Determination of global land surface temperature using data from only five selected thermal infrared channels: Method extension and accuracy assessment

Xiaopo Zheng, Zhao-Liang Li, Tianxing Wang, Huabing Huang, Françoise Nerry

To cite this version:

Xiaopo Zheng, Zhao-Liang Li, Tianxing Wang, Huabing Huang, Françoise Nerry. Determination of global land surface temperature using data from only five selected thermal infrared channels: Method extension and accuracy assessment. Remote Sensing of Environment, 2022, 268, pp.112774. 10.1016/j.rse.2021.112774 . hal-03835395

HAL Id: hal-03835395
https://hal.science/hal-03835395v1

Submitted on 4 Nov 2022

Determination of global land surface temperature using data from only five selected thermal infrared channels: method extension and accuracy assessment

Xiaopo Zheng [a, c], Zhao-Liang Li [b, c, *], Tianxing Wang [a], Huabing Huang [a], Françoise Nerry [c]

[a] School of Geospatial Engineering and Science, Sun Yat-Sen University, 519082 Guangdong, China
[b] Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, 100081 Beijing, China
[c] ICube, UdS, CNRS, 300 Bld Sébastien Brant, CS 10413, 67412 Illkirch, France

Abstract:

Land surface temperature (LST) is an essential input for modeling the processes of energy exchange and balance of the earth's surface. Thermal infrared (TIR) remote sensing is considered to be the most efficient way to obtain accurate LST, both regionally and globally. Currently, many LST retrieval algorithms have been developed, including the up-to-date SW-TES (SW: split window; TES: temperature-emissivity separation) method, which is claimed to be able to accurately derive LST without the need for atmospheric information and land surface emissivity (LSE) based on the selected multiple TIR channel configuration. However, this hybrid method is actually not applicable to observations with large viewing angles and was only preliminarily evaluated in Australia. In this study, this method was extended for application to global TIR measurements with different viewing angles. Additionally, the performance of this extended SW-TES method was assessed globally for different seasons by using the MODIS LST product as a reference, and was also validated using in-situ LST measurements from the SURFRAD (SURFace RADiation budget network) sites. The results showed that the LST

---

[*] Corresponding author: lizl@unistra.fr

retrievals using the extended SW-TES method were comparable to the MODIS LST product, with discrepancies of <2.7 K and <1.8 K for global daytime and nighttime observations, respectively. Validations based on the SURFRAD in-situ LST measurements indicated that the extended method could be used to retrieve LST accurately with a root-mean-square error (RMSE) of approximately 3.6 K during the daytime and 2.4 K during the nighttime. However, special attention should be paid when applying the extended method to daytime observations on grasslands and shrublands during hot seasons, considering the relatively large discrepancy when using this method compared with that obtained with the MODIS LST product (>4.0 K). Overall, in this study, the SW-TES method was extended, and the performance was comprehensively evaluated at the global scale, which may help in facilitating its potential applications.

Keywords: Thermal infrared data, land surface temperature (LST), split window, temperature-emissivity separation, accuracy assessment

1. Introduction

Land surface temperature (LST) is an important factor that influences water and energy exchange between the Earth’s surface and the atmosphere (Kappas and Phan, 2018; Sobrino et al., 2016; Tomlinson et al., 2011). It has been widely used in many studies, including climatology, meteorology, hydrology, and ecology (Anderson et al., 2008; Kappas and Phan, 2018; Li et al., 2009; O'Connor et al., 2020; Yao et al., 2019). Therefore, LST has been listed as a high-priority parameter for the International Geosphere-Biosphere Programme (IGBP)

(Townshend et al., 1994). It has also been approved as an important Earth Surface Data Record (ESDR) by NASA and is defined as an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS) (O'Connor et al., 2020).

Currently, thermal infrared (TIR) remote sensing is considered the most appropriate method for obtaining accurate LST, both regionally and globally (Li et al., 2013a; Sánchez-Aparicio et al., 2020). Over several decades, a wide variety of LST retrieval algorithms have been developed and can be roughly classified into four categories as follows (Li et al., 2013a; Li et al., 2013b; Sattari and Hashim, 2014). First, the single-channel method, which requires both atmospheric profiles and LSE to obtain the LST from the observed TIR radiance (Duan et al., 2019b; Jimenez-Munoz and Sobrino, 2003; Ottlé and Vidal-Madjar, 1992; Qin et al., 2001b) (Jimenez-Munoz et al., 2009; Sánchez-Aparicio et al., 2020). Second, the split window (SW) method that can eliminate atmospheric effects by directly using the observations of two adjacent channels (Becker and Li, 1990; Coll and Caselles, 1997; McMillin, 1975; Qin et al., 2001a; Wan and Dozier, 1996; Wang et al., 2019b). But the performance of this method is shown to be sensitive to the channel LSE uncertainties, implying that accurate LSE of the study area must be known in advance (Jiang and Li, 2008; Wang et al., 2019a). Third, the temperature/emissivity separation (TES) method that can be used to simultaneously separate LST and LSE (Gillespie et al., 1998). However, accurate atmospheric correction is needed before applying this method (Hulley and Hook, 2011; Hulley et al., 2018; Ren et al., 2020). Four, the day/night methods that could also be used to retrieve LST and LSE at the same time (Wan and Li, 1997; Wan, 2008; Wan, 2014). By using two geographically matched images and assuming the LSE is unchanged in the two images, this method requires only the shape information of atmospheric profiles (Wan and Li, 1997). But the assumption may not be true because of the spatial misregistration or variations in land surface conditions (Gao et al., 2014; Wan and Li, 1997).

In fact, these existing LST retrieval methods were all designed to remove the effect of the atmosphere and LSE from satellite TIR observations under specific assumptions (Becker, 1987; Sobrino et al., 1991). Based on the above analysis, either atmospheric parameters, LSE, or both are required to be known in advance for these methods to recover accurate LST. Unfortunately, such prior knowledge at the pixel scale with high quality is not always available, which may lead to a decrease in LST retrieval accuracy. Inspired by the fact that SW and TES methods have complementary advantages and disadvantages, a hybrid method was proposed by Zheng et al. (2019) (hereafter referred to as the SW-TES method), and the required multi-channel configuration was also recommended (five channels centered at 8.6, 9.0, 10.4, 11.3, and 12.5 μm). According to their study, the SW method was first used to remove the atmospheric effect from satellite TIR observations to obtain the ground observed radiance. Subsequently, the TES method was used to separate the LST and LSE simultaneously (Gillespie et al., 1998). Thus, neither atmospheric information nor LSE is required in this method, which is believed to be a promising strategy for eliminating the dependence on prior knowledge during LST retrieval. However, this method can only be applied to observations with viewing zenith angles <20 ° from the nadir. Extensions are therefore needed because of the large viewing angles of commonly used satellite observations (such as ±55 ° for MODIS and AVHRR, leading to a local viewing zenith angle >65 °). Moreover, detailed performance assessment of the extended method at the global scale and validation using ground measured LST are also required to facilitate potential applications.

The objective of this study was to extend the SW-TES method for application to satellite observations with different viewing angles and assess its performance in detail. To achieve this aim, the extended SW-TES method was applied globally to obtain the LST of the entire Earth's land surface for four months, corresponding to different seasons. Next, using the MODIS LST products as reference, the performance of the extended method was investigated. In addition, the LST retrieval accuracy of the extended SW-TES method was also validated using the in-situ LST measured at seven American SURFRAD (SURFace RADiation budget network) sites. This paper is organized as follows: Section 2 introduces the data used to apply and evaluate the extended SW-TES method. Section 3 recalls the SW-TES method for LST retrieval and demonstrates the extensions of this study. Section 4 compares the retrieved LST with the MODIS LST product and in-situ LST measurements. Section 5 summarizes the main findings.

2. Data

2.1 AIRS data

Currently, satellite TIR observations using the above-mentioned five channels are still not available. The atmospheric infrared sounder (AIRS) onboard the Aqua satellite can acquire the ultra-fine hyperspectral radiance of the Earth's surface. The Aqua satellite operates in a polar sun-synchronous orbit at an altitude of 705 km, resulting in the AIRS passing overhead at almost the same local time every day with the Sun in the same position. The ascending node is approximately 1:35 pm, and the descending node is approximately 1:35 am. Therefore, these measurements are ideal for synthesizing theoretical multispectral observations of the Earth's surface during both daytime and nighttime. Therefore, AIRS data was introduced in this study to generate the five required multichannel TIR observations in theory. The detailed specifications of the AIRS instrument are listed in Table 1.

Table 1. Characteristics of the AIRS instrument.

Spatial Resolution	**13.5 km at nadir**
Spatial Sampling	1.1˚ footprints
Swath Width	1650 km (±49.5 degrees)
Spectral Range	2378 channels from 3.75–15.4 μm
Spectral Resolution	$\lambda / \Delta\lambda \approx 1200$ nominal
Spectral Sampling	$\lambda / \Delta\lambda \approx 2400$ nominal
Radiometric Accuracy	0.2 K at 265 K
Ground Coverage	>95% global daily

2.2 MODIS LST product for cross validation

Because it is difficult to obtain the ground-measured LST globally, the well-validated and commonly used MODIS LST product (MYD11_L2) was introduced as a reference for intercomparison (Wan et al., 2015). The MYD11_L2 product was generated at a spatial resolution of 1 km from the generalized split-window algorithm (GSW), which requires prior-known LSE information (Wan and Dozier, 1996). Similar to other SW methods, the performance of the GSW algorithm is sensitive to uncertainties in the emissivity difference between the two adjacent channels. Currently, the LSE of MODIS bands 31 and 32 are estimated by the classification-based method (Snyder et al., 1998). This method firstly classifies each pixel according to the input data in Land Cover (MCDLC1KM) and dynamic and seasonal factors. For different land cover types, the appropriate kernel models were applied to determine the bidirectional reflectance distribution function models (BRDF), and then to obtain the hemispherical-directional reflectance, $\rho(\theta)$. Finally, the emissivity could be computed by $1-\rho(\theta)$ (Snyder et al., 1998). However, the misclassification of the land cover type or inhomogeneity inside a single MODIS pixel still may decrease the LST retrieval accuracy.

According to the validation results using temperature-based (T-based) and radiance-based (R-based) methods over different land cover types, such as grassland, cropland, inland water, silt playa, and bare soil, the MYD11_L2 product is believed to have an uncertainty of approximately 1.0 K in most cases, except for bare soil sites (Coll et al., 2009; Duan et al., 2018; Duan et al., 2019a; Wan, 2008; Wan and Li, 2008; Wan, 2014). The largest biases in the MYD11_L2 LST retrieval results were found to occur in the hot and warm bare soil zones (HAWBSZ), with an underestimation up to 4.5 K in the collection 5 (C5) product. Fortunately, the accuracy of the retrieved LST over HAWBSZ pixels has been greatly improved in the collection 6 (C6) product after several refinements have been made and the mean error decreases to <2.0 K (Wan, 2014).

In addition to the MYD11_L2 product, a newer MODIS LST product, MYD21_L2, was also available currently (Hulley, 2015). The MODIS MYD21_L2 product was generated using the TES method which was originally designed for the Advanced Spaceborne Thermal Emission Reflection Radiometer (ASTER) instrument (Hulley and Hook, 2011). Comparing with the GSW method used in the MYD11_L2 product, the LST and LSE were retrieved simultaneously in the MYD21_L2 product. However, accurate atmospheric correction was firstly required before the LST retrieval (Hulley and Hook, 2011). The auxiliary atmospheric products (Hulley et al., 2016) and the Water Vapor Scaling (WVS) Method (Tonooka, 2005) therefore were used to remove the atmospheric effect. The uncertainty of the MYD21_L2 LST product was announced to be <1.5 K (Hulley et al., 2016; Hulley et al., 2012).

With the merits of good accuracy, long-term consistency, frequent revisit cycles, and global coverage, MODIS LST products have been widely used worldwide. Because the MODIS and AIRS are onboard the same satellite platform, Aqua, and provide observations of the same Earth surface patch within 5 minutes, it is easy to find pixel pairs matching each other both temporally and spatially. Thus, the MYD11_L2 C6 and MYD21_L2 C6 products for Aqua downloaded from the NASA website (https://search.earthdata.nasa.gov/) was used as the LST reference to evaluate the performance of the extended SW-TES method.

2.3 In-situ LST measurements for direct validation

To provide a direct evaluation of the performance of the extended SW-TES method, the T-based validation method was used in this study to compare the derived LST with the in-situ LST measured at the satellite overpass in a simple manner. The T-based method requires the ground LST to be measured accurately at the satellite passing time and is adequate for representing the LST at the satellite pixel scale (Li et al., 2013a). Therefore, collection of in-situ LST measurements is quite difficult, and the available datasets are very limited. The SURFRAD sites built in 1993 provide continuous, long-term, accurate, and quality-controlled measurements of broadband hemispherical upwelling and downwelling longwave radiation (3-50 μm) at seven different locations in the United States (Augustine et al., 2000). After applying the Stefan–Boltzmann law, in-situ LST could be obtained every 3 minutes (before 2009) or every 1 minute (after 2009) from the measured longwave radiation (Duan et al., 2019a). The SURFRAD measurements are currently widely used to evaluate the uncertainties of various satellite LST products, such as ASTER (Wang and Liang, 2009), AATSR (Ghent et al., 2017), MODIS (Duan et al., 2019a; Li et al., 2014; Wang and Liang, 2009), and VIIRS (Guillevic et al., 2014; Liu et al., 2015). In this study, the SURFRAD measurements were chosen to provide a direct validation of the extended SW-TES method described in Section 3. Relevant information on the seven SURFRAD sites is listed in Table 2.

Table 2. Detailed information of the seven SURFRAD sites.

Site name	**Latitude**	**Longitude**	**Land cover**
Penn State (PSU)	40.72033 °N	77.93100 °W	Cropland/natural vegetation
Fort Peck (FPK)	48.30798 °N	105.10177 °W	Grass
Bondville (BND)	40.05155 °N	88.37325 °W	Cropland
Table Mountain (TBL)	40.12557 °N	105.23775 °W	Grass
Desert Rock (DRA)	36.62320 °N	116.01962 °W	Open shrub
Sioux Falls (SXF)	43.73431 °N	96.62334 °W	Cropland
Goodwin Creek (GCM)	34.25470 °N	89.87290 °W	Cropland/natural vegetation

2.4 MODIS land cover product

To evaluate the performance of the extended SW-TES method in detail, the MODIS land cover product was used to study the consistency of the retrieved LST with the referenced LST for each land cover type separately. In this study, the annually published IGBP classification layer in the MODIS land cover climate modeling grid product (MCD12C1) was introduced.

According to the user guide (Sulla-Menashe and Friedl, 2018), the MCD12C1 product is provided with a pixel size of 0.05 ° × 0.05 °by spatially aggregating and re-projecting the tiled MODIS land cover type product (MCD12Q1). According to previous studies, the MCD12Q1 product was created using supervised classification of MODIS reflectance data, with an overall accuracy of 73.6% for the C6 product (Friedl et al., 2010; Sulla-Menashe et al., 2019). Following the IGBP scheme, the land surface was classified into 17 types (Fig. 1).

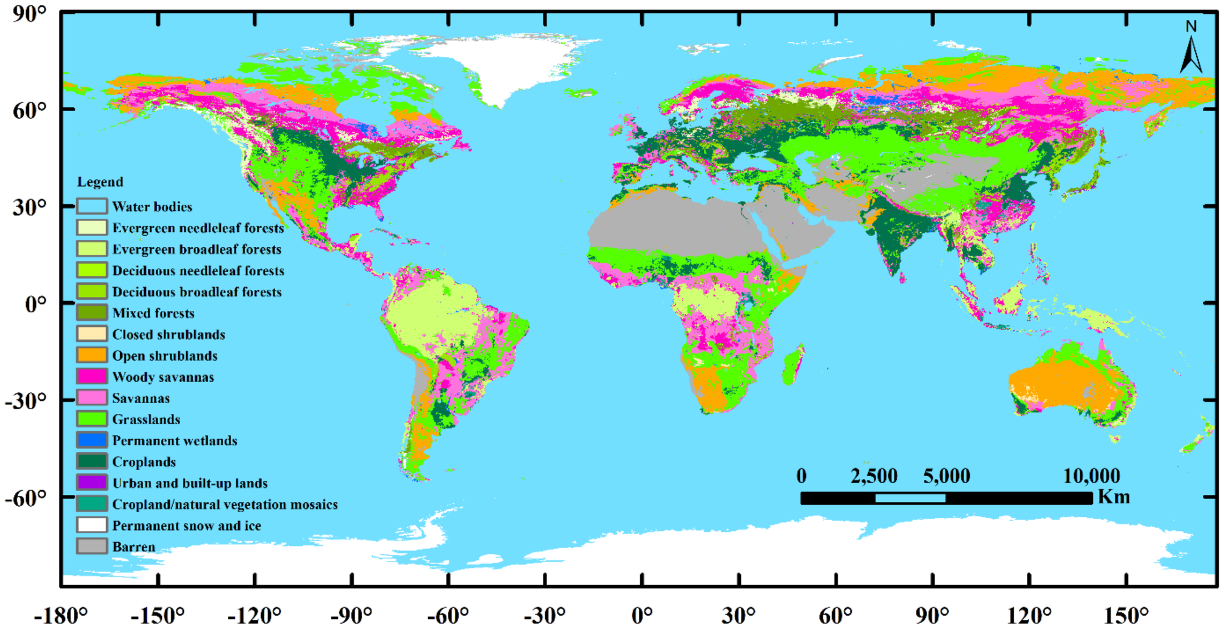


Fig. 1. Global land cover types according to the IGBP layer in MCD12C1 C6 product for the year 2018.

However, such a detailed classification is not necessary for evaluating the performance of the extended LST retrieval method. Therefore, similar land cover types were merged in this study (Table 3). Then, the consistency of the retrieved LST with the MODIS LST product was separately investigated for each of the land cover types in Section 4.2. Another advantage of combining similar land cover types is that it avoids the insufficiency of valid pixels regarding some of the land cover types of the Earth's surface, such as the deciduous needleleaf forest, urban and built-up lands, etc.

Table 3. Legends of the 17 original MODIS IGBP land cover types and the merged land cover types used in this study.

Original MODIS IGBP land cover type	**Merged land cover type**
Evergreen needleleaf forests	Forests
Evergreen broadleaf forests	
Deciduous needleleaf forests	
Deciduous broadleaf forests	
Mixed forests	
Closed shrublands	Shrublands
Open shrublands	
Woody savannas	Savannas
Savannas	
Grasslands	Grasslands
Permanent wetlands	Wetlands
Permanent snow and ice	Snow/Ice
Water bodies	Water
Croplands	Croplands
Cropland/natural vegetation mosaics	
Urban and built-up lands	Urban
Barren	Barren

2.5 Global climate classification map

To investigate the potential impact of climate type on the performance of the extended SW-

TES method, a Köppen-Geiger climate classification map was introduced (Beck et al., 2018). According to previous studies (Kriticos et al., 2012; Peel et al., 2007), the Köppen-Geiger system classifies the global climate into five main climate types (i.e., tropical, arid, temperate, cold, and polar) and 30 sub-types (Table 4). The map is generated according to the air temperature and precipitation data of the world using a threshold-based method (Beck et al., 2018). The Köppen-Geiger climate classification map used in this study was derived at a spatial resolution of 1-km by combining climatic air temperature and precipitation data from multiple independent sources, as shown in Fig. 2 (Beck et al., 2018, 2020). The spatial resolution was unprecedented compared to previous versions (Beck et al., 2018). Moreover, the topographic effects were also corrected explicitly in this version, which could provide more accurate results in mountainous areas (Beck et al., 2018).

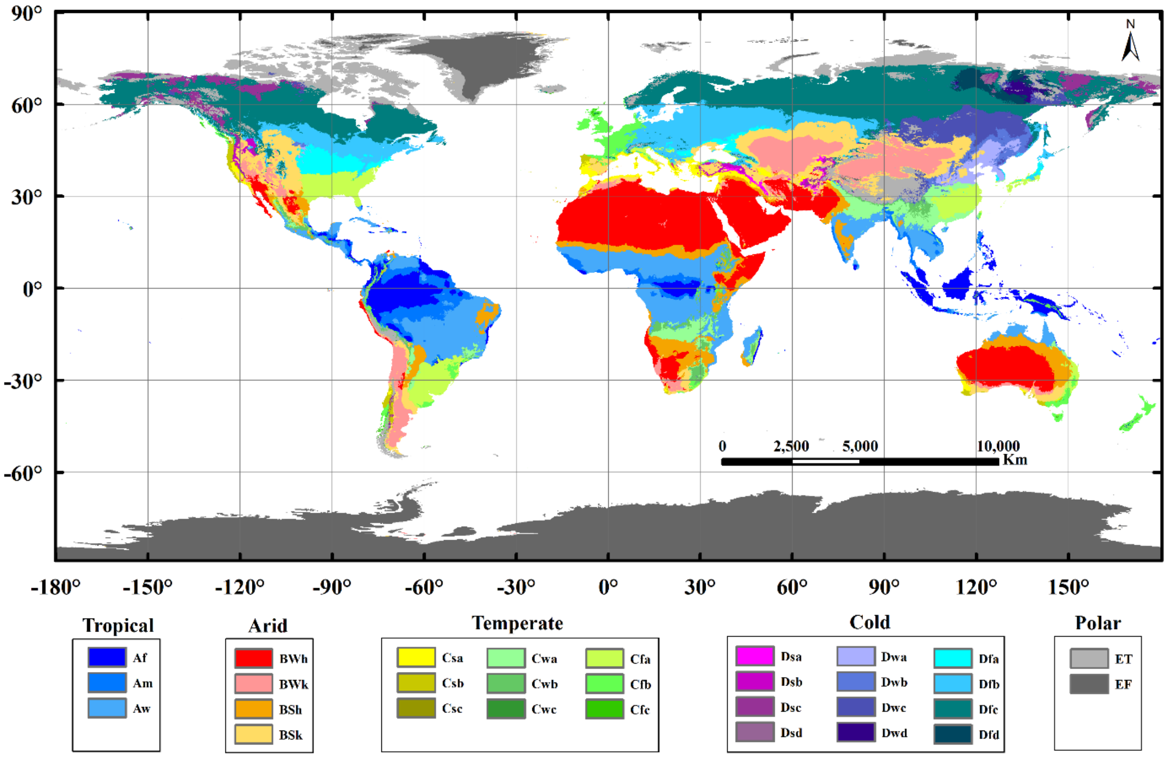


Fig. 2. Global Köppen-Geiger climate classification map (referring to Table 4 for abbreviations of each sub-classes).

Table 4. Overview of the Global Köppen-Geiger climate classes.

Main classes	**First sub-classes**	**Second sub-classes**	**Abbreviation**
Tropical	-	Rainforest	Af
	-	Monsoon	Am
	-	Savannah	Aw
Arid	Desert	Hot	BWh
		Cold	BWk
	Steppe	Hot	BSh
		Cold	BSk
Temperate	Dry summer	Hot summer	Csa
		Warm summer	Csb
		Cold summer	Csc
	Dry winter	Hot summer	Cwa
		Warm summer	Cwb
		Cold summer	Cwc
	No dry season	Hot summer	Cfa
		Warm summer	Cfb
		Cold summer	Cfc
Cold	Dry summer	Hot summer	Dsa
		Warm summer	Dsb
		Cold summer	Dsc
		Very cold winter	Dsd
	Dry winter	Hot summer	Dwa
		Warm summer	Dwb
		Cold summer	Dwc
		Very cold winter	Dwd
	No dry season	Hot summer	Dfa
		Warm summer	Dfb
		Cold summer	Dfc
		Very cold winter	Dfd
Polar	-	Tundra	ET
	-	Frost	EF

3. Methodology

3.1 Extension of the SW-TES method to different local viewing zenith angles

In the study by Zheng et al. (2019), a procedure was first designed to select specific TIR channel pairs that are eligible to recover accurate ground brightness temperature using only satellite measurements based on an SW-like equation. Then, the TES algorithm was introduced to separate the LST and LSE from the recovered ground brightness temperatures. Finally, the LST could be retrieved from satellite TIR measurements directly without knowing both atmospheric information and LSE on the basis of the elaborately selected channel configuration. In their study, five channels centered at 8.6, 9.0, 10.4, 11.3, and 12.5 μm with the full width at half maximum (FWHM) of 0.1 μm were suggested to obtain the ground brightness temperature of three channels centered at 8.6, 9.0, and 10.4 μm (Zheng et al., 2019). However, the coefficients of the SW-like equation provided in their study were only applicable to quasi-nadir observations. According to Table 1, the largest viewing angle of AIRS could be up to 49.5° from the nadir, resulting in a local viewing zenith angle (θ_v) near 60°. The viewing angle of other instruments, such as MODIS, was even larger. Extensions, therefore, are needed before applying to observations on a global scale.

In this study, the SW-like equation was optimized following previous studies to suppress the angular effects during ground brightness temperature retrieval (Wan and Dozier, 1996). Specifically, the coefficients in Eq. (1) were obtained as a function of the local viewing zenith angle, rather than constants (Niclos et al., 2011; Pérez-Planells et al., 2021). Similar to the traditional SW method, these coefficients could be determined by regression analysis of the simulation data (Wan and Dozier, 1996). First, the simulated satellite TIR measurements for seven local viewing zenith angles (1 / cos(θ_v) = 1.00, 1.15, 1.30, 1.45, 1.60, 1.75, and 1.90, with θ_v = 0.00°, 29.59°, 39.72°, 46.40°, 51.32°, 55.15°, and 58.24°, respectively) were generated. Then, the ground brightness temperature at different local viewing zenith angles could be acquired feasibly using this angle-dependent SW-like equation [Eq. (1)].

$$T_{gi} = A_0(\theta_v) + A_1(\theta_v)T_i + A_2(\theta_v)\left(T_i - T_j\right) + A_3(\theta_v)\left(T_i - T_j\right)^2 \quad (1)$$

where T_{gi} is the ground brightness temperature of channel *i*. $A_0(\theta_v)$, $A_1(\theta_v)$, $A_2(\theta_v)$, and $A_3(\theta_v)$ are the regression coefficients that vary with the local viewing zenith angle θ_v. Here, *i* and *j* represent the channels. T_i and T_j are the brightness temperatures of channels *i* and *j*, respectively.

Finally, the TES method could be applied to separate the LST and LSE based on the retrieved ground brightness temperatures.

3.2 Simulation dataset

To generate the simulation dataset for determining the coefficients of the SW-like equation [Eq. (1)], the five suggested channel response functions in the study by Zheng et al. (2019) were introduced. According to their algorithm sensitivity analysis, the performance of the SW-TES method decreased as the channel width (full width at half maximum, FWHM) and channel noise (noise equivalent differential temperature, NEΔT) increased (Zheng et al., 2019). Based on numerical simulations, the channel FWHM and NEΔT of possible future sensors should be respectively limited within 0.1 μm and 0.1 K if the LST retrieval accuracy of 1.0 K is required. For those sensors with channel FWHM <0.6 μm, the LST could be retrieved with RMSE <1.2 K and <1.5 K provided that the channel NEΔT is <0.1 K and <0.2 K, respectively (Zheng et al., 2019). In this study, considering that the spectral resolution of AIRS is ultra-fine, the ideal channel FWHM of 0.1 μm was therefore used in the following calculations (Fig. 3). Together with different ground-measured LSE spectra, atmospheric profiles, LST inputs, and sensor viewing angles, the satellite-observed channel brightness temperature under various imaging conditions could be simulated using the MODTRAN (MODerate resolution atmospheric TRANsmission) computer code, version 5.2.2 (Berk et al., 2004; Berk et al., 2006).

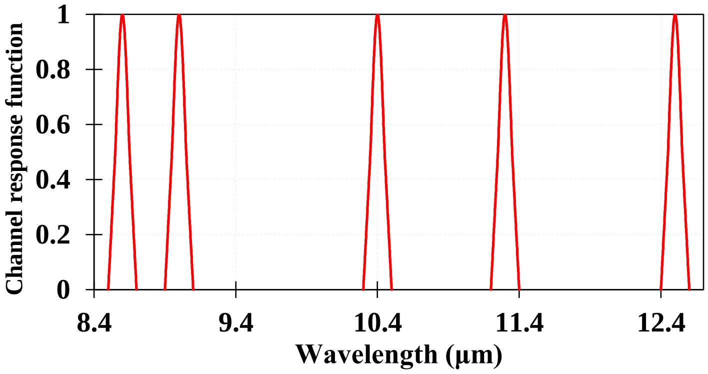


Fig. 3. Channel response functions of the five required channels.

In this study, the TIGR (The Institute for Genomic Research, http://ara.abct.lmd.polytechnique.fr/) database (Chevallier et al., 1998) was introduced as the MODTRAN inputs to generate simulation dataset of various atmospheric conditions (Chen et al., 2017; Ren et al., 2020; Wang et al., 2019b). After quality check and cloudy-sky removal, 946 atmospheric profiles remained which were considered to be redundant because many of them were similar to each other (Fig. 4). Therefore, it is necessary to select different atmospheric profiles to make the total atmospheric water vapor content (WVC) and the bottom layer temperature (T_0) uniformly distributed from dry to moist and from cold to warm, respectively. Following these criteria, a subset with a moderate sample number of 98 was selected and believed to be representative enough for determining the coefficients of the SW-like equation. As shown in Fig. 4, the selected 98 atmospheric profiles have varying characteristics with the water vapor content ranging from 0.09 g/cm^2 to 6.15 g/cm^2 and the bottom layer temperature (T_0) ranging from 236.25 K to 311.95 K.

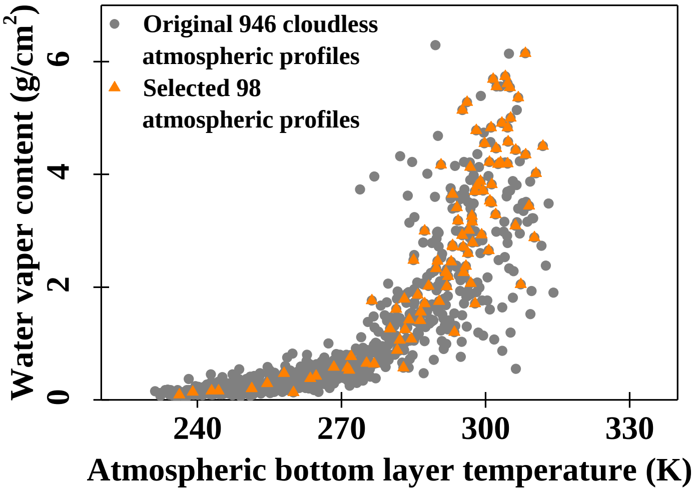


Fig. 4. Scatter plot of bottom layer temperature and total atmospheric water vapor content of the original 946 atmospheric profiles (represented by circles) and the selected 98 ones (represented by triangles).

The input LST was then assigned to vary with an interval of 5 K from $T_0 - 5$ K to $T_0 + 15$ K if $T_0 \geq 280$ K and from $T_0 - 5$ K to $T_0 + 5$ K if $T_0 < 280$ K. Additionally, the emissivity dataset used in the study of Zheng et al. (2019) was refined to be more representative at the global scale. First, the soil samples with similar characteristics were removed. Second, the vegetation samples were expanded to cover typical dry and fresh vegetations. Third, the typical rock samples were included to cover the three rock types (i.e. igneous, sedimentary, and metamorphic). Forth, the manmade samples were introduced to include typical paving, constructing, and roofing materials. In addition, the number of each kind of emissivity sample was determined referring to the study of Ren et al. (2020). Finally, a new dataset of 83 LSE spectra, containing 30 soil, 15 vegetation, 8 water/snow/ice, 25 rock, and 5 manmade samples were selected from the ECOSTRESS library (https://speclib.jpl.nasa.gov/) to represent different

Earth surface coverages globally (Fig. 5).

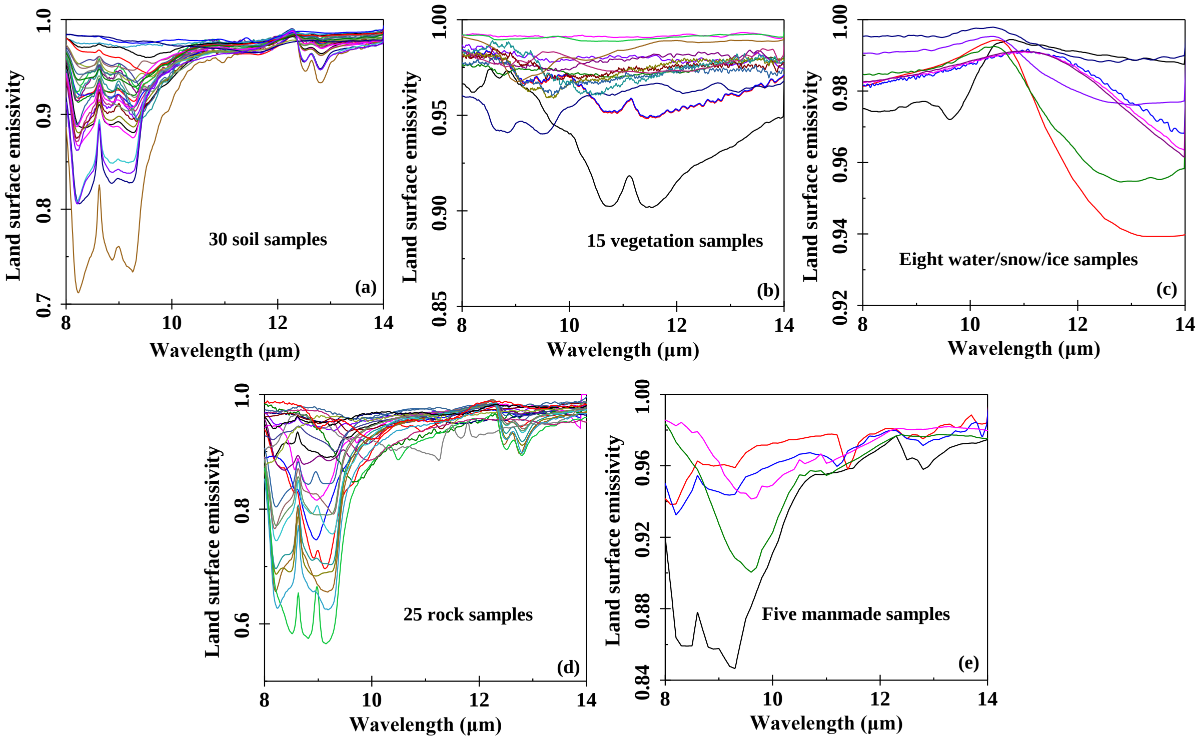


Fig. 5. LSE spectra of the (a) 30 soil samples, (b) 15 vegetation samples, (c) 8 water/snow/ice samples, (d) 25 rock samples, and (e) 5 man-made samples.

Finally, under each of the seven local viewing zenith angles (0.00°, 29.59°, 39.72°, 46.40°, 51.32°, 55.15°, 58.24°), a dataset containing 37,848 cases could be simulated using the parameters described above.

3.3 LST retrieval with different local viewing zenith angle

Based on the simulation dataset described in Section 3.2, the coefficients of Eq. (1) for each local viewing zenith angle could be obtained using the least-squares fitting method separately. As shown in Fig. 6, all the coefficients varied almost linearly as functions of the secant local viewing zenith angle. With increasing local viewing zenith angle, $A_0(\theta_v)$, $A_2(\theta_v)$, and $A_3(\theta_v)$ also increased, while $A_1(\theta_v)$ decreased, implying that it is necessary to extend the SW-TES method for its application to global TIR measurements with different viewing angles.

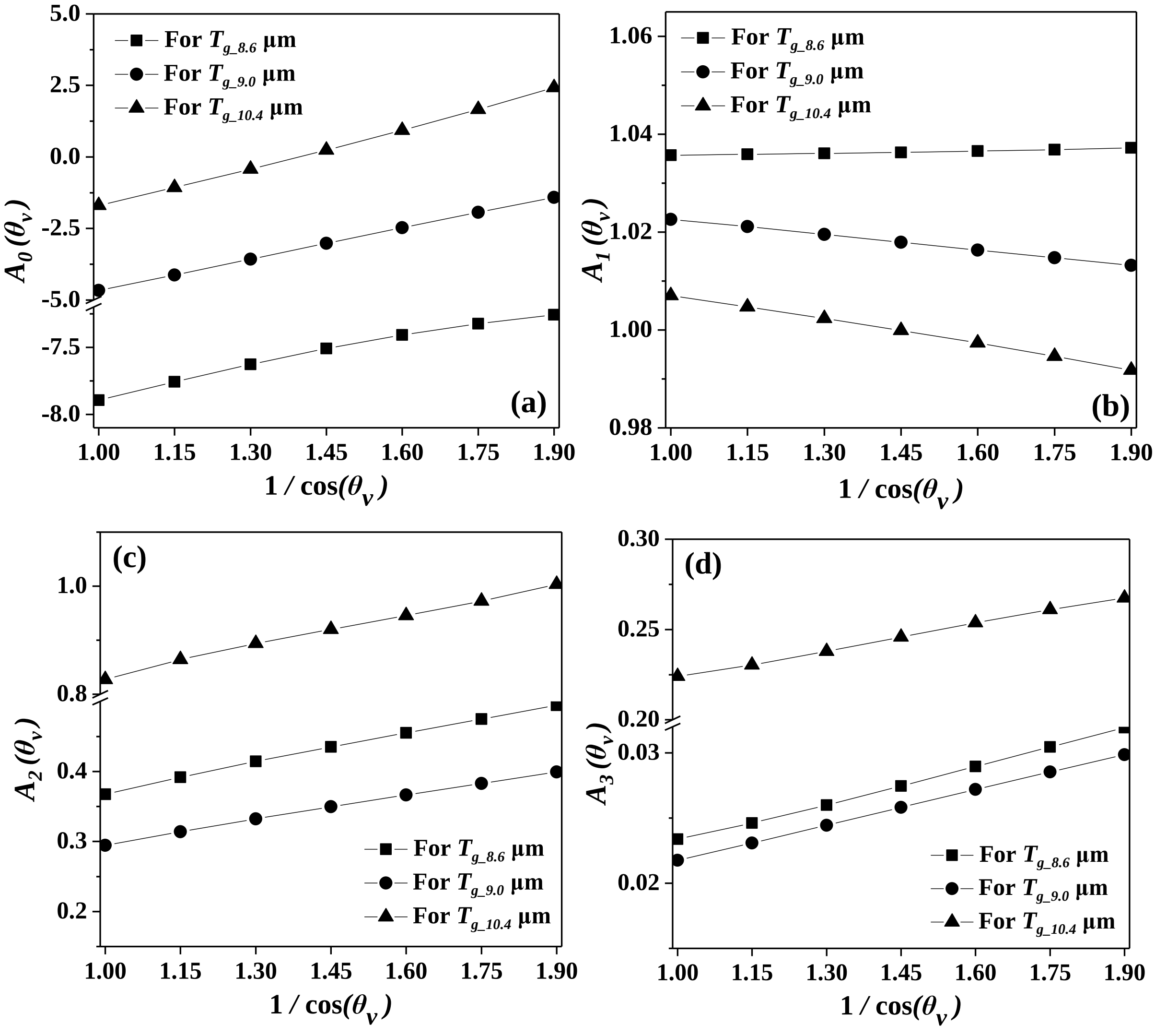


Fig. 6. Coefficients of Eq. (1) as functions of the secant local viewing zenith angle regarding to (a) $A_0(\theta_v)$, (b) $A_1(\theta_v)$, (c) $A_2(\theta_v)$, and (d) $A_3(\theta_v)$.

Using the coefficients obtained above, the ground brightness temperature of the three channels centered at 8.6, 9.0, and 10.4 μm could be obtained based on Eq. (1). Then, the TES method was introduced to separate the LST and LSE. This method combines three modules (Gillespie et al., 1998): the NEM algorithm, used to provide a first guess of the temperature and the emissivity spectrum; the Ratio algorithm, used to preserve shape of the emissivity spectrum; and the MMD (maximum-minimum difference) algorithm, used to find the accurate temperature and emissivity by relating the minimum emissivity ($\varepsilon_{\min}$) to the spectra contrast (*MMD*). Therefore, an empirical relationship between $\varepsilon_{\min}$ and *MMD* is very important and required in the TES method (Hulley and Hook, 2011). On the basis of previously selected 83 emissivity spectra, this $\varepsilon_{\min}$ ~ *MMD* relationship was adjusted for the spectra response functions of the three relevant channels with an uncertainty of about 0.008 (Fig. 7).

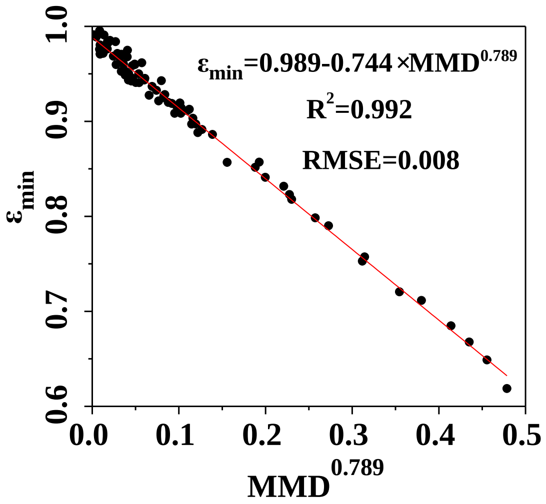


Fig. 7. The empirical relationship between $\varepsilon_{\min}$ and *MMD*, based on the selected 83 emissivity spectra samples.

To provide an overall look of the extended SW-TES method, a brief illustration is presented in Fig. 8.

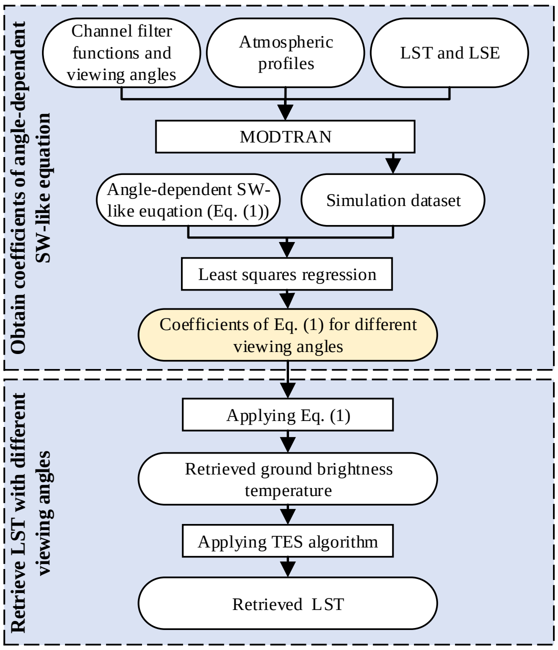


Fig. 8. Illustration of the extended SW-TES method for global land surface temperature retrieval.

Using the simulation dataset above, the uncertainty of the extended SW-TES method could be accessed. Firstly, the simulated satellite observations under each of the seven local viewing zenith angles were separately inputted into the extended SW-TES method. The LST retrievals then were compared with the MODTRAN LST inputs used for generating the simulated dataset under each viewing zenith angle. As shown in Fig. 9a, the root-mean-square error (RMSE) of retrieved LST increased by about 0.4 K from 0.94 K to 1.37 K, as the secant local viewing zenith angle increased from 1.0 to 1.9. When the simulations and the relevant LST retrievals regarding the seven local viewing zenith angles were considered together, the overall LST retrieval uncertainty was approximately 1.16 K, as shown in Fig. 9b. Although the increase of local viewing zenith angle affects the performance of the extended SW-TES method, the overall LST retrieval accuracy was still acceptable according to the simulation results.

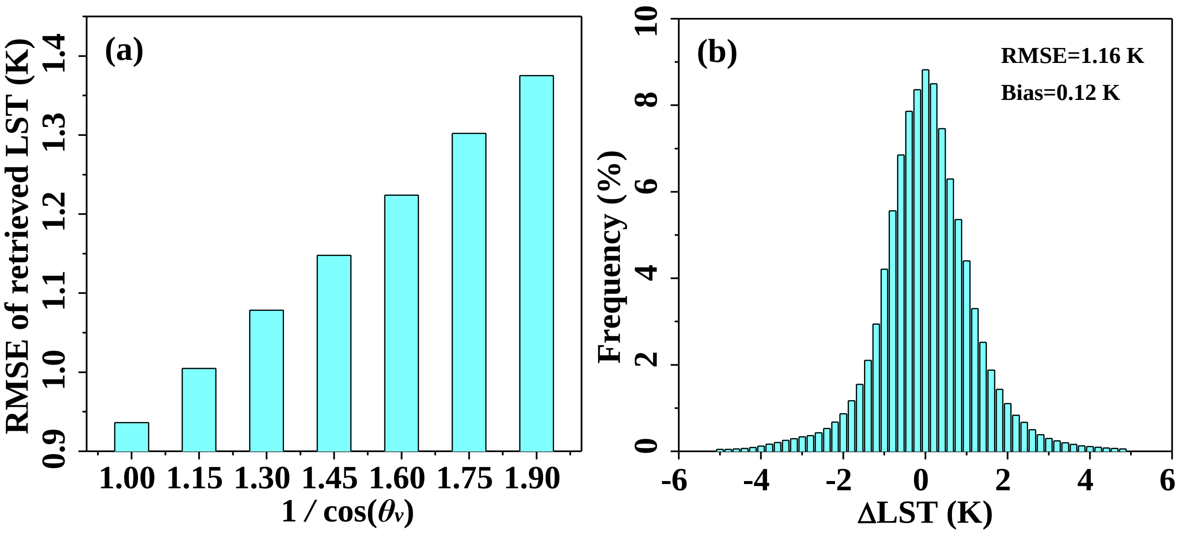


Fig. 9. Land surface temperature (LST) retrieval accuracy on the basis of simulation dataset. (a) RMSE of retrieved LST as a function of the secant local viewing zenith angle. (b) Residual histograms of retrieved LST when considering the simulations of the seven viewing zenith angles all together.

4. Application to real satellite data

4.1 LST retrieval from real AIRS data

Currently, there are no operational satellites providing multispectral TIR observations of the five suggested channels. In this study, to apply the extended SW-TES method to real satellite data, theoretical space measurements were first generated by convoluting the five artificial channel response functions (Fig. 3) with the hyperspectral radiance of the AIRS instrument onboard the Aqua platform. Observations of the entire Earth's land surface, obtained in four months (July 2018, October 2018, January 2019, and April 2019), representing the four seasons, were collected and processed.

To apply the angle-dependent SW-like equation [Eq. (1)] to real satellite data, a look-up table of the coefficients for different viewing angles is required. Because the coefficients in Eq. (1) varied smoothly with the viewing zenith angle, according to Fig. 6, it was sufficient to establish this look-up table using the coefficients at the seven viewing zenith angles described in Section 3.3. The coefficients at any other viewing angle could thus be linearly interpolated according to this look-up table. Then, the ground brightness temperature of each pixel could be easily obtained. Consequently, LST could be retrieved without knowing any prior knowledge after applying the TES algorithm. Finally, the LST retrieval accuracy was assessed through cross-validation with the MODIS LST product and direct validation with the in-situ LST measurements. The detailed processing flow for applying the extended SW-TES method to real AIRS data is shown in Fig. 10.

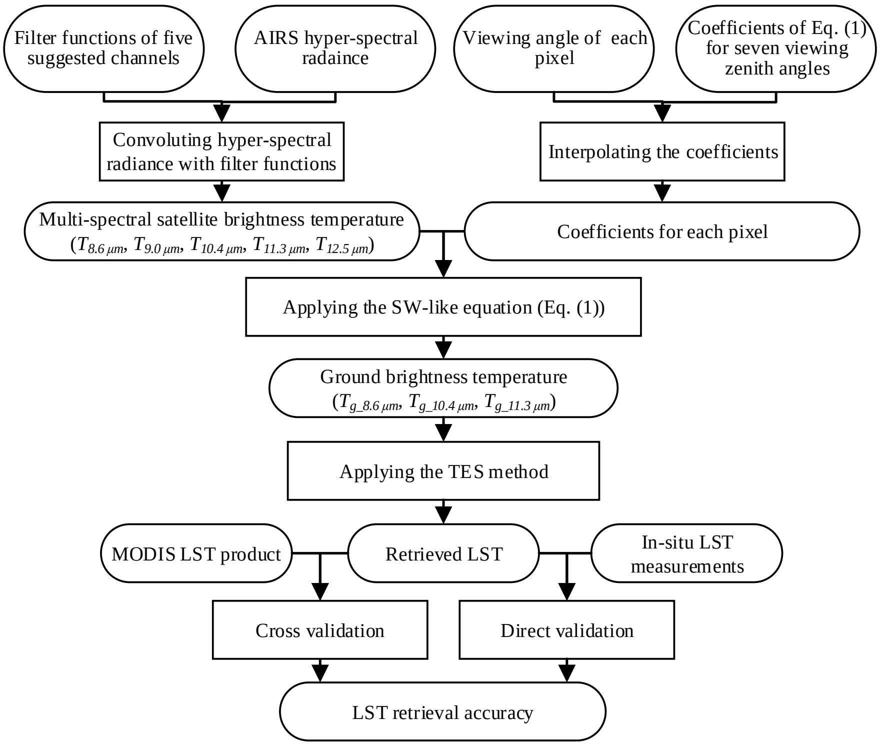


Fig. 10. The flow diagram of applying the extended SW-TES method to real AIRS data.

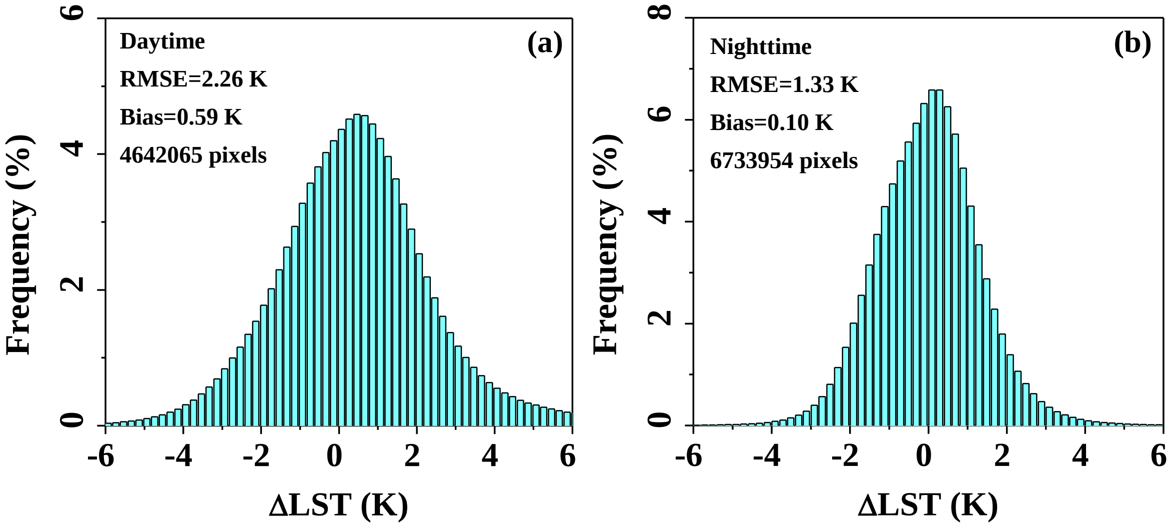
4.2 Intercomparison with MYD11 LST product

4.2.1 Intercomparison residuals for daytime and nighttime observations

Following the processing flow shown in Fig. 10, the extended SW-TES method was implemented to retrieve LST globally from the real AIRS data. Then, the well-validated MODIS MYD11_L2 LST product was considered as a reference to evaluate the global LST retrieval results. Based on the quality control layers distributed along with the MYD11_L2 LST product, only cloudless pixels with announced LST retrieval accuracy which were better than 1.5 K were retained. Moreover, considering that the pixel size of MODIS is around 1.0 km, while that of AIRS is 13.5 km, the valid MODIS LST pixels were first aggregated to match the AIRS spatial resolution using the area-weighted pixel aggregation method (Gao et al., 2012; Qian et al., 2013). For example, all the MODIS LST pixels inside or overlapped with the corresponding AIRS pixel footprint were weighted and aggregated according to their overlapping areas. Please note that the pixel size varies along with the view angles for both instruments. Therefore, the actual coordinates of the four vertexes of each AIRS and MODIS pixel were calculated to determine their accurate overlapping areas. In addition, for each AIRS pixel, the corresponding MODIS LST product was aggregated only if all the involved MODIS pixels had valid LST values and the standard deviation was <1.5 K. We applied these two criteria to ensure the representativeness of the aggregated MODIS LST on the AIRS pixel scale and to guarantee the quality and homogeneity of the selected pixel pairs for intercomparison.

The discrepancies between the retrieved LST and aggregated MODIS MYD11_L2 LST at the global scale for the four months were then calculated and analyzed (Fig. 11). Considering that the thermal environment should be more homogeneous at nighttime than during the daytime owing to the absence of solar loadings, results were separately generated for daytime and nighttime observations. It was observed that the LST retrieval results were overestimated during the daytime with the RMSE of 2.26 K and bias of 0.59 K when using the MODIS MYD11_L2

LST product as a reference (Fig. 11a). For nighttime observations, the LST retrieval results were much more consistent with the MODIS LST product, with an RMSE of 1.33 K (Fig. 11b), which was close to the algorithm uncertainty of the extended SW-TES method, as shown in Fig. 9b. Besides, the LST retrieval uncertainty (including RMSE and Bias) were also calculated as functions of MODIS LST (Fig. 11). As the LST increased, results showed that the RMSE of retrieved LST gradually increased during the daytime (Fig. 11c) while decreased during the nighttime (Fig. 11d). But the bias of retrieved LST decreased with increasing LST during both daytime and nighttime. Since the thermal homogeneity of land surface may decrease with increasing the LST during the daytime because of solar heating process. Therefore, the thermal heterogeneity of the land surface was suspected to be the main factor affecting the intercomparison results.



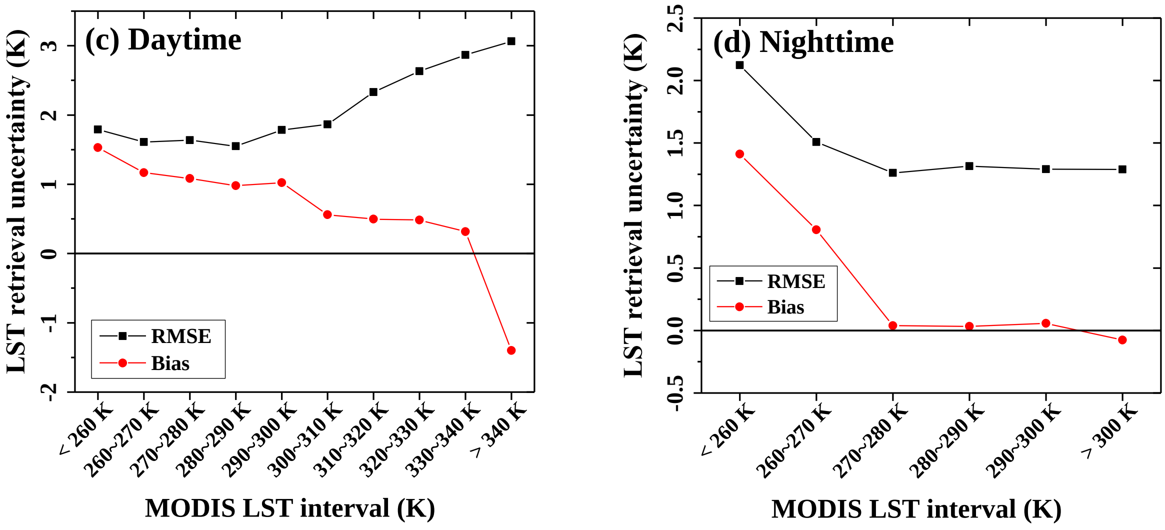


Fig. 11. Residual histograms of retrieved land surface temperature (LST) intercomparing with MODIS MYD11_L2 LST product (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime; and LST retrieval uncertainty as functions of MODIS LST for global observations acquired during (c) the daytime and (d) the nighttime.

A global distribution map of the discrepancy between the retrieved LST and aggregated MODIS LST products was also produced for investigation. The RMSE and bias for each pixel were calculated if more than two valid LST retrieval results were available in the period of the selected four months. Moreover, similar to the processes above, the statistical analysis for the daytime and nighttime observations was done separately.

According to the global RMSE distribution map for daytime observations (Fig. 12a), a large difference between the retrieved LST and MODIS LST products was found in some places, such as the regions in the east of the Caspian Sea, south of the Arabian Peninsula, north and southeast of Australia, and central Africa. However, during the nighttime, almost all pixels had relatively smaller LST retrieval RMSE and the magnitude generally did not vary with geolocation (Fig. 12b). The spatial continuity of the global RMSE distribution map for nighttime observations was also better than that of daytime observations. According to the MODIS land cover product for 2018 (Fig. 1), the land surface of the above-mentioned regions was much more likely dominated by sparse shrublands and grasslands. Therefore, the severe surface thermal heterogeneity was considered as the most probable reason for the large discrepancies in these regions during the daytime.

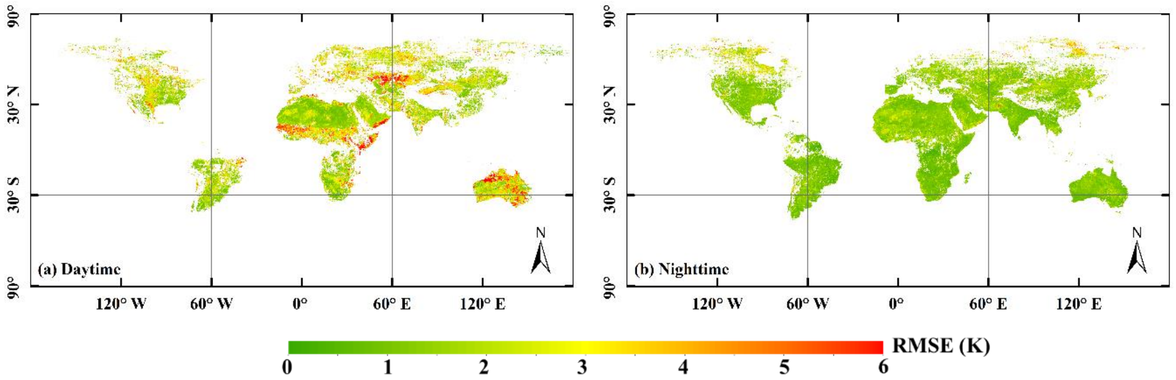


Fig. 12. RMSE distribution map of the retrieved land surface temperature (LST) intercomparing with the aggregated MYD11_L2 LST product for global observations acquired during (a) the daytime and (b) the nighttime.

Fig. 13 shows that the distribution pattern of the global bias coincided with the RMSE distribution maps. The results showed that the retrieved LST was slightly underestimated compared with the aggregated MODIS LST product in desert areas (such as the regions of Sahara, Taklimakan, and Arabian) and overestimated in the rest areas for both daytime and nighttime observations. To investigate the possible reason that caused the LST underestimation in these areas, the retrieved LSE was compared with those used in the MODIS MYD11_L2 LST product. However, the variation of the surface emissivity spectra around 8.6 μm is quite large and the LSE for MODIS channel 29 is not provided by the MYD11_L2 product. In addition, none of the five channels used in this study overlap with MODIS channel 32 when evaluated using their channel response functions. Therefore, we went on to compare the retrieved LSE for the 10.4 μm channel and the LSE from MODIS channel 31. We corrected for the potential impact of differences in the channel response functions by fitting the retrieved LSE for 10.4 μm channel to that of MODIS channel 31 using linear regression. The adaption coefficients were obtained using the emissivity dataset used in this study. Finally, we used this data to produce the bias distribution maps for the retrieved and adapted LSE values for the 10.4 μm channel using the aggregated MYD11_L2 LSE for MODIS band 31 as a reference (Fig. 14). Results showed that the retrieved LSE in some regions of the Sahara, Arabian, and the Taklimakan was a bit larger than the corresponding MODIS LSE product during both daytime and nighttime. But in other regions, it was generally smaller than the MODIS LSE. This may be one of the reasons that the retrieved LST was slightly underestimated in some desert areas comparing with the MODIS MYD11_L2 LST product while overestimated in the rest areas. Moreover, it should be noticed that the LST retrieval accuracy regarding to bare soil pixels in the C6 MYD11_L2 product was generally lower than that of other land cover types, the absolute uncertainty of which was validated to be around ±2.0 K according to previous studies (Duan et al., 2017; Duan et al., 2018; Wan, 2014). This may also influence the comparison results.

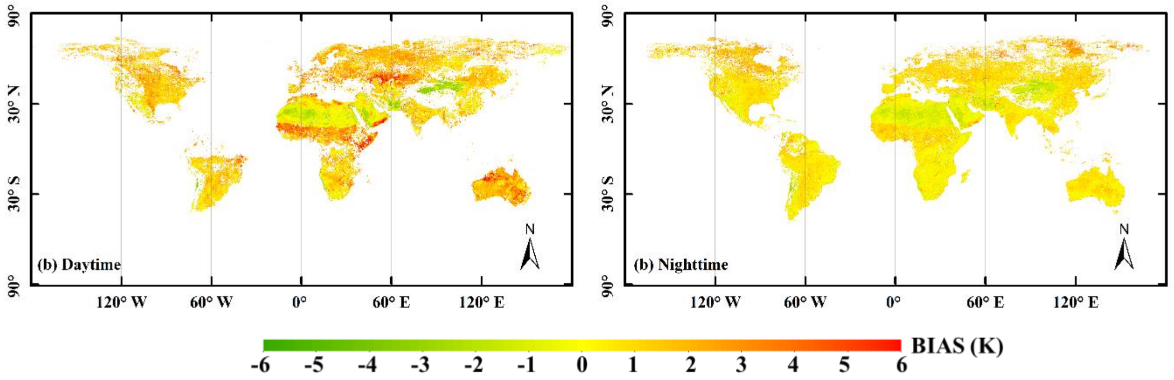


Fig. 13. Bias distribution map of the retrieved land surface temperature (LST) intercomparing with the aggregated MYD11_L2 LST product (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime.

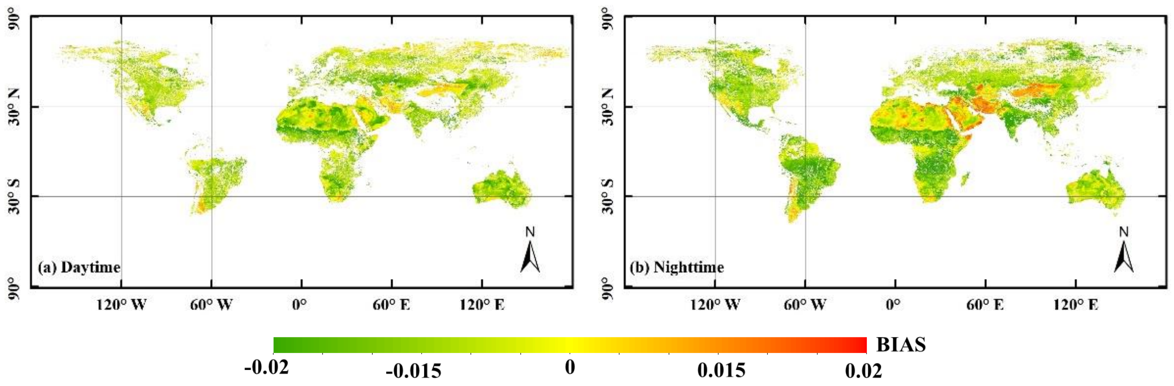


Fig. 14. Bias distribution map of the retrieved land surface emissivity (LSE) for the 10.4 μm channel intercomparing with the aggregated MYD11_L2 LSE for MODIS band 31 (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime. Please note that the retrieved LSE for the 10.4 μm channel has been adapted to the response function of MODIS channel 31 by linear regression.

These results indicated that the discrepancy between the retrieved LST and MODIS MYD11_L2 LST product was much smaller for nighttime observations than for daytime observations. During the nighttime, because of the absence of solar loading, the LST is considered to be more spatially homogeneous and closer to the effective air temperature. Thus, less uncertainty was introduced in both the aggregated MODIS MYD11_L2 LST product and the retrieved LST. Therefore, the retrieved LST coincided better with the MODIS MYD11_L2 LST product during the nighttime than during the daytime.

4.2.2 Intercomparison residuals for different seasons

To evaluate the influence of seasons on the LST retrieval results, the discrepancy between the retrieved LST and MODIS LST products was calculated daily. Because massive cloud coverage usually exists globally, the number of available pixel pairs is limited to a single day. Thus, the results obtained in the same month were combined to produce the statistics of the LST retrieval residuals. Moreover, one should notice that the four seasons are opposite in the two hemispheres of the Earth. Thus, the results for the four months (July 2018, October 2018, January 2019, and April 2019) were rearranged accordingly. In addition, the statistics were made for daytime and nighttime observations separately (Fig. 15).

During the daytime, the results showed that the RMSE (Fig. 15a) and bias (Fig. 15b) between the retrieved LST and MODIS LST products increased as the season changed from spring to summer and then decreased in autumn and winter. The magnitude of RMSE was <3.0 K, while the bias was <1.0 K based on the results during the four seasons. In autumn and winter, the RMSE and bias decreased further to <2.0 K and <0.5 K, respectively. In addition, the positive biases indicated that the retrieved LST was overestimated by taking the MODIS LST product as a reference for daytime observations in the four seasons. For nighttime observations, the discrepancy between the retrieved LST and MODIS LST product was much smaller than the daytime results. The RMSEs for all nighttime observations were <1.5 K, and no obvious variation in pattern was found as the seasons changed (Fig. 15a). Moreover, the magnitudes of LST retrieval biases during the nighttime were all <0.2 K, which were close to zero. Both RMSE and bias implied that the LST retrieval results were much closer to the MODIS MYD11_L2 LST product in the nighttime than during the daytime. However, it should be noted that negative biases occurred in the results during the nighttime in autumn, as shown in Fig. 15b.

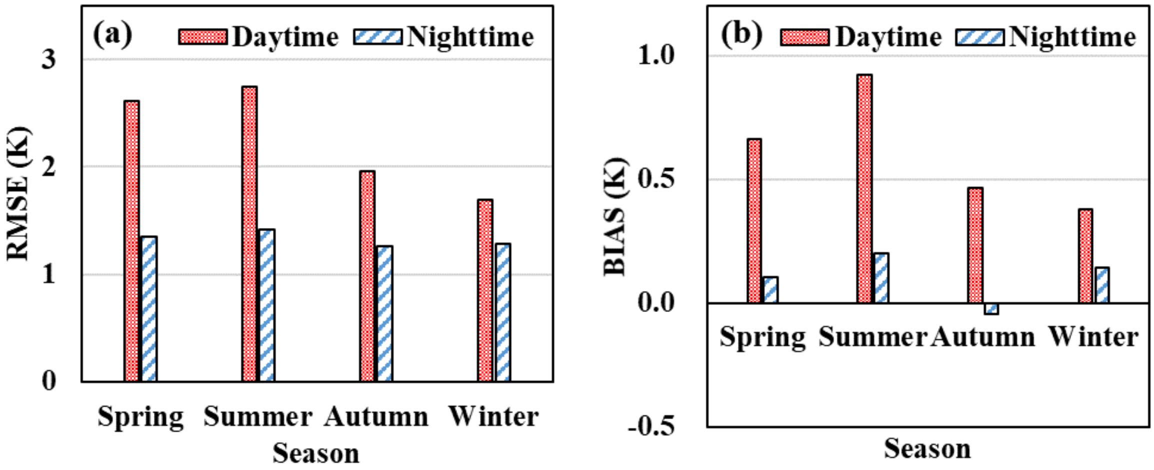


Fig. 15. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST) intercomparing with the MYD11_L2 LST product (the former minus the latter) for different seasons at the global scale. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern hemisphere, while refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere. Therefore, the results for these four months were rearranged accordingly for statistical analysis.

In addition, variations in the RMSE and bias with latitude were investigated for different months (Fig. 16). Statistics were made for a latitude interval of five degrees. During the daytime, with a decrease in latitude from 80 °N to 20 °N, the RMSE and bias for the observations in July 2018 first increased and then decreased. When the latitude approached the equator, the RMSE and bias increased again, but gradually decreased as the latitude continually decreased to the South Pole. As for the observations in January 2019, the varying pattern of RMSE and bias was reversed as the latitude increased from the South Pole to the North Pole. The results of the spring and autumn were generally between summer and winter. However, during the nighttime, the LST retrieval RMSE of the four months increased gradually with increasing latitude from the South Pole to the North Pole, and no other obvious variations were found. But there was a decrease in the biases with respect the latitude between 20 °N and 40 °N for both daytime and nighttime observations. This is believed to be relevant to desert regions, as shown in Fig. 13.

Overall, the retrieved LST during the nighttime coincided better with the MODIS LST product than the daytime. However, the discrepancy between the retrieved LST and MODIS LST product during summer was generally much larger than that in other seasons. In addition, most of the biases were found to be positive, indicating that the retrieved LST was overestimated compared with the MODIS LST product under most circumstances. According to the sensitivity analysis provided by Zheng et al. (2019), LST retrieval bias increases with an increase in the noise-equivalent differential temperature (NEΔT). Because of the quadratic term introduced in the SW-like equation, the noise always contributes positively to the LST retrievals. This could probably explain the overestimation of LST retrievals by the proposed method.

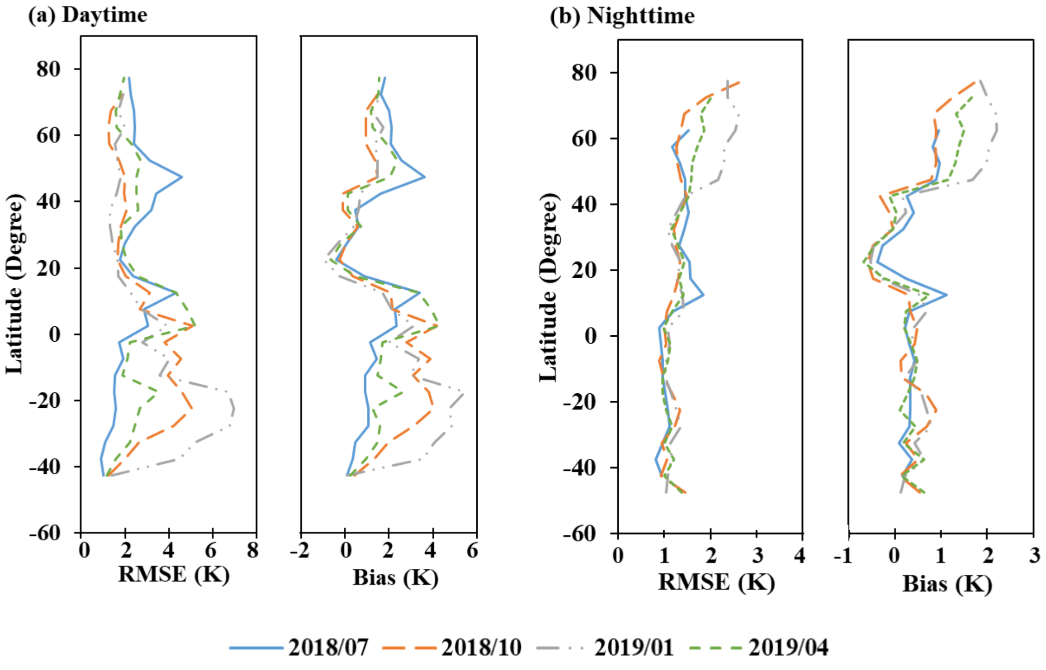


Fig. 16. RMSE and bias of the retrieved land surface temperature (LST) intercomparing with the MYD11_L2 LST product as functions of latitude (the former minus the latter) during (a) the daytime, (b) nighttime.

4.2.3 Intercomparison residuals for different land cover types

To further study the performance of the proposed LST retrieval method, statistics of the discrepancy between the retrieved LST and MODIS MYD11_L2 LST products were processed separately for different land cover types. The MCD12C1 product published in 2018 was introduced to identify the land cover type for each pixel. Because the pixel size of the retrieved LST was larger than that of the MCD12C1 product, the land cover type possessed by the majority of MODIS pixels inside the corresponding AIRS pixel's footprint was taken as the classification result of the retrieved LST pixel. Moreover, if the MODIS pixel number of the dominated land cover type inside the corresponding AIRS pixel's footprint is <60%, this pixel would be removed from the following statistics.

During the daytime, it was shown that the RMSE of LST retrievals was <2.5 K for different land cover types, except for shrublands and grasslands, with a magnitude of approximately 3.0 K (Fig. 17a). Compared with the daytime results, the RMSE was <1.4 K for all land cover types during the nighttime, except for the water and wetlands (Fig. 17a). Nevertheless, the LST retrieval RMSE of nighttime observations over the water and wetlands were about 1.6 K and 2.2 K respectively, which were still lower than most daytime results. Moreover, as indicated by the statistics of bias, the retrieved LST was underestimated compared with the MODIS MYD11_L2 LST product for barren and overestimated for the remaining land cover types (Fig. 17b).

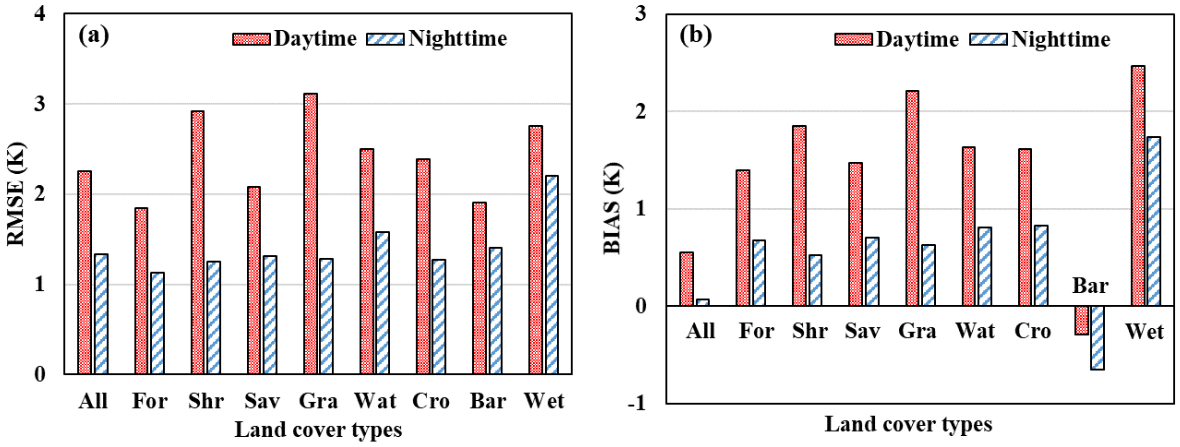
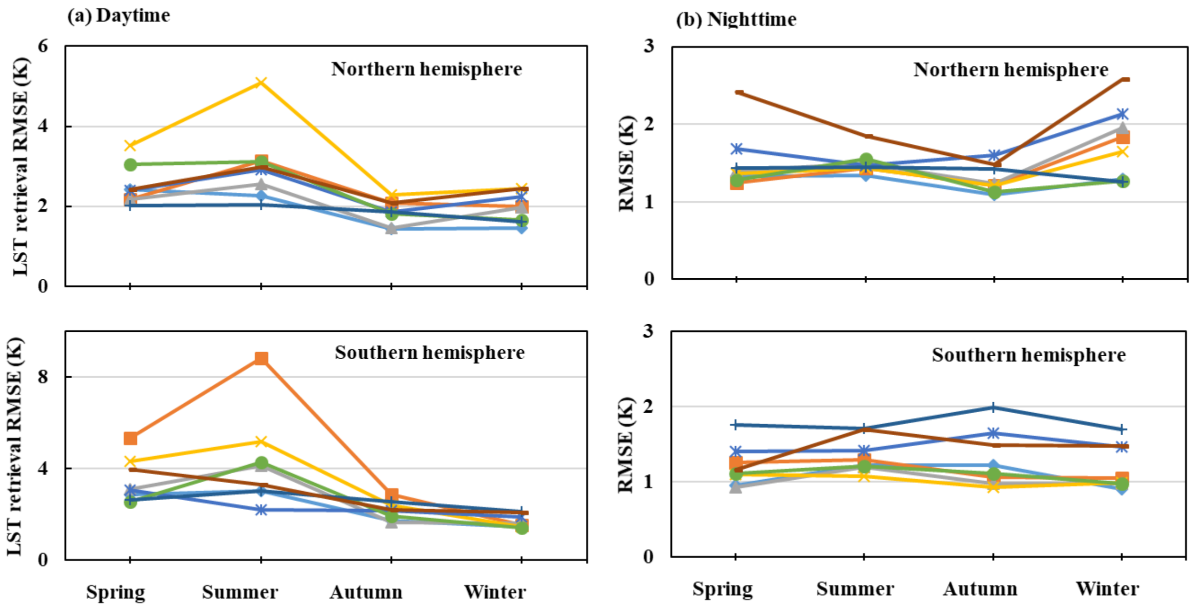


Fig. 17. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD11_L2 LST product (the former minus the latter) for different land cover types. Observations acquired in the four months were merged to generate the statistics. The meaning of abbreviations follows: All=All land cover types together, For=Forests, Shr=Shrublands, Sav=Savannas, Gra=Grasslands, Wat=Water, Cro=Croplands, Bar=Barren, Wet=Wetlands.

Variations in the LST retrieval RMSE for each land cover type as a function of season are also shown in Fig. 18. During the daytime, the results showed that the RMSE of the retrieved LST generally increased first and then decreased as the season changed from spring to winter for most land cover types (Fig. 18a). However, a large discrepancy was found during the summer for grasslands in the northern hemisphere (mainly in central Africa according to Fig. 13), with RMSE >4.0 K, and for shrublands in the southern hemisphere (mainly in Australia according to Fig. 13), with RMSE >8.0 K. This may be due to three reasons. First, radiation from vegetation and bare soil contribute to satellite TIR observations because the shrublands and grasslands are usually covered by sparse vegetation. Because the surface heating speed is quite different between vegetation and bare soil, the surface heterogeneity should therefore increase during the daytime in hot seasons. Consequently, the upscaling error in the aggregated

MODIS LST product may also increase. Second, the LSE was retrieved simultaneously with the LST in the extended SW-TES method, and it was assigned according to the land cover types. For heterogeneous land surfaces, the LSE may be generated differently using these two strategies. This may also introduce additional biases in the comparison results. As shown in Fig. 18, during the daytime of summer, the RMSE between the retrieved LSE for the 10.4 μm channel and the aggregated LSE for the MODIS band 31 for grasslands is the largest in the northern hemisphere. In the southern hemisphere, the retrieved LSE for the 10.4 μm channel for shrublands, grasslands, and wetlands are generally more different from the aggregated LSE for the MODIS band 31 than other land cover types. However, the LSE retrieval RMSE using the MODIS product as a reference is <0.02, which generally could not lead to an LST retrieval discrepancy of >4.0 K. Therefore, the third reason may be that the atmospheric correction using the SW-like equation (1) was failed under these conditions.

As for the nighttime observations (Fig. 18b), the retrieved LST in both hemispheres was close to the MODIS LST product for all land cover types in all seasons, with RMSE <2.0 K, except the results for water and wetlands in some seasons of the northern hemisphere. One reason for the slightly larger RMSE under this condition may be that pure pixels are rare for water and wetlands during the winter in the northern hemisphere. In addition, the thermal heterogeneity of the land surface may increase because of the freezing water and snow cover in winter. This is perhaps another reason for the relatively larger RMSE for these land cover types during winter in the northern hemisphere. Nevertheless, as shown in Fig. 18b, the magnitude of RMSE was less than 1.5 K in most cases, indicating that the performance of the extended SW-TES method was consistent with the MODIS MYD11_L2 LST product.



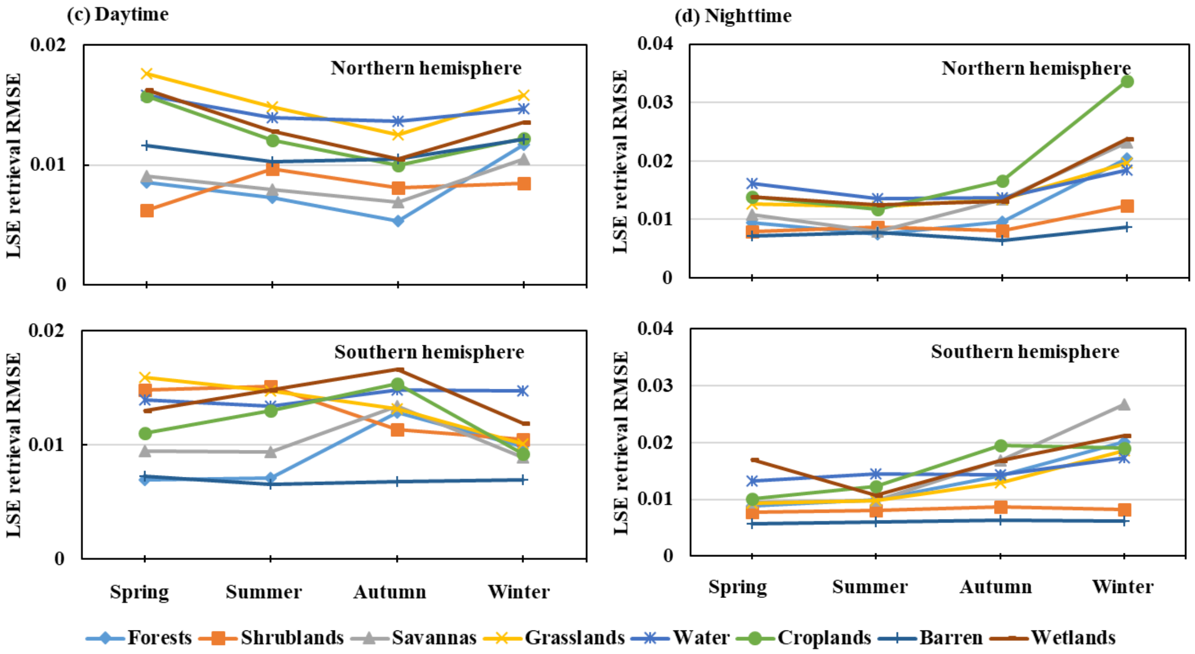


Fig. 18. RMSE of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product as a function of season for different land cover types separately for observations during (a) the daytime, and (b) the nighttime; and the RMSE between the retrieved land surface emissivity (LSE) for the 10.4 µm channel and the aggregated MYD11_L2 LSE for MODIS band 31 as a function of season for different land cover types separately for observations during (c) the daytime, and (d) the nighttime. Please note that the retrieved LSE for 10.4 µm channel has been adapted to the response function of MODIS channel 31 by linear regression. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January

2019, respectively, for the northern hemisphere, and refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere.

The biases between the retrieved LST and MODIS MYD11_L2 LST products were also calculated and are shown in Fig. 19. Except for the results of barren areas, the retrieved LST was overestimated compared with the MODIS MYD11_L2 LST product for both daytime and nighttime observations during all seasons in the two hemispheres. Statistics showed that the biases were generally larger during the daytime than during the nighttime. As shown in Fig. 19a, the biases were almost all >1.0 K for different land cover types, except for the barren during the daytime, while the magnitude was <1.0 K under most circumstances during the nighttime. The largest biases for daytime observations occurred in summer in the grasslands in the northern hemisphere (≈4.2 K) and shrublands in the southern hemisphere (≈7.8 K). The main reason for this observation may be the larger thermal heterogeneity of shrublands during daytime caused by solar loading, as explained above. For nighttime observations, the largest bias occurs in winter with respect to the wetlands in the northern hemisphere (≈2.2 K) and during autumn in the built-up lands/barren in the southern hemisphere (≈-1.2 K). However, the magnitude was still smaller than in most cases during daytime.

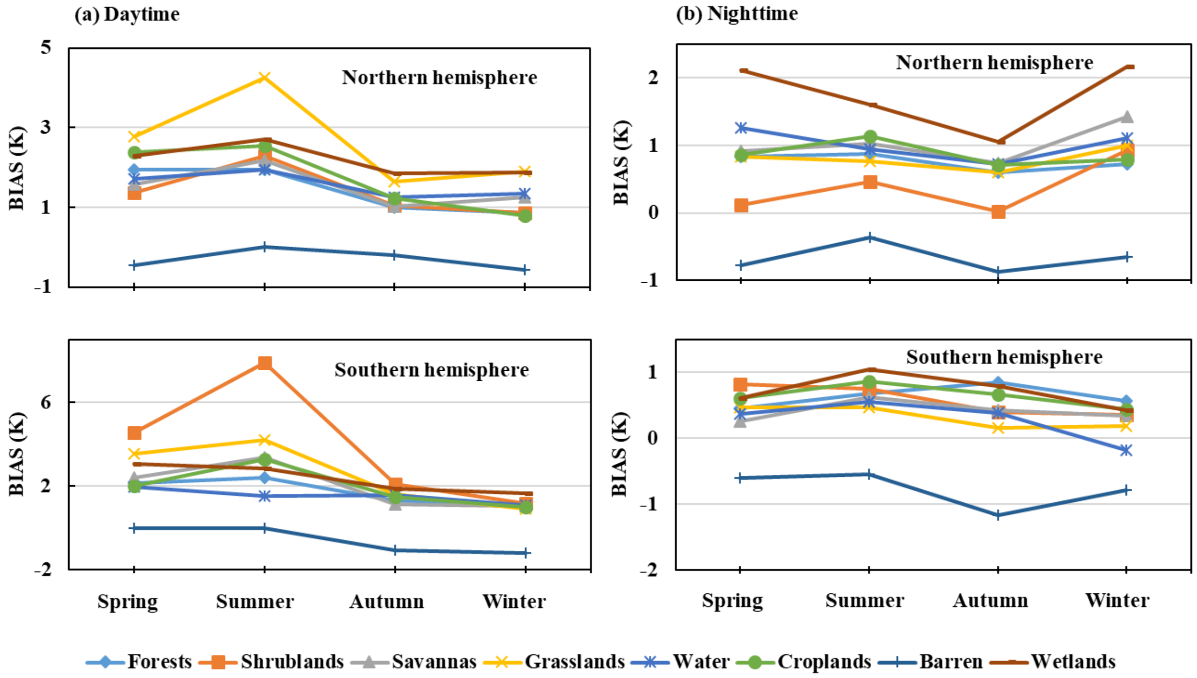


Fig. 19. Bias of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product (the former minus the latter) as a function of season for different land cover types separately. The statistics were processed for observations acquired in the (a) daytime, and (b) nighttime. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern hemisphere, and refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere.

Overall, these results suggested that the strong solar loading during the daytime could have a considerable influence on the comparison results. The relatively smaller RMSE and bias for nighttime observations than the daytime observations implied that thermal heterogeneity was likely the main reason for the difference between the retrieved LST and MODIS MYD11_L2 LST product.

4.2.4 Intercomparison residuals for different climate types

Climate type may also impact the performance of the extended SW-TES method. Statistics of the discrepancy between retrieved LST and MODIS MYD11_L2 LST products were processed separately for different climate types. In this section, the RMSE and bias of retrieved LST with MODIS MYD11_L2 LST product as a reference were calculated for each of the five main climate types classified by the Köppen-Geiger system (Fig. 2 and Table 4).

As shown in Fig. 20, as the climate transited from “tropical” to “polar,” the RMSE of retrieved LST gradually decreased from 2.4 K to 1.9 K for daytime observations, while the RMSE increased from 1.1 K to 1.9 K for nighttime observations (Fig. 20a). In addition, the biases for observations of the arid climate type were generally smaller than those of other climate types. The bias regarding the results of the arid climate type during the nighttime was even below zero (Fig. 20b). Considering that only the LST retrieval biases for “barren” were negative, as shown in Fig. 17, large overlaps must exist between the areas covered by the land cover type “barren” and the climate type “arid.”

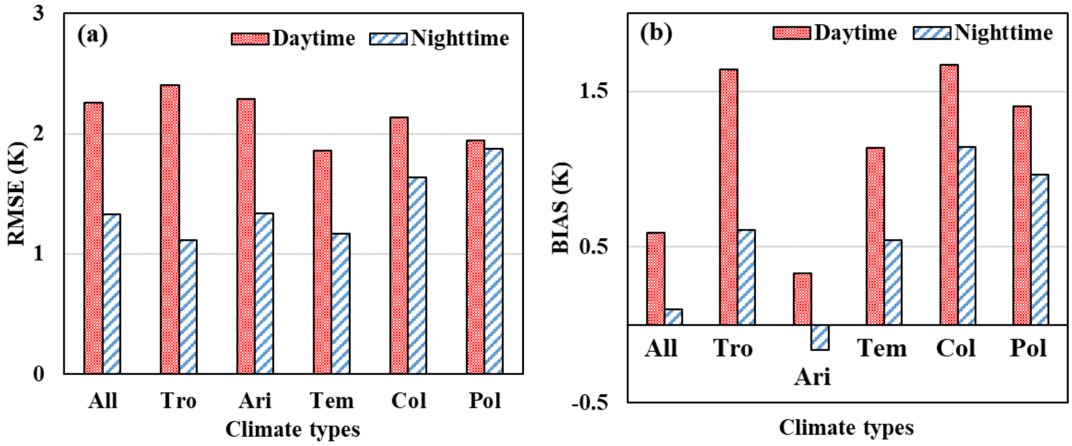


Fig. 20. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD11_L2 LST product (the former minus the latter) for different climate types. Observations acquired in the four months were merged to generate the statistics. The meaning of abbreviations follows: All=All climate types together, Tro=Tropical, Ari=Arid, Tem=Temperate, Col=Cold, Pol=Polar.

Similar to land cover types, variations in the LST retrieval RMSE for each climate type as a function of season were also investigated, as shown in Fig. 21. For daytime observations (Fig. 21a), it was observed that the LST retrieval RMSE varied much more considerably during spring and summer in both the hemispheres. During autumn and winter, the RMSE of the retrieved LST was almost the same referring to the MODIS MYD11_L2 LST product. This was perhaps caused by the severe thermal heterogeneity of the land surface during the daytime of hot seasons. The largest value was observed in spring with respect to the tropical climate in northern hemisphere (≈4.0 K) and the arid climate in southern hemisphere (≈6.8 K). For all nighttime observations (Fig. 21b), the RMSE of the retrieved LST were <2.0 K, except for cold and polar climates during the winter. During the night, the highest values were observed in winter with respect to the cold climate in northern hemisphere (≈2.4 K) and the polar climate in southern hemisphere (≈2.1 K). Obviously, the LST retrieval RMSE during nighttime coincided with the MODIS MYD11_L2 LST product better than that during the daytime.

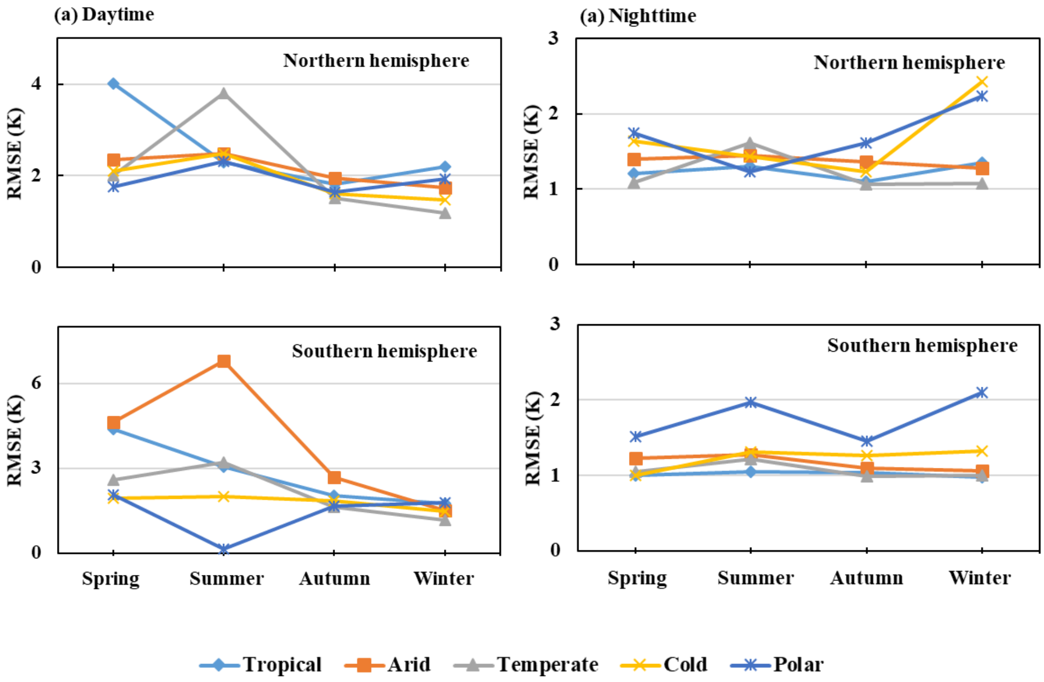


Fig. 21. RMSE of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product as a function of season for different climate types. The statistics were processed for observations acquired in the (a) daytime, and (b) nighttime. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern hemisphere, and refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere.

Fig. 22 showed that the variations in the LST retrieval bias for each climate type as a function of season. It was observed that the retrieved LST was overestimated compared with the MODIS MYD11_L2 LST product for almost all climate types during the daytime (Fig. 22a). However, during nighttime, the biases were negative for the arid climate in the northern hemisphere in all seasons and for the polar climate in the southern hemisphere, except winter (Fig. 22b). For most cases, the magnitude of bias was <2.0 K. But attention should be paid to the daytime results of tropical and temperate climates in the northern hemisphere, as well as the results of tropical and arid climates in the southern hemisphere, the magnitude of which could be larger than 3.0 K.

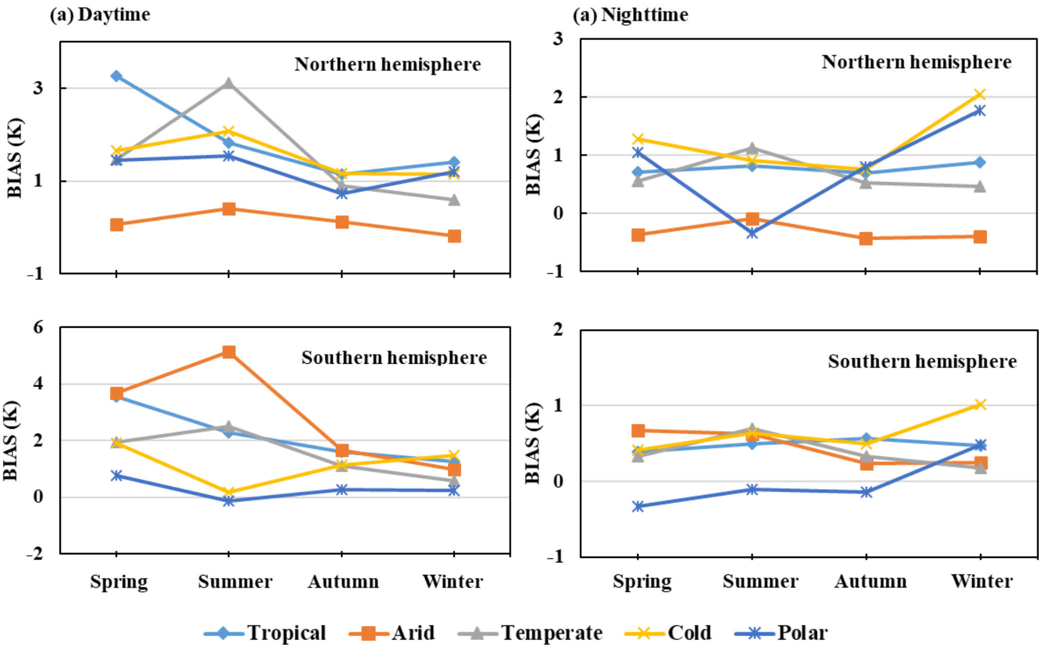


Fig. 22. Bias of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product (the former minus the latter) as a function of season for different climate types. The statistics were processed for observations acquired in the (a) daytime, and (b) nighttime. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern hemisphere, and refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere.

4.3 Intercomparison with MYD21 LST product

To have a better understanding about the performance of the extended SW-TES method, the retrieved results were also compared with the MODIS MYD21_L2 product (Fig. 23-Fig. 30). Our data showed that the LST retrieval RMSE was increased by approximately 0.4 K than that when using the MYD11_L2 LST as a reference for both daytime and nighttime observations at the global scale (Fig. 23). This bias indicated that the retrieved LST was underestimated by -0.97 K and -0.86 K for the daytime and nighttime observations, respectively (Fig. 23). In contrast, the retrieved LST was seen to be overestimated when compared with the MYD11_L2 LST products (Fig. 11). In fact, the MYD21_L2 LST product was shown to be slightly larger than the MYD11_L2 LST product for most cases according to previous studies (Hulley et al., 2016; Yao et al., 2020). This indicated that the LST retrieval results using the extended SW-TES method fell in between these two MODIS LST products for most situations when evaluated at the global scale. When considering the discrepancies between our LST retrievals and the MYD21_L2 LST product for different land cover types, we found that the desert areas had larger RMSE (Fig. 24) as the retrieved LST values for these regions was shown to be underestimated (Fig. 25). One reason for this may be that the LSE for these regions was slightly overestimated (Fig. 26).

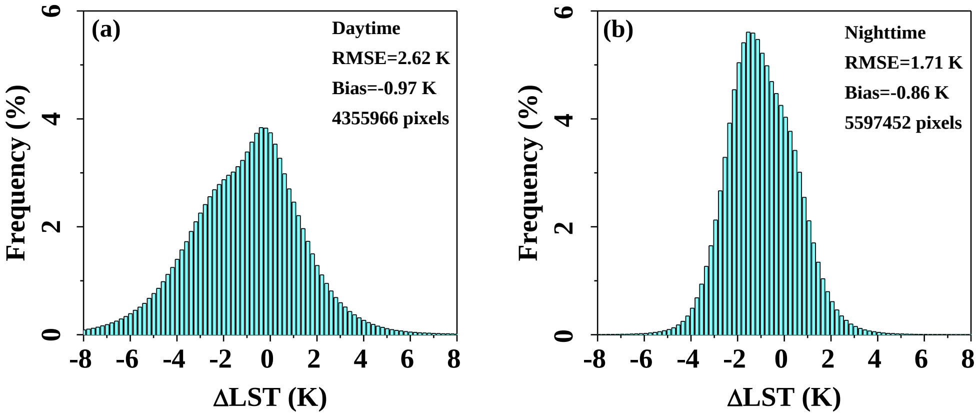


Fig. 23. Residual histograms of retrieved land surface temperature (LST) intercomparing with MYD21_L2 LST product (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime.

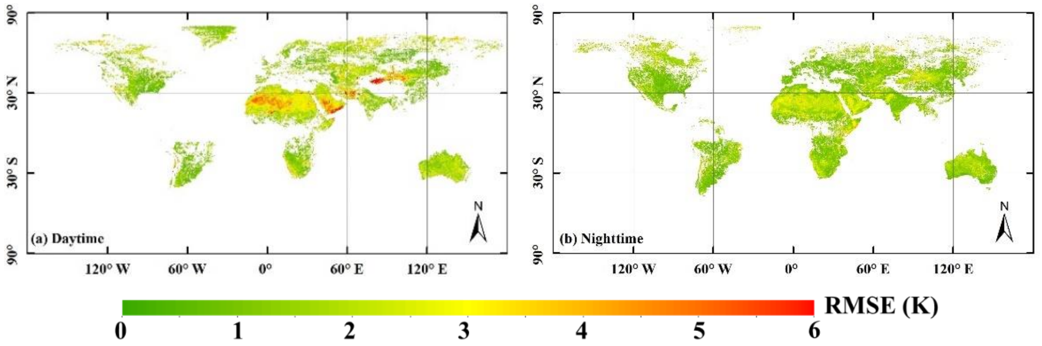


Fig. 24. RMSE distribution map of the retrieved land surface temperature (LST) intercomparing with the aggregated MYD21_L2 LST product for global observations acquired during (a) the daytime and (b) the nighttime.

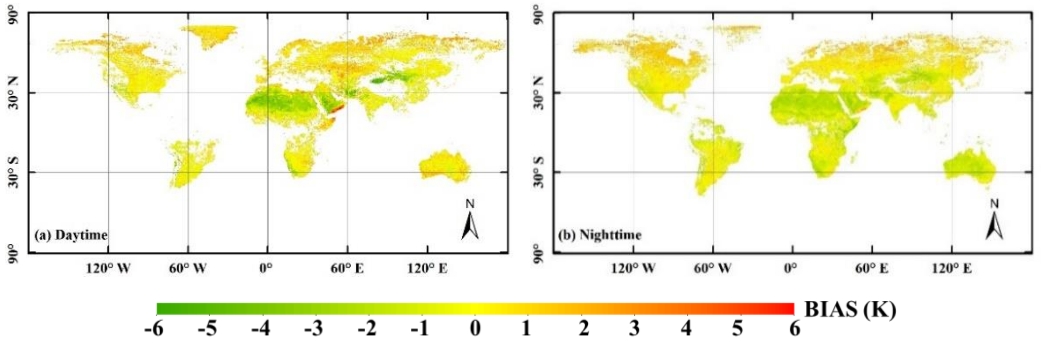


Fig. 25. Bias distribution map of the retrieved land surface temperature (LST) intercomparing with the aggregated MYD21_L2 LST product (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime.

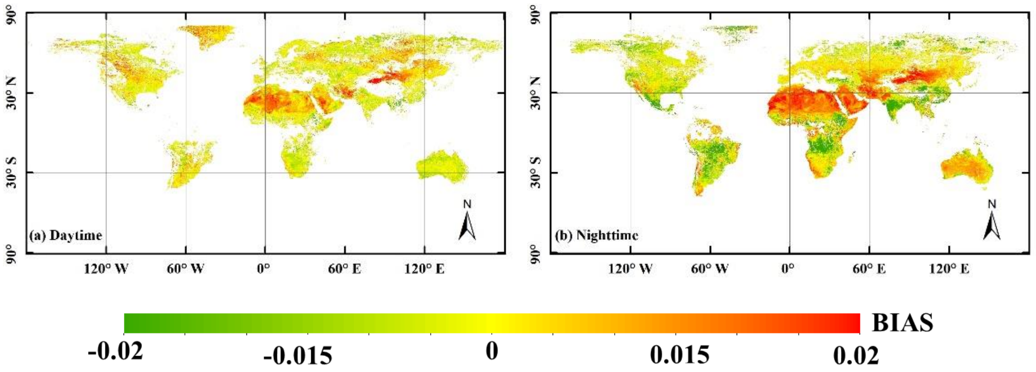


Fig. 26. Bias distribution map of the retrieved land surface emissivity (LSE) for the 10.4 μm channel intercomparing with the aggregated MYD21_L2 LSE for MODIS band 31 (the former minus the latter) for global observations acquired during (a) the daytime and (b) the nighttime. Please note that the retrieved LSE for the 10.4 μm channel has been adapted to the response function of MODIS channel 31 by linear regression.

The RMSE for the retrieved LST for each season was almost the same as those calculated using the MYD11_L2 LST product as a reference (Fig. 27a). However, all the biases were negative (Fig. 27b), suggesting that the LST retrievals were all smaller than the MYD21_L2 LST products for each of the different seasons. Fig. 28 showed that the largest discrepancy appears in the land cover type values for barren data with an RMSE of >3.0 K and a bias of <-2.0 K. The other land cover types had an RMSE of <2.0 K using the MYD21_L2 LST as a reference (Fig. 28) while the RMSE for the LSE retrievals in the 10.4 μm channel when evaluated against the aggregated MYD21_L2 LSE for MODIS band 31 was <0.02 (Fig. 29). Additionally, for those regions with an “Arid” climate type, the RMSE and bias values for the retrieved LST were >2.5 K (Fig. 30a) and <-1.0 K (Fig. 30b) respectively, which was significantly poorer than the results for the other climate types. In fact, barren land cover was generally shown to overlap with the arid climate type. Therefore, the results shown in Fig. 30 agreed with those of Fig. 28. Moreover, a comparison of the LST retrieval biases when using the MYD21_L2 LST as a reference with those results shown in section 4.2 once again showed that the retrieved LST values fell between these two MODIS LST products under most conditions.

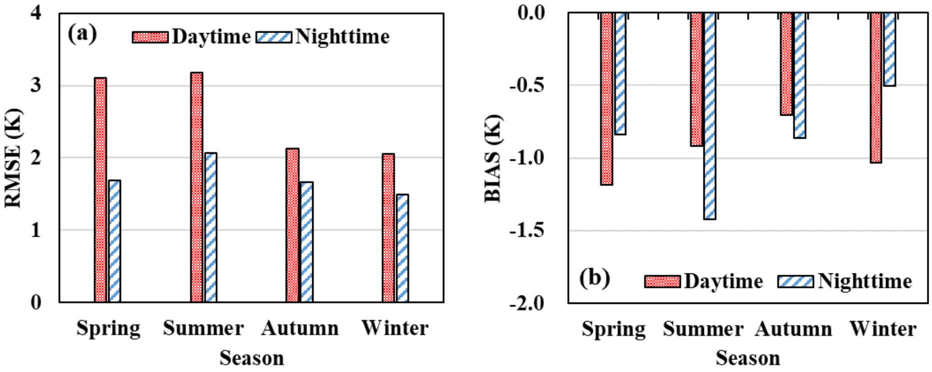


Fig. 27. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST) intercomparing with the MYD21_L2 LST product (the former minus the latter) for different seasons at the global scale. The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern hemisphere, while refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for the southern hemisphere. Therefore, the results for these four months were rearranged accordingly for statistical analysis.

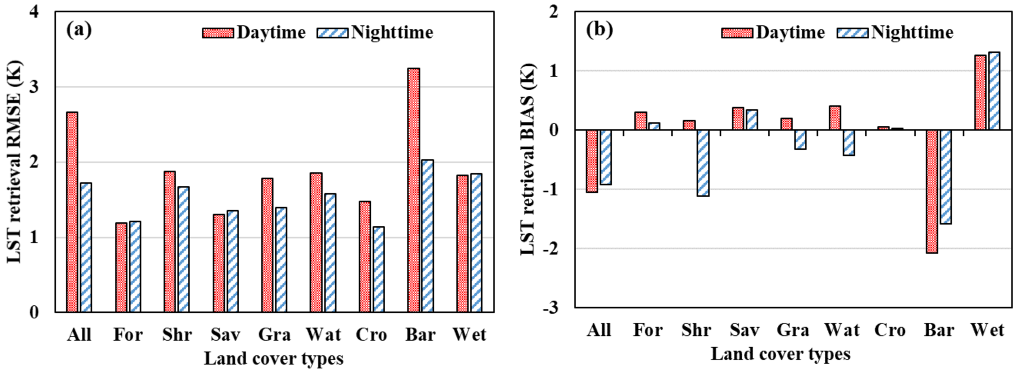


Fig. 28. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD21_L2 LST product (the former minus the latter) for different land cover types. Observations acquired in the four months were merged to generate the statistics. The meaning of abbreviations follows: All=All land cover types together, For=Forests, Shr=Shrublands, Sav=Savannas, Gra=Grasslands, Wat=Water, Cro=Croplands, Bar=Barren, Wet=Wetlands.

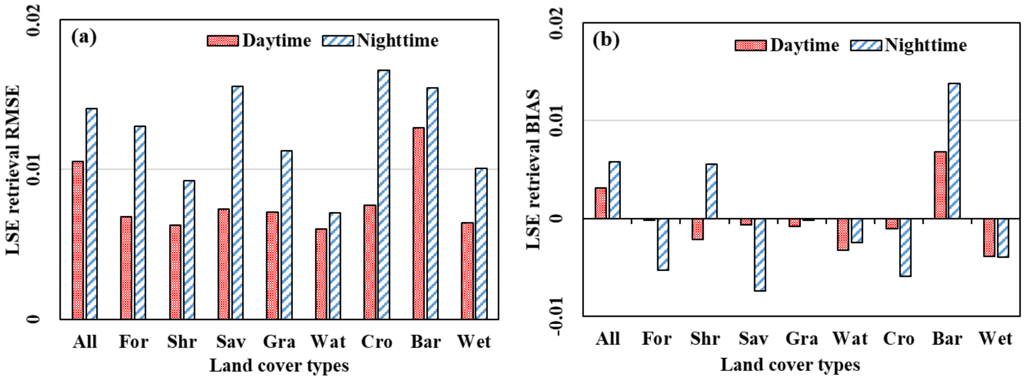


Fig. 29. (a) RMSE and (b) Bias between retrieved land surface emissivity (LSE) for the 10.4 μm channel and the aggregated MYD21_L2 LSE for MODIS band 31 (the former minus the latter) for different land cover types. Observations acquired in the four months were merged to generate the statistics. The meaning of abbreviations follows: All=All land cover types together, For=Forests, Shr=Shrublands, Sav=Savannas, Gra=Grasslands, Wat=Water, Cro=Croplands, Bar=Barren, Wet=Wetlands.

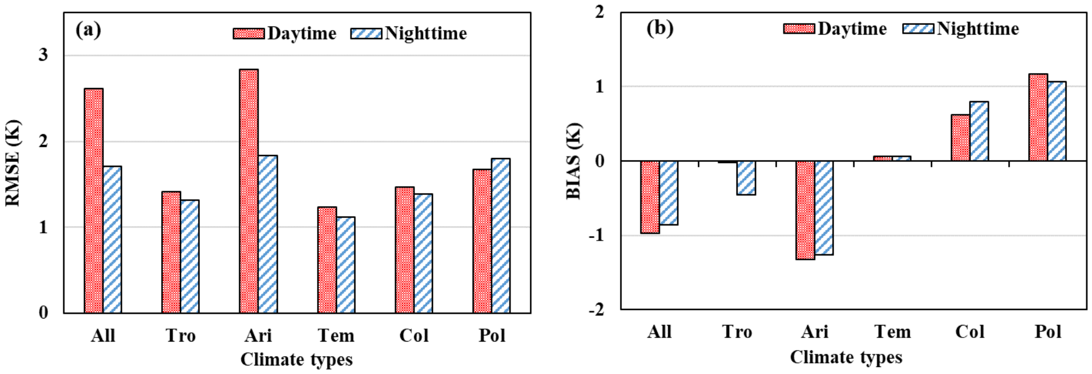


Fig. 30. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD21_L2 LST product (the former minus the latter) for different climate types. Observations acquired in the four months

were merged to generate the statistics. The meaning of abbreviations follows: All=All climate types together, Tro=Tropical, Ari=Arid, Tem=Temperate, Col=Cold, Pol=Polar.

4.4 Validation with in-situ LST measurements

In this section, the T-based method was applied to validate the performance of the extended SW-TES method based on in-situ LST measurements from seven American SURFRAD sites (Duan et al., 2019a). However, the LST was not measured directly in the SURFRAD sites while the broadband hemispherical upwelling and downwelling infrared radiation (3-50 μm) was provided. The Stephan-Boltzmann's law then was used to calculate the in-situ LST (Eq. 2).

$$LST = \left[\frac{R_{\uparrow} - (1-\varepsilon_b)R_{\downarrow}}{\sigma\varepsilon_b} \right]^{\frac{1}{4}} \qquad (2)$$

where $R_{\uparrow}$ and $R_{\downarrow}$ are respectively the upwelling and downwelling infrared radiation (W/m^2), ε_b is the surface broadband emissivity, σ is the Stephan-Boltzmann constant (5.67×10^{-8} W/m^2/K^4).

According to Eq. 2, the in-situ LST could be obtained only if the broadband emissivity of the SURFRAD sites was known. Previous studies indicated that the uncertainty of 0.01 in the broadband emissivity will introduce an error of about 0.3 K in the derived LST (Heidinger et al., 2013; Xing et al., 2021). In the seven SURFRAD sites, the surfaces are normally covered or partly covered by vegetation and their surface broadband emissivities are shown to be high and with limited variations (Duan et al., 2019a). Therefore, a fixed value of 0.97 was used in some studies to calculate the in-situ LST (Heidinger et al., 2013; Xing et al., 2021). The consequent error introduced by using this fixed value was announced to be not a dominant source of uncertainty in the SURFRAD-based performance metrics (Heidinger et al., 2013).

Therefore, the broadband emissivity of the seven SURFRAD sites were also assumed to be 0.97 in this study.

Considering that the pixel size of retrieved LST was around 13.5 km × 13.5 km, the standard deviation of MODIS LSTs inside each AIRS pixel footprint was calculated to select pixel pairs with high spatial thermal homogeneity. The corresponding ARIS pixel was considered homogenous only if the standard deviation of MODIS LSTs was <1.5 K. In addition, the data contaminated by clouds should be removed to ensure robust statistics before evaluating the discrepancy between the retrieved LST and in-situ LST measurements. However, no cloud mask layers were provided along with AIRS radiance. Thus, for each AIRS image, the MODIS cloud mask in the corresponding MYD11_L2 granule was used to identify cloudless AIRS pixels.

Taking the in-situ LST measurements from the seven SURFRAD sites in the year 2018 and 2019 as ground truth, the validation results showed that the LST could be retrieved with an RMSE of 3.58 K during the daytime (Fig. 31a). As a comparison, the uncertainties of MODIS MYD11_L2 and MYD21_L2 LST products were also evaluated and found to have the RMSE of 2.93 K (Fig. 31b) and 2.46 K (Fig. 31c), respectively. The bias of retrieved LST was around 0.42 K, which was larger than that of the MODIS MYD11_L2 LST product with the bias of about -0.95 K (Fig. 31b). Therefore, the retrieved LST was generally larger than the MODIS MYD11_L2 LST product, as described in the previous sections.

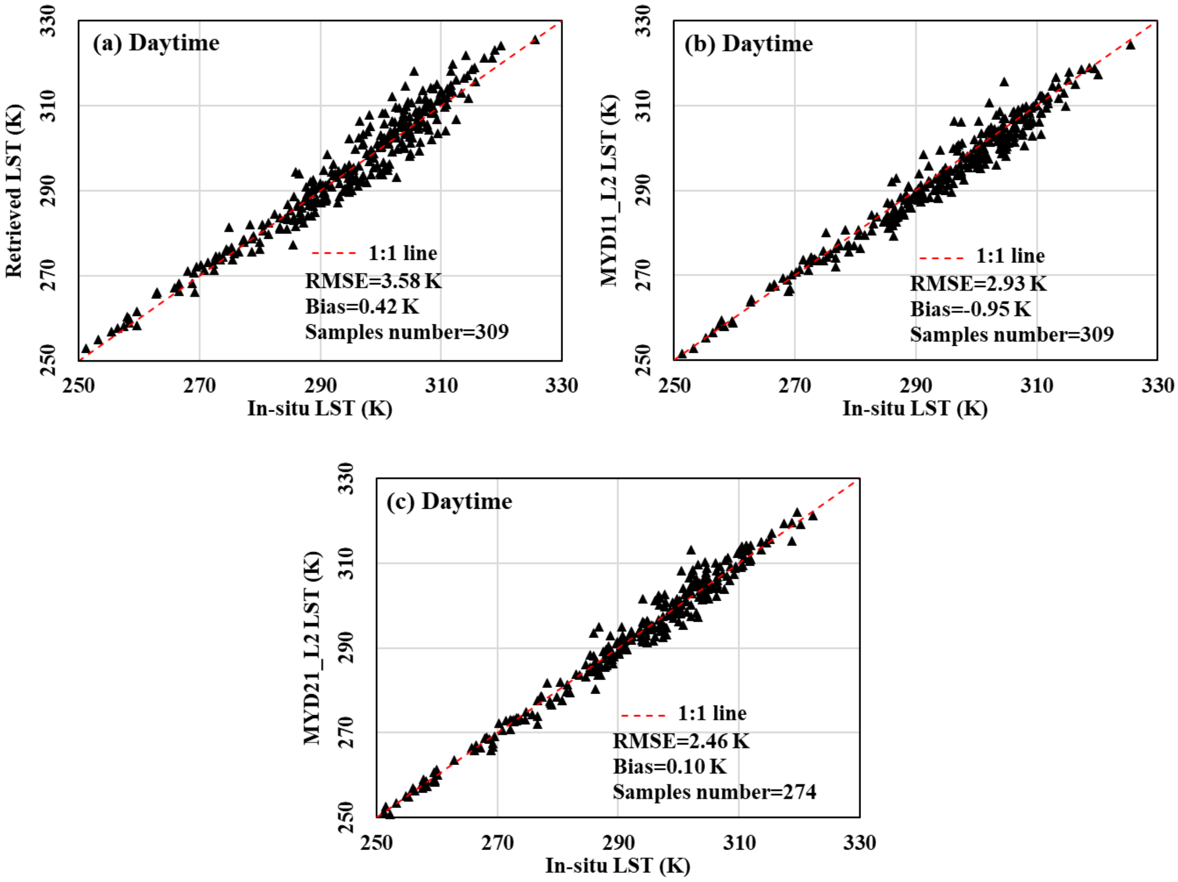


Fig. 31. In-situ land surface temperature (LST) versus (a) the retrieved LST using the extended SW-TES (SW: split window; TES: temperature-emissivity separation) method, (b) the MYD11_L2 LST product, and (c) the MYD21_L2 LST product. The observations acquired in the daytime of year 2018 and 2019 were merged to generate the statistics.

For the nighttime observations, the validation results showed that the LST could be retrieved with an RMSE of 2.35 K using the extended SW-TES method (Fig. 32a). Based on the in-situ LST measurements, the RMSEs of the MODIS MYD11_L2 and MYD21_L2 LST products were also calculated with the value of 2.29 K (Fig. 32b) and 1.64 K (Fig. 32c), respectively. In addition, the biases showed that the retrieved LST was overestimated, while the MODIS MYD11_L2 LST product was underestimated when compared with the in-situ LST. This was consistent with the results in Section 4.2, in which the retrieved LST was generally larger than the MODIS MYD11_L2 LST product under most conditions.

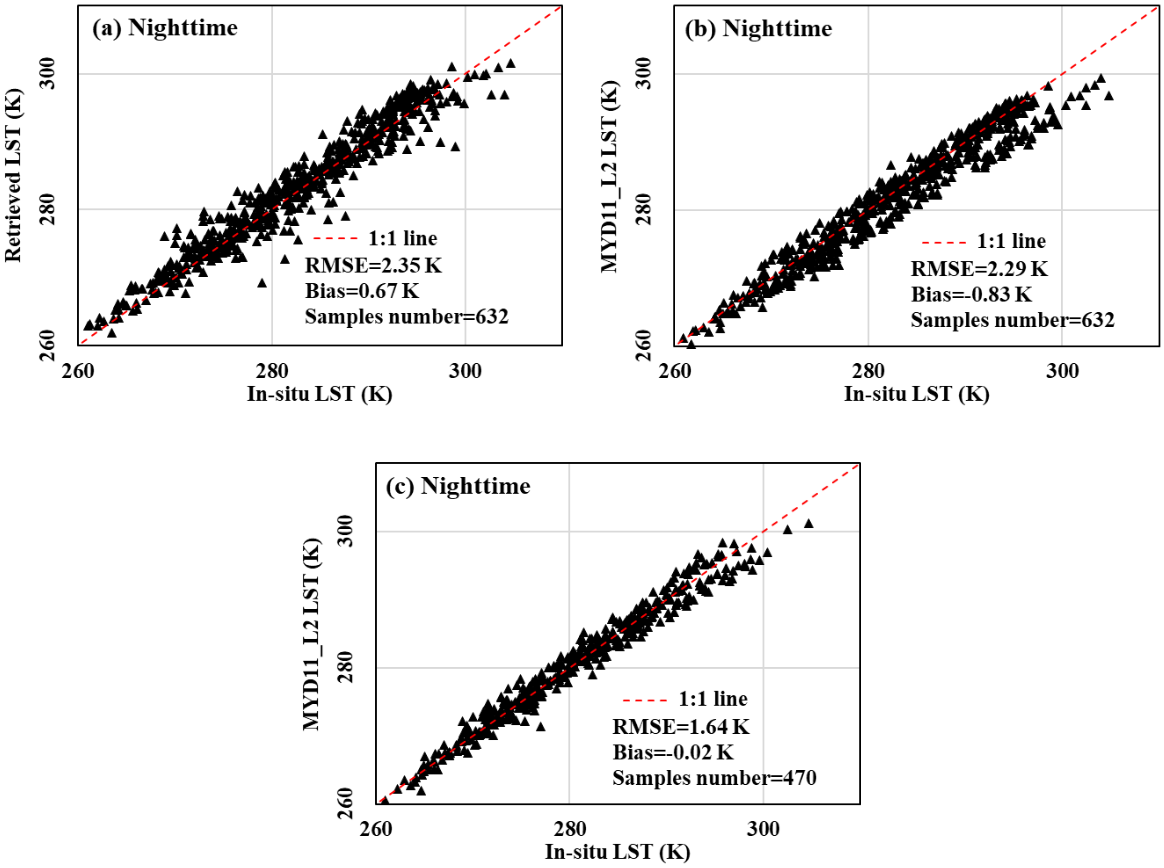


Fig. 32. In-situ land surface temperature (LST) versus (a) the retrieved LST using the extended SW-TES (SW: split window; TES: temperature-emissivity separation) method, (b) the MYD11_L2 LST product, and (c) the MYD21_L2 LST product. The observations acquired in the nighttime of year 2018 and 2019 were merged to generate the statistics.

We then investigated the performance of the extended SW-TES method in detail by calculating the individual RMSE and bias values for the retrieved LST at each SURFRAD site (Fig. 33 and Fig. 34). The uncertainties in the MYD11_L2 and MYD21_L2 LST products were also calculated using the in-situ data for comparison. The daytime evaluations revealed that the LST could be retrieved with an RMSE of <2.0 K at only two sites (Fig. 33a): the Table Mountain site (grass) and the Desert Rock site (open shrub). For the MYD11_L2 LST product, only the

Table Mountain site (grass) and Fort Peck (grass) sites produced an RMSE of <2.0 K. For the MYD21_L2 LST product, the RMSEs at the Desert Rock (open shrub) and Fort Peck (grass) sites were <2.0 K. All the observations at the other SURFRAD sites (dominated by cropland) showed that both the retrieved LST and the MYD11_L2 LST product exhibited a large discrepancy between the in-situ LST measurements. Moreover, the LST seems to be greatly overestimated at the Bondville site (cropland) while underestimated at the Goodwin Creek site (cropland/natural vegetation) with the bias being recorded at around 4.0 K and -4.0 K, respectively (Fig. 33b).

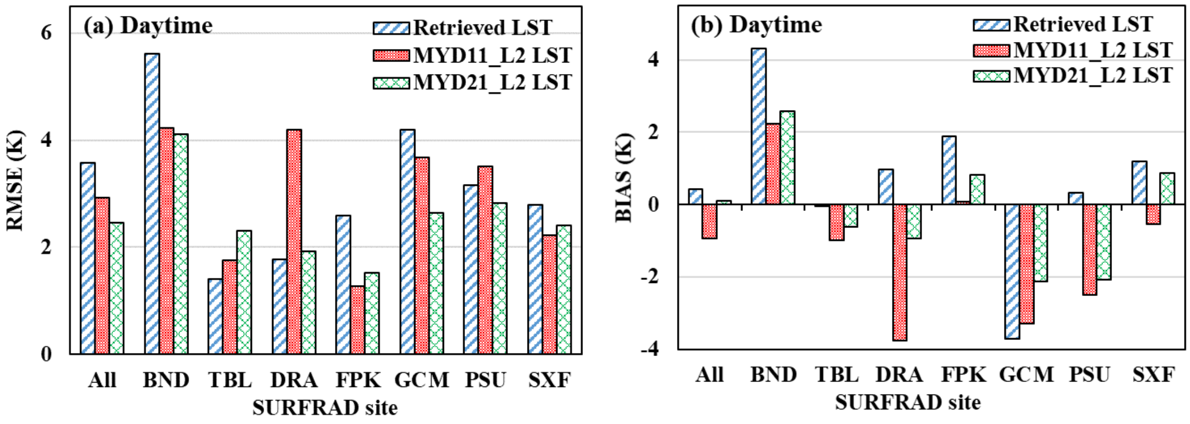


Fig. 33. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST), the MYD11_L2 LST product, and the MYD21_L2 LST product taking in-situ LST from the seven SURFRAD sites as references (retrieved LST, MYD11_LST, and MYD21_LST minus the in-situ LST). Observations acquired in the daytime of year 2018 and 2019 were merged to generate the statistics for each site separately. The meaning of abbreviations follows: All=All sites together, BND= Bondville, TBL= Table Mountain, DRA= Desert Rock, FPK= Fort Peck, GCM= Goodwin Creek, PSU= Penn State, SXF=Sioux Falls.

During the nighttime, the LST could be retrieved with an RMSE of <2.0 K except at the Desert Rock (open shrub), Goodwin Creek (cropland/natural vegetation), and the Penn State (cropland/natural vegetation) sites (Fig. 34a). But for the MYD11_L2 LST product, the RMSE was over 2.0 K only at the Desert Rock (open shrub) site. For the MYD21_L2 LST product, the RMSEs at all sites were <2.0 K except the Fort Peck site (grass) which has the RMSE of 2.07 K. Results also showed that the MYD11_L2 LST was greatly underestimated at the Desert Rock site (with the bias recorded at around -4.0 K), while the extended SW-TES method overestimated the retrieved LST at the Goodwin Creek (cropland/natural vegetation), and the Penn State (cropland/natural vegetation) sites with the bias being recorded at around 2.5 K (Fig. 34b).

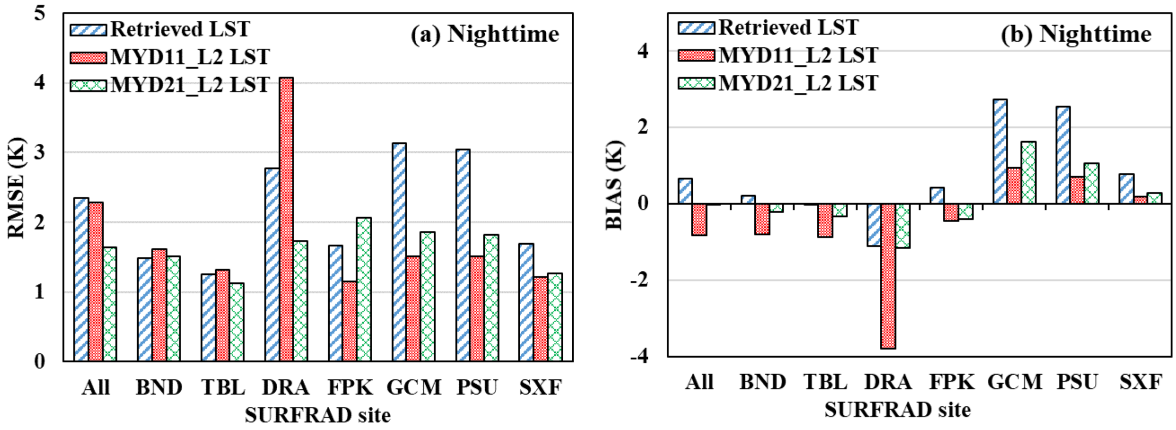


Fig. 34. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST), the MYD11_L2 LST product, and the MYD21_L2 LST product taking in-situ LST from the seven SURFRAD sites as references (retrieved LST, MYD11_LST, and MYD21_LST minus the in-situ LST). Observations acquired in the nighttime of year 2018 and 2019 were merged to generate the statistics for each site separately. The meaning of abbreviations follows: All=All sites together, BND= Bondville, TBL= Table Mountain, DRA= Desert Rock, FPK= Fort Peck, GCM= Goodwin Creek, PSU= Penn State, SXF=Sioux Falls.

5. Discussions

5.1 Uncertainties in the cross-validation results when referencing the MODIS LST products

Validation of the LST values from space is challenging, especially at the global scale. Currently, there are three commonly applied methods: The T-based method, the R-based method, and the cross-validation method (Li et al., 2013a). The cross-validation method uses the well characterized and validated LST products from other satellites as a reference to validate any new LST values retrieved from these systems. Therefore, it does not require any ground measurements.

Here, we used the widely applied and accepted MODIS LST values including the MYD11_L2 and MYD21_L2 products to complete the cross-validation analyses. The MODIS LST products are well characterized having been validated using both the T-based and R-based methods at various sites in several previous studies (Coll et al., 2009; Duan et al., 2019a; Hulley et al., 2018; Li et al., 2021; Wan, 2014). However, most of these validation sites are located within flat, homogeneous regions that meet the requirements of the T-based and R-based validation methods. The satellite observations for the other areas are generally found to be mixed signals from various ground features and may even include the effects of the surface structures. This means that the absolute accuracy of the MODIS LST products for the pixels from heterogeneous or ragged areas remains largely unknown. Thus, the discrepancy between our retrieved LST data and the MODIS LST products is likely to be associated with four key factors: first, the uncertainty in the extended SW-TES method; second, the uncertainty in the MODIS LST products; third, the uncertainty resulting from spatial-temporal misregistration and the viewing angle misregistration of the two sensors; and fourth, the uncertainty introduced during the downscaling process.

In addition to the uncertainties discussed above, differences in the channel response functions of the two sensors may also introduce additional uncertainties into the LSE comparisons as there may be differences in the emissivity signatures captured by each channel (Jacob et al., 2004). Here, we adapted the 10.4 μm channel LSE results to the channel 31 MODIS data using linear regression on the basis of the 83 typical emissivity samples and the relevant channel response functions. This method was adopted in the hopes of suppressing the influence of these additional uncertainties. However, it should be noted that the uncertainty within this linear relationship was around 0.005, which was then transferred to the LSE comparisons and may even be magnified during our downstream analyses. This means that our LSE comparison results may not be as reliable as our LST comparison results.

These factors mean that it is very challenging to complete a delicate and accurate comparison between our retrieved LST with the MODIS LST data. Additionally, differences in the channel configurations (including the channel center and width) make it even harder to compare LSE values. However, the cross-validation method could be performed across different seasons and geographies, indicating that the performance of the proposed method could be evaluated for a variety of conditions including different hemispheres, land cover types, and climatic zones, amongst others. Thus, the cross-validation method was still used in many studies (Gao et al., 2017; Qian et al., 2013; Trigo et al., 2008). In this study, a set of selection criteria was created for the pixel pairs used in each comparison according to their quality control layers and viewing geometries in an effort to increase the reliability of the comparison results. Taken together we believe that the results of the cross-validation studies still produce valuable data around the possible limitations of the proposed method as a first step.

5.2 Issues associated with the T-based validation method using ground measurements from SURFRAD sites

The SURFRAD sites were originally constructed to monitor the global surface energy budget to help advance our understanding of different regional climates (Augustine et al., 2000; Augustine et al., 2005), but have also been used to validate the LST retrieval data from satellites. These sites acquire independent measures of both upwelling and downwelling, solar and infrared radiation at each location and the surface skin temperature then could be derived using the downward and upward longwave radiation values generated from the pyrgeometers via the Stefan-Boltzmann's law (Duan et al., 2019a; Guo et al., 2020; Krishnan et al., 2020). Although these measurements have been successfully designed and applied in several previous studies to validate the LST products of various satellites (Guillevic et al., 2014; Malakar et al., 2018; Pinker et al., 2019; Wang and Liang, 2009), there have been several more recent reports that have started to identify some issues with this data.

First, the accuracy of the in-situ LST values derived from the pyrgeometer readings may be limited. For example, Guillevic et al. (2012) indicated that the instrumental uncertainty alone leads to an error of about 1.0 K in the derived LSTs from the SURFRAD sites. In addition, when the derived LST for the grassland sites were compared with those from the TIR radiometers, the authors noted an average difference in readings of up to 2.0 K (Krishnan et al., 2020). Gerace et al. (2020) also announced that the standard deviation in the differences between the LST derived data from the SURFRAD sites and those retrieved from the spaceborne TIR instruments was around 2.0 K. Since this 2.0 K residual error consistently exists, the SURFRAD measurements are likely to be particularly challenging in T-based validations requiring high degrees of accuracy. This error is likely the result of reflected solar radiation as the pyrgeometers are sensitive to radiation from 3 to 50 µm (Hulley et al., 2021).

Second, the differences in the spatial representativeness of the ground sites and satellite pixels are also a challenge for T-based validation using SURFRAD measurements especially during the daytime. Previous study has demonstrated that the shadow issues, phenological changes in the crop/grassland sites, and differences in the surface geology surrounding the Desert rock site lead to more heterogeneous surfaces (Guillevic et al., 2014; Malakar et al., 2018; Wang and Liang, 2009), making the measurements from the SURFRAD sites less reliable at the satellite pixel scale. Malakar et al. (2018) showed that the Sioux Falls, Penn State, and Goodwin Greek sites are more suitable for LST validation when using sensors with a spatial resolution of >100 m than the other four SURFRAD sites.

Given this, several studies have suggested that it is necessary to introduce more robust and homogeneous ground-measured datasets to clarify the uncertainty source in the T-based validation results (Guo et al., 2020; Hulley et al., 2021). For example, Hulley et al. (2021) suggest that the Jet Propulsion Laboratory (JPL) sites, University of Valencia (UV) sites, and Karlsruhe Institute of Technology (KIT) sites are more suitable for performing the T-based validations. These sites are all located in homogeneous areas with known emissivity characterizations and provide TIR based radiance instead of pyrgeometer based measurements. However, the SURFRAD data is an open access dataset and remains easier to obtain. Therefore, we completed our validations using the SURFRAD data. Although the RMSEs of the T-based validations using the SURFRAD data were higher than the Global Observing System for Climate recommended thresholds on accuracy, it should be noted that the uncertainties stated above were also included in the comparisons. In addition, the uncertainty in our results is comparable with that of the MODIS LST products evaluated using the same in-situ measurements as their reference.

6. Conclusions

Previous studies have successfully proposed many LST retrieval algorithms, including the widely used SW and TES methods. However, atmospheric information, LSE, or both are required in these methods to obtain accurate LST. Combining the advantages of the SW and TES algorithms, a new hybrid method was proposed by Zheng et al. (2019), which could be used to retrieve the LST without prior known atmospheric information and LSE. However, their method was not applicable to observations with large viewing angles and was only preliminarily evaluated in Australia. In this study, this method was first extended for application to global TIR observations with different viewing angles. The overall algorithm uncertainty of the extended SW-TES method was found to be around 1.16 K, according to the simulation analysis. Consequently, the performance of this extended method was satisfactorily assessed globally by referring to the MODIS LST product, and was also validated using in-situ LST measurements from seven American SURFRAD sites.

Taking the MODIS MYD11_L2 LST product as the reference, we found that the LST could be retrieved with RMSE <2.5 K for global daytime observations using the extended SW-TES method. For nighttime observations, the RMSE decreased to <1.5 K because of the absence of solar loading. The generally positive biases indicated that the retrieved LST was overestimated compared with the MODIS MYD11_L2 LST product. It should also be noted that the uncertainties of the MYD11_L2 LST product are approximately 1.0 K (Duan et al., 2018; Duan et al., 2019a; Wan, 2014), which may have also contributed to the discrepancies between the retrieved LST and MYD11_L2 LST product. Therefore, it can be concluded that the extended SW-TES method could be used to retrieve the global LST from the suggested multichannel satellite observations with acceptable accuracy.

Separate intercomparison results by season showed that for both hemispheres, the largest RMSE of retrieved LST during daytime occurred in summer, followed by spring, autumn, and winter. The RMSE exceeded 4.0 K, while the bias exceeded 3.0 K based on the results of summer and spring in the southern hemisphere. In contrast, LST retrievals coincided with the MYD11_L2 LST product much better in the northern hemisphere with RMSE <2.5 K and bias <1.0 K. For nighttime observations, no obvious relations were found between the LST retrieval RMSE and the seasons. In addition, the RMSE was shown to be <1.5 K in both the hemispheres while the bias was <0.7 K, thus indicating that the discrepancy between the retrieved LST and MYD11_L2 LST product was much smaller than the corresponding daytime results. Moreover, the variations in RMSE and bias as a function of latitude for daytime observations showed that the LST retrievals during the local cold seasons generally coincided with the MYD11_L2 LST product better than during the local hot seasons. However, for nighttime observations, the results for each latitude during different seasons were almost the same and increased gradually with increasing latitude from the South Pole to the North Pole.

From the intercomparison residuals and the statistics for different land cover types, we found that the retrieved LST was underestimated in barren regions. In contrast, the retrieved

LST was greatly overestimated in grasslands and shrublands during hot seasons. For example, in summer, the RMSE was >4.0 K for grasslands in the northern hemisphere and >8.0 K for shrublands in the southern hemisphere. Therefore, special attention should be paid when applying the extended SW-TES method to such land cover types under the above-mentioned circumstances. However, during the nighttime, the RMSE of retrieved LST compared with the MYD11_L2 LST product was within 2.0 K for all land cover types in both the hemispheres and in all seasons, except the water in the winter and the wetlands in spring and winter in northern hemisphere, the magnitude of which was approximately 2.5 K. Because the same method and coefficients were used for both daytime and nighttime observations, thermal heterogeneity resulting from the strong solar loading is believed to be the main reason for the large discrepancy in the daytime LST retrieval results.

As for the residual statistics made for different climate types, results showed that the RMSE of retrieved LST decreased gradually as the climate transited from “tropical” to “polar” for daytime observations. However, for nighttime observations, the variation pattern reversed, that is, the LST retrieval RMSE increased gradually as the climate transited from “tropical” to “polar.” Moreover, attention should be paid to daytime observations regarding tropical, arid, and temperate climates during spring and summer. The magnitude of RMSE under these circumstances could reach 4.0 K in the northern hemisphere and 6.8 K in the southern hemisphere. A possible reason for this may be that the thermal surroundings of the land surface would be more heterogeneous.

In addition to the comparison results with the MYD11_L2 LST product, the retrieved LST was also compared with the MYD21_L2 LST product. Results showed that the LST retrieval RMSE was larger by about 0.4 K than that when using MYD11_L2 LST product as a reference for both daytime and nighttime observations at the global scale. Moreover, the biases showed that the retrieved LST was generally underestimated with magnitude of about -1.0 K at the global scale when taking the MYD21_L2 LST product as a reference, while was overestimated when taking the MYD11_L2 LST product as a reference. This indicated that the LST retrieval results using the extended SW-TES method fell in between these two MODIS LST products for most circumstances globally.

In addition to the cross-validation using the MODIS LST product, the extended SW-TES method was validated based on in-situ LST measurements. It was demonstrated that the LST could be retrieved with an RMSE of 3.58 K for daytime observations and 2.35 K for nighttime observations. As a comparison, the RMSE of the MODIS MYD11_L2 LST product comparing with the same in-situ LST measurements was around 2.93 K for daytime observations and 2.29 K during the nighttime. Additionally, the RMSE of the MODIS MYD21_L2 LST product comparing with the in-situ LST was around 2.46 K for daytime observations and 1.64 K during the nighttime. Therefore, it was concluded that the extended SW-TES method could be used to retrieve LST with an accuracy close to that of the MODIS LST product. Compared with previous studies, the extended SW-TES method had the advantages of the method proposed by Zheng et al. (2019), which does not require any prior knowledge of the atmosphere and LSE during LST retrieval. In addition, this method could be used to retrieve LST from global TIR measurements at different viewing angles.

Nevertheless, additional efforts are required to further improve LST retrieval accuracy. For example, because the biases of retrieved LST were found to vary with the land cover and climate types, and were different for daytime and nighttime observations, obtaining the coefficients of the SW-like equation separately according to the land cover type, climate type, and day/nighttime may improve the performance of the proposed method. Moreover, the coefficients of the SW-like equation can also be parameterized as a function of the atmospheric water vapor content to increase the ground brightness temperature retrieval accuracy under humid atmospheric conditions (Niclos et al., 2011; Pérez-Planells et al., 2021). Detailed investigations are planned in one of our succeeding works. Besides, the angle effect of surface emissivity maybe significant for TIR measurements with large viewing angles. Therefore, the empirical relationship between minimum emissivity and emissivity contrast could also be adapted for large viewing angles to improve the LST and LSE separation accuracy.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 41921001) and also by the China Scholarship Council (CSC). Thanks would be given to Dr. Xiaoyue GONG for her help in the English language editing and in the collection of satellite data. Thanks would also be given to the anonymous reviewers for their valuable comments, which have been very useful in improving the manuscript quality.

References

Anderson, M.C., Norman, J.M., Kustas, W.P., Houborg, R., Starks, P.J., 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 Sens. Environ. 112, 4227-4241.

Augustine, J.A., DeLuisi, J.J., Long, C.N., 2000. Surfrad—a national surface radiation budget network for atmospheric research. Bull. Am. Meteorol. Soc. 81, 2341-2357.

Augustine, J.A., Hodges, G.B., Cornwall, C.R., Michalsky, J.J., Medina, C.I., 2005. An update on surfrad—the gcos surface radiation budget network for the continental united states. J. Atmos. Ocean. Technol. 22, 1460-1472.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future koppen-geiger climate classification maps at 1-km resolution. Sci. Data. 5, 1-12.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2020. Publisher correction: Present and future koppen-geiger climate classification maps at 1-km resolution. Sci. Data. 7, 1-2.

Becker, F., 1987. The impact of spectral emissivity on the measurement of land surface temperature from a satellite. Int. J. Remote Sens. 8, 1509-1522.

Becker, F., Li, Z.-L., 1990. Towards a local split window method over land surfaces. Int. J. Remote Sens. 11, 369-393.

Berk, A., Cooley, T.W., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Lewis, P.E., 2004. Modtran5: A reformulated atmospheric band model with auxiliary species and practical multiple scattering options. In, *Proc. SPIE 5571, Remote Sensing of Clouds and the Atmosphere IX* (pp. 78-85). Maspalomas, Canary Islands, Spain: SPIE.

Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E., Shettle, E.P., 2006. Modtran5: 2006 update. In, *SPIE 6233, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII* (pp. 1-8). Orlando (Kissimmee), Florida, United States: SPIE.

Chen, Y., Duan, S.-B., Ren, H., Labed, J., Li, Z.-L., 2017. Algorithm development for land surface temperature retrieval: Application to chinese gaofen-5 data. Remote Sens. 9, 1-19.

Chevallier, F., Chéruy, F., Scott, N.A., Chédin, A., 1998. A neural network approach for a fast and accurate computation of a longwave radiative budget. J. Appl. Meteorol. 37, 1385-1397.

Coll, C., Caselles, V., 1997. A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison. J. Geophys. Res. 102, 16697-16713.

Coll, C., Wan, Z., Galve, J.M., 2009. Temperature-based and radiance-based validations of the v5 modis land surface temperature product. J. Geophys. Res. 114, 1-15.

Duan, S.-B., Li, Z.-L., Cheng, J., Leng, P., 2017. Cross-satellite comparison of operational land surface temperature products derived from modis and ASTER data over bare soil surfaces. ISPRS J. Photogramm. Remote Sens. 126, 1-10.

Duan, S.-B., Li, Z.-L., Wu, H., Leng, P., Gao, M., Wang, C., 2018. Radiance-based validation of land surface temperature products derived from collection 6 modis thermal infrared data. Int. J. Appl. Earth Obs. Geoinf. 70, 84-92.

Duan, S.-B., Li, Z.-L., Li, H., Göttsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., Coll, C., 2019a. Validation of collection 6 modis land surface temperature product using in situ measurements. Remote Sens. Environ. 225, 16-29.

Duan, S.-B., Li, Z.-L., Wang, C., Zhang, S., Tang, B.-H., Leng, P., Gao, M.-F., 2019b. Land-surface temperature retrieval from landsat 8 single-channel thermal infrared data in combination with ncep reanalysis data and ASTER ged product. Int. J. Remote Sens. 40, 1763-1778.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sens. Environ. 114, 168-182.

Gao, C., Jiang, X., Wu, H., Tang, B., Li, Z., Li, Z.-L., 2012. Comparison of land surface temperatures from msg-2/seviri and terra/modis. J. Appl. Remote Sens. 6, 1-16.

Gao, C., Li, Z.-L., Qiu, S., Tang, B., Wu, H., Jiang, X., 2014. An improved algorithm for retrieving land surface emissivity and temperature from msg-2/seviri data. IEEE Trans. Geosci. Remote Sens. 52, 3175-3191.

Gao, C., Qiu, S., Zhao, E.-Y., Li, C., Tang, L.-L., Ma, L.-L., Jiang, X., Qian, Y., Zhao, Y., Wang, N., Ren, L., 2017. Land surface temperature retrieval from fy-3c/virr data and its cross-validation with terra/modis. IEEE J-Stars. 10, 4944-4953.

Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an operational, split window-derived surface temperature product for the thermal infrared sensors onboard landsat 8 and 9. Remote Sens. 12, 1-15.

Ghent, D.J., Corlett, G.K., Göttsche, F.M., Remedios, J.J., 2017. Global land surface temperature from the along-track scanning radiometers. J. Geophys. Res. Atmos. 122, 12167-12193.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J.S., Hook, S., Kahle, A.B., 1998. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. IEEE Trans. Geosci. Remote Sens. 36, 1113-1126.

Guillevic, P.C., Privette, J.L., Coudert, B., Palecki, M.A., Demarty, J., Ottlé, 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 Sens. Environ. 124, 282-298.

Guillevic, P.C., Biard, J.C., Hulley, G.C., Privette, J.L., Hook, S.J., Olioso, A., Göttsche, F.M., Radocinski, R., Román, M.O., Yu, Y., Csiszar, I., 2014. Validation of land surface temperature products derived from the visible infrared imaging radiometer suite (viirs) using ground-based and heritage satellite measurements. Remote Sens. Environ. 154, 19-37.

Guo, J., Ren, H., Zheng, Y., Lu, S., Dong, J., 2020. Evaluation of land surface temperature retrieval from landsat 8/tirs images before and after stray light correction using the surfrad dataset. Remote Sens. 12, 1-21.

Heidinger, A.K., Laszlo, I., Molling, C.C., Tarpley, D., 2013. Using surfrad to verify the noaa single-channel land surface temperature algorithm. J. Atmos. Ocean. Technol. 30, 2868-2884.

[dataset] Hulley, G. (2015), Myd21 modis/aqua land surface temperature/3-band emissivity 5-min l2 1km v006. U.S.A, Accessed: [Aug. 12, 2021], 10.5067/MODIS/MOD11_L2.006.

Hulley, G., Malakar, N., Freepartner, R. (2016). Moderate resolution imaging spectroradiometer (modis) land surface temperature and emissivity product (mxd21) algorithm theoretical basis document collection-6. Pasadena, California. J.P. Laboratory.

Hulley, G.C., Hook, S.J., 2011. Generating consistent land surface temperature and emissivity products between ASTER and modis data for earth science research. IEEE Trans. Geosci. Remote Sens. 49, 1304-1315.

Hulley, G.C., Hughes, C.G., Hook, S.J., 2012. Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and modis thermal infrared data. J. Geophys. Res. Atmos. 117, 1-18.

Hulley, G.C., Malakar, N.K., Islam, T., Freepartner, R.J., 2018. Nasa's modis and viirs land surface temperature and emissivity products: A long-term and consistent earth system data record. IEEE J-Stars. 11, 522-535.

Hulley, G.C., Gottsche, F.M., Rivera, G., Hook, S.J., Freepartner, R.J., Martin, M.A., Cawse-Nicholson, K., Johnson, W.R., 2021. Validation and quality assessment of the ecostress level-2 land surface temperature and emissivity product. IEEE Trans. Geosci. Remote Sens. Early Access, 1-23.

Jacob, F., Petitcolin, F.o., Schmugge, T., Vermote, É., French, A., Ogawa, K., 2004. Comparison of land surface emissivity and radiometric temperature derived from modis and ASTER sensors. Remote Sens. Environ. 90, 137-152.

Jiang, G.-M., Li, Z.-L., 2008. Split‐window algorithm for land surface temperature estimation from msg1‐seviri data. Int. J. Remote Sens. 29, 6067-6074.

Jimenez-Munoz, J.C., Sobrino, J.A., 2003. A generalized single-channel method for retrieving land surface temperature from remote sensing data. J. Geophys. Res. 108, 1-9.

Jimenez-Munoz, J.C., Cristobal, J., Sobrino, J.A., Soria, G., Ninyerola, M., Pons, X., Pons, X., 2009. Revision of the single-channel algorithm for land surface temperature retrieval from landsat thermal-infrared data. IEEE Trans. Geosci. Remote Sens. 47, 339-349.

Kappas, M., Phan, T.N., 2018. Application of modis land surface temperature data: A systematic literature review and analysis. J. Appl. Remote Sens. 12, 1-20.

Krishnan, P., Meyers, T.P., Hook, S.J., Heuer, M., Senn, D., Dumas, E.J., 2020. Intercomparison of in situ sensors for ground-based land surface temperature measurements. Sensors (Basel) 20, 1-26.

Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J., Scott, J.K., 2012. Climond: Global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. Methods Ecol. Evol. 3, 53-64.

Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., Liu, Q., 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 Trans. Geosci. Remote Sens. 59, 1794-1807.

Li, S., Yu, Y., Sun, D., Tarpley, D., Zhan, X., Chiu, L., 2014. Evaluation of 10 year aqua/modis land surface temperature with surfrad observations. Int. J. Remote Sens. 35, 830-856.

Li, Z.-L., Tang, R., Wan, Z., Bi, Y., Zhou, C., Tang, B., Yan, G., Zhang, X., 2009. A review of current methodologies for regional evapotranspiration estimation from remotely sensed data. Sensors (Basel) 9, 3801-3853.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., Sobrino, J.A., 2013a. Satellite-derived land surface temperature: Current status and perspectives. Remote Sens. Environ. 131, 14-37.

Li, Z.-L., Wu, H., Wang, N., Qiu, S., Sobrino, J.A., Wan, Z., Tang, B.-H., Yan, G., 2013b. Land surface emissivity retrieval from satellite data. Int. J. Remote Sens. 34, 3084-3127.

Liu, Y., Yu, Y., Yu, P., Göttsche, F., Trigo, I., 2015. Quality assessment of s-npp viirs land surface temperature product. Remote Sens. 7, 12215-12241.

Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An operational land surface temperature product for landsat thermal data: Methodology and validation. IEEE Trans. Geosci. Remote Sens. 56, 5717-5735.

McMillin, L.M., 1975. Estimation of sea surface temperatures from two infrared window measurements with different absorption. J. Geophys. Res. 80, 5113-5117.

Niclos, R., Galve, J.M., Valiente, J.A., Estrela, M.J., Coll, C., 2011. Accuracy assessment of land surface temperature retrievals from msg2-seviri data. Remote Sens. Environ. 115, 2126-2140.

O'Connor, B., Bojinski, S., Röösli, C., Schaepman, M.E., 2020. Monitoring global changes in biodiversity and climate essential as ecological crisis intensifies. Ecol. Inform. 55, 1-8.

Ottlé, C., Vidal-Madjar, D., 1992. Estimation of land surface temperature with noaa9 data. Remote Sens.

Environ. 40, 27-41.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the köppen-geiger climate classification. Hydrol. Earth Syst. Sci. 11, 1633-1644.

Pérez-Planells, L., Niclòs, R., Puchades, J., Coll, C., Göttsche, F.-M., Valiente, J.A., Valor, E., Galve, J.M., 2021. Validation of sentinel-3 slstr land surface temperature retrieved by the operational product and comparison with explicitly emissivity-dependent algorithms. Remote Sens. 13, 1-25.

Pinker, R.T., Ma, Y., Chen, W., Hulley, G., Borbas, E., Islam, T., Hain, C., Cawse-Nicholson, K., Hook, S., Basara, J., 2019. Towards a unified and coherent land surface temperature earth system data record from geostationary satellites. Remote Sens. 11, 1-23.

Qian, Y.-G., Li, Z.-L., Nerry, F., 2013. Evaluation of land surface temperature and emissivities retrieved from msg/seviri data with modis land surface temperature and emissivity products. Int. J. Remote Sens. 34, 3140-3152.

Qin, Z., Dall'Olmo, G., Karnieli, A., Berliner, P., 2001a. Derivation of split window algorithm and its sensitivity analysis for retrieving land surface temperature from noaa-advanced very high resolution radiometer data. J. Geophys. Res.-Atmos. 106, 22655-22670.

Qin, Z., Karnieli, A., Berliner, P., 2001b. A mono-window algorithm for retrieving land surface temperature from landsat tm data and its application to the israel-egypt border region. Int. J. Remote Sens. 22, 3719-3746.

Ren, H.Z., Dong, J.J., Liu, R.Y., Zheng, Y.T., Guo, J.X., Chen, S.S., Nie, J., Zhao, Y., 2020. New hybrid algorithm for land surface temperature retrieval from multiple-band thermal infrared image without atmospheric and emissivity data inputs. Int. J. Digit. Earth. 13, 1430-1453.

Sánchez-Aparicio, M., Andrés-Anaya, P., Del Pozo, S., Lagüela, S., 2020. Retrieving land surface temperature from satellite imagery with a novel combined strategy. Remote Sens. 12, 1-18.

Sattari, F., Hashim, M., 2014. A breife review of land surface temperature retrieval methods from thermal satellite sensors. Middle East J. Sci. Res. 22, 757-768.

Snyder, W.C., Wan, Z., Zhang, Y., Feng, Y.Z., 1998. Classification-based emissivity for land surface temperature measurement from space. Int. J. Remote Sens. 19, 2753-2774.

Sobrino, J.A., Coll, C., Caselles, V., 1991. Atmospheric correction for land surface temperature using noaa-11 avhrr channels 4 and 5. Remote Sens. Environ. 38, 19-34.

Sobrino, J.A., Del Frate, F., Drusch, M., Jimenez-Munoz, J.C., Manunta, P., Regan, A., 2016. Review of thermal infrared applications and requirements for future high-resolution sensors. IEEE Trans. Geosci. Remote Sens. 54, 2963-2972.

Sulla-Menashe, D., Friedl, M.A. (2018). User guide to collection 6 modis land cover (mcd12q1 and mcd12c1) product. South Dakota USA. N.E.L.P. DAAC.

Sulla-Menashe, D., Gray, J.M., Abercrombie, S.P., Friedl, M.A., 2019. Hierarchical mapping of annual global land cover 2001 to present: The modis collection 6 land cover product. Remote Sens. Environ. 222, 183-194.

Tomlinson, C.J., Chapman, L., Thornes, J.E., Baker, C., 2011. Remote sensing land surface temperature for meteorology and climatology: A review. Meteorol. Appl. 18, 296-306.

Tonooka, H., 2005. Accurate atmospheric correction of ASTER thermal infrared imagery using the wvs method. IEEE Trans. Geosci. Remote Sens. 43, 2778-2792.

Townshend, J.R.G., Justice, C.O., Skole, D., Malingreau, J.P., Cihlar, J., Teillet, P., Sadowski, F., Ruttenberg, S., 1994. The 1 km resolution global data set: Needs of the international geosphere biosphere programme. Int. J. Remote Sens. 15, 3417-3441.

Trigo, I.F., Monteiro, I.T., Olesen, F., Kabsch, E., 2008. An assessment of remotely sensed land surface temperature. J. Geophys. Res. 113, 1-12.

Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 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 Trans. Geosci. Remote Sens. 35, 980-996.

Wan, Z., 2008. New refinements and validation of the modis land-surface temperature/emissivity products. Remote Sens. Environ. 112, 59-74.

Wan, Z., Li, Z.-L., 2008. Radiance ‐ based validation of the v5 modis land ‐ surface temperature product. Int. J. Remote Sens. 29, 5373-5395.

Wan, Z., 2014. New refinements and validation of the collection-6 modis land-surface temperature/emissivity product. Remote Sens. Environ. 140, 36-45.

[dataset] Wan, Z., Hook, S., Hulley, G. (2015), Mod11_l2 modis/terra land surface temperature/emissivity 5-min l2 swath 1km v006. U.S.A, Accessed: [Feb. 5, 2019], 10.5067/MODIS/MOD11_L2.006.

Wang, H., Mao, K., Mu, F., Shi, J., Yang, J., Li, Z., Qin, Z., 2019a. A split window algorithm for retrieving land surface temperature from fy-3d mersi-2 data. Remote Sens. 11, 1-25.

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 Sens. Environ. 113, 1556-1565.

Wang, M., He, G., Zhang, Z., Wang, G., Wang, Z., Yin, R., Cui, S., Wu, Z., Cao, X., 2019b. A radiance-based split-window algorithm for land surface temperature retrieval: Theory and application to modis data. Int. J. Appl. Earth Obs. Geoinf. 76, 204-217.

Xing, Z., Li, Z.-L., Duan, S.-B., Liu, X., Zheng, X., Leng, P., Gao, M., Zhang, X., Shang, G., 2021. Estimation of daily mean land surface temperature at global scale using pairs of daytime and nighttime modis instantaneous observations. ISPRS J. Photogramm. Remote Sens. 178, 51-67.

Yao, R., Wang, L., Huang, X., Gong, W., Xia, X., 2019. Greening in rural areas increases the surface urban heat island intensity. Geophys. Res. Lett. 46, 2204-2212.

Yao, R., Wang, L., Wang, S., Wang, L., Wei, J., Li, J., Yu, D., 2020. A detailed comparison of myd11 and myd21 land surface temperature products in mainland china. Int. J. Digit. Earth. 13, 1-17.

Zheng, X., Li, Z.-L., Nerry, F., Zhang, X., 2019. A new thermal infrared channel configuration for accurate land surface temperature retrieval from satellite data. Remote Sens. Environ. 231, 1-12.