological Society, 26*, 1929-1936
766 Duan, S.-B., Han, X.-J., Huang, C., Li, Z.-L., Wu, H., Qian, Y., Gao, M., & Leng, P. (2020). Land
767 Surface Temperature Retrieval from Passive Microwave Satellite Observations: State-of-the-Art
768 and Future Directions. *Remote Sensing, 12*, 2573
769 Duan, S.-B., Li, Z.-L., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., & Coll, C.
770 (2019). Validation of Collection 6 MODIS land surface temperature product using in situ
771 measurements. *Remote sensing of environment, 225*, 16-29

Duan, S.-B., Li, Z.-L., Wang, N., Wu, H., & Tang, B.-H. (2012). Evaluation of six land-surface diurnal temperature cycle models using clear-sky in situ and satellite data. *Remote sensing of environment, 124*, 15-25

Duan, S.-B., Li, Z.-L., Wu, H., Tang, B.-H., Jiang, X., & Zhou, G. (2013). Modeling of day-to-day temporal progression of clear-sky land surface temperature. *IEEE Geoscience and Remote Sensing Letters, 10*, 1050-1054

Fu, P., & Weng, Q. (2018). Variability in annual temperature cycle in the urban areas of the United States as revealed by MODIS imagery. *ISPRS journal of photogrammetry and remote sensing, 146*, 65-73

Göttsche, F.-M., & Olesen, F.S. (2001). Modelling of diurnal cycles of brightness temperature extracted from METEOSAT data. *Remote sensing of environment, 76*, 337-348

Guillevic, P.C., Privette, J.L., Coudert, B., Palecki, M.A., Demarty, J., Ottle, C., & Augustine, J.A. (2012). Land Surface Temperature product validation using NOAA's surface climate observation networks—Scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS). *Remote sensing of environment, 124*, 282-298

Heidinger, A.K., Laszlo, I., Molling, C.C., & Tarpley, D. (2013). Using SURFRAD to verify the NOAA single-channel land surface temperature algorithm. *Journal of Atmospheric and Oceanic Technology, 30*, 2868-2884

Hu, L., Sun, Y., Collins, G., & Fu, P. (2020). Improved estimates of monthly land surface temperature from MODIS using a diurnal temperature cycle (DTC) model. *ISPRS journal of photogrammetry and remote sensing, 168*, 131-140

Hu, X., Ren, H., Tansey, K., Zheng, Y., Ghent, D., Liu, X., & Yan, L. (2019). Agricultural drought monitoring using European Space Agency Sentinel 3A land surface temperature and normalized difference vegetation index imageries. *Agricultural and Forest Meteorology, 279*, 107707

Huang, C., Li, X., & Lu, L. (2008). Retrieving soil temperature profile by assimilating MODIS LST products with ensemble Kalman filter. *Remote sensing of environment, 112*, 1320-1336

Huang, F., Zhan, W., Voogt, J., Hu, L., Wang, Z., Quan, J., Ju, W., & Guo, Z. (2016). Temporal upscaling of surface urban heat island by incorporating an annual temperature cycle model: A tale of two cities. *Remote sensing of environment, 186*, 1-12

Huang, R., Zhang, C., Huang, J., Zhu, D., Wang, L., & Liu, J. (2015). Mapping of daily mean air temperature in agricultural regions using daytime and nighttime land surface temperatures derived from TERRA and AQUA MODIS data. *Remote Sensing, 7*, 8728-8756

Jiang, L., & Islam, S. (1999). A methodology for estimation of surface evapotranspiration over large areas using remote sensing observations. *Geophysical Research Letters, 26*, 2773-2776

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., & Woollen, J. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society, 77*, 437-472

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

Leng, P., Li, Z.-L., Duan, S.-B., Gao, M.-F., & Huo, H.-Y. (2017). A practical approach for deriving all-weather soil moisture content using combined satellite and meteorological data. *ISPRS journal of photogrammetry and remote sensing, 131*, 40-51

Lenssen, N.J., Schmidt, G.A., Hansen, J.E., Menne, M.J., Persin, A., Ruedy, R., & Zyss, D. (2019). Improvements in the GISTEMP uncertainty model. *Journal of Geophysical Research: Atmospheres, 124*, 6307-6326

Li, H., Sun, D., Yu, Y., Wang, H., Liu, Y., Liu, Q., Du, Y., Wang, H., & Cao, B. (2014). Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China. *Remote sensing of environment, 142*, 111-121

Li, X., Li, X., Li, Z., Ma, M., Wang, J., Xiao, Q., Liu, Q., Che, T., Chen, E., & Yan, G. (2009). Watershed allied telemetry experimental research. *Journal of Geophysical Research: Atmospheres, 114*

Li, Y., Zhao, M., Motesharrei, S., Mu, Q., Kalnay, E., & Li, S. (2015). Local cooling and warming effects of forests based on satellite observations. *Nature communications, 6*, 6603

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., & Sobrino, J.A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote sensing of environment, 131*, 14-37

Liu, S., Xu, Z., Zhu, Z., Jia, Z., & Zhu, M. (2013). Measurements of evapotranspiration from

eddy-covariance systems and large aperture scintillometers in the Hai River Basin, China. *Journal of Hydrology, 487*, 24-38

Liu, X., Tang, B.-H., Yan, G., Li, Z.-L., & Liang, S. (2019a). Retrieval of global orbit drift corrected land surface temperature from long-term AVHRR data. *Remote Sensing, 11*, 2843

Liu, Z., Zhan, W., Lai, J., Hong, F., Quan, J., Bechtel, B., Huang, F., & Zou, Z. (2019b). Balancing prediction accuracy and generalization ability: A hybrid framework for modelling the annual dynamics of satellite-derived land surface temperatures. *ISPRS journal of photogrammetry and remote sensing, 151*, 189-206

Marty, C., & Philipona, R. (2000). The clear- sky index to separate clear- sky from cloudy- sky situations in climate research. *Geophysical Research Letters, 27*, 2649-2652

Nordli, P., Alexandersson, H., Frisch, P., Førland, E., Heino, R., Jónsson, T., Steffensen, P., Tuomenvirta, H., & Tveito, O. (1996). The effect of radiation screens on Nordic temperature measurements. *DNMI Report, 4*, 96

Novick, K.A., Biederman, J., Desai, A., Litvak, M., Moore, D.J., Scott, R., & Torn, M. (2018). The AmeriFlux network: A coalition of the willing. *Agricultural and Forest Meteorology, 249*, 444-456

Ogawa, K., Schmugge, T., Jacob, F., & French, A. (2002). Estimation of broadband land surface emissivity from multi-spectral thermal infrared remote sensing

Ogawa, K., Schmugge, T., & Rokugawa, S. (2008). Estimating broadband emissivity of arid regions and its seasonal variations using thermal infrared remote sensing. *IEEE Transactions on Geoscience and Remote Sensing, 46*, 334-343

Ouyang, B., Che, T., DAI, L.-y., & WANG, Z.-y. (2012). Estimating Mean Daily Surface Temperature over the Tibetan Plateau Based on MODIS LST Products [J]. *Journal of Glaciology and Geocryology, 2*

Qin, J., Yang, K., Lu, N., Chen, Y., Zhao, L., & Han, M. (2013). Spatial upscaling of in-situ soil moisture measurements based on MODIS-derived apparent thermal inertia. *Remote sensing of environment, 138*, 1-9

Rao, Y., Liang, S., Wang, D., Yu, Y., Song, Z., Zhou, Y., Shen, M., & Xu, B. (2019). Estimating daily average surface air temperature using satellite land surface temperature and top-of-atmosphere

radiation products over the Tibetan Plateau. *Remote sensing of environment, 234*, 111462

Sandholt, I., Rasmussen, K., & Andersen, J. (2002). A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote sensing of environment, 79*, 213-224

Semenov, M.A., & Stratonovitch, P. (2010). Use of multi-model ensembles from global climate models for assessment of climate change impacts. *Climate Research, 41*, 1-14

Sobrino, J.A., & Irakulis, I. (2020). A Methodology for Comparing the Surface Urban Heat Island in Selected Urban Agglomerations Around the World from Sentinel-3 SLSTR Data. *Remote Sensing, 12*, 2052

Son, N.T., Chen, C., Chen, C., Chang, L., & Minh, V.Q. (2012). Monitoring agricultural drought in the Lower Mekong Basin using MODIS NDVI and land surface temperature data. *International Journal of Applied Earth Observation and Geoinformation, 18*, 417-427

Stine, A.R., Huybers, P., & Fung, I.Y. (2009). Changes in the phase of the annual cycle of surface temperature. *Nature, 457*, 435-440

Townshend, J., Justice, C., Skole, D., Malingreau, J.-P., Cihlar, J., Teillet, P., Sadowski, F.a., & Ruttenberg, S. (1994). The 1 km resolution global data set: needs of the International Geosphere Biosphere Programme. *International Journal of Remote Sensing, 15*, 3417-3441

Tran, D.X., Pla, F., Latorre-Carmona, P., Myint, S.W., Caetano, M., & Kieu, H.V. (2017). Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS journal of photogrammetry and remote sensing, 124*, 119-132

Trenberth, K.E., Anthes, R.A., Belward, A., Brown, O.B., Habermann, T., Karl, T.R., Running, S., Ryan, B., Tanner, M., & Wielicki, B. (2013). Challenges of a sustained climate observing system. *Climate Science for Serving Society* (pp. 13-50): Springer

Van den Bergh, F., Van Wyk, M., & Van Wyk, B. (2006). Comparison of data-driven and model-driven approaches to brightness temperature diurnal cycle interpolation

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*, 449-465

Vincent, L.A., Wang, X.L., Milewska, E.J., Wan, H., Yang, F., & Swail, V. (2012). A second generation

of homogenized Canadian monthly surface air temperature for climate trend analysis. *Journal of Geophysical Research: Atmospheres, 117*

Wan, Z. (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote sensing of environment, 140*, 36-45

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on Geoscience and Remote Sensing, 34*, 892-905

Wan, Z., & Li, Z.-L. (1997). A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Transactions on Geoscience and Remote Sensing, 35*, 980-996

Wang, G., & Dillon, M.E. (2014). Recent geographic convergence in diurnal and annual temperature cycling flattens global thermal profiles. *Nature Climate Change, 4*, 988-992

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*, 1556-1565

Wang, T., Shi, J., Ma, Y., Letu, H., & Li, X. (2020). All-sky longwave downward radiation from satellite measurements: General parameterizations based on LST, column water vapor and cloud top temperature. *ISPRS journal of photogrammetry and remote sensing, 161*, 52-60

Wang, W., Liang, S., & Meyers, T. (2008). Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote sensing of environment, 112*, 623-635

Warren, R. (2006). Impacts of global climate change at different annual mean global temperature increases. *Avoiding dangerous climate change, 93*, 93-94

Weng, Q., & Fu, P. (2014a). Modeling annual parameters of clear-sky land surface temperature variations and evaluating the impact of cloud cover using time series of Landsat TIR data. *Remote sensing of environment, 140*, 267-278

Weng, Q., & Fu, P. (2014b). Modeling diurnal land temperature cycles over Los Angeles using downscaled GOES imagery. *ISPRS journal of photogrammetry and remote sensing, 97*, 78-88

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote sensing of environment, 89*, 467-483

Williamson, S.N., Hik, D.S., Gamon, J.A., Kavanaugh, J.L., & Flowers, G.E. (2014). Estimating

temperature fields from MODIS land surface temperature and air temperature observations in a sub-arctic alpine environment. *Remote Sensing*, *6*, 946-963

Wooster, M.J., Xu, W., & Nightingale, T. (2012). Sentinel-3 SLSTR active fire detection and FRP product: Pre-launch algorithm development and performance evaluation using MODIS and ASTER datasets. *Remote sensing of environment*, *120*, 236-254

Xing, Z., Yu, Y., Duan, S.-B., Li, Z.-L., Gao, M., Leng, P., Zhang, X., & Shang, G. (2020). Modeling Year-to-Year Variations of Clear-Sky Land Surface Temperature Using Aqua/MODIS Data. *IEEE Access*, *8*, 114541-114553

Yoo, C., Im, J., Cho, D., Yokoya, N., Xia, J., & Bechtel, B. (2020). Estimation of all-weather 1 km MODIS land surface temperature for humid summer days. *Remote Sensing*, *12*, 1398

Zhang, X., Zhou, J., Liang, S., Chai, L., Wang, D., & Liu, J. (2020). Estimation of 1-km all-weather remotely sensed land surface temperature based on reconstructed spatial-seamless satellite passive microwave brightness temperature and thermal infrared data. *ISPRS journal of photogrammetry and remote sensing*, *167*, 321-344

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

Zou, Z., Zhan, W., Liu, Z., Bechtel, B., Gao, L., Hong, F., Huang, F., & Lai, J. (2018). Enhanced modeling of annual temperature cycles with temporally discrete remotely sensed thermal observations. *Remote Sensing*, *10*, 650