



HAL
open science

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

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

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

4 Xiaopo Zheng ^{a, c}, Zhao-Liang Li ^{b, c, *}, Tianxing Wang ^a, Huabing Huang ^a, Françoise Nerry ^c

5 ^a School of Geospatial Engineering and Science, Sun Yat-Sen University, 519082 Guangdong, China

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

8 ^c ICube, UdS, CNRS, 300 Bld Sébastien Brant, CS 10413, 67412 Illkirch, France

9

10 **Abstract:**

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

* Corresponding author: lizl@unistra.fr

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

35

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

38

39 **1. Introduction**

40 Land surface temperature (LST) is an important factor that influences water and energy
41 exchange between the Earth's surface and the atmosphere (Kappas and Phan, 2018; Sobrino et
42 al., 2016; Tomlinson et al., 2011). It has been widely used in many studies, including
43 climatology, meteorology, hydrology, and ecology (Anderson et al., 2008; Kappas and Phan,
44 2018; Li et al., 2009; O'Connor et al., 2020; Yao et al., 2019). Therefore, LST has been listed
45 as a high-priority parameter for the International Geosphere-Biosphere Programme (IGBP)

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

49 Currently, thermal infrared (TIR) remote sensing is considered the most appropriate method
50 for obtaining accurate LST, both regionally and globally (Li et al., 2013a; Sánchez-Aparicio et
51 al., 2020). Over several decades, a wide variety of LST retrieval algorithms have been
52 developed and can be roughly classified into four categories as follows (Li et al., 2013a; Li et
53 al., 2013b; Sattari and Hashim, 2014). First, the single-channel method, which requires both
54 atmospheric profiles and LSE to obtain the LST from the observed TIR radiance (Duan et al.,
55 2019b; Jimenez-Munoz and Sobrino, 2003; Otlé and Vidal-Madjar, 1992; Qin et al., 2001b)
56 (Jimenez-Munoz et al., 2009; Sánchez-Aparicio et al., 2020). Second, the split window (SW)
57 method that can eliminate atmospheric effects by directly using the observations of two adjacent
58 channels (Becker and Li, 1990; Coll and Caselles, 1997; McMillin, 1975; Qin et al., 2001a;
59 Wan and Dozier, 1996; Wang et al., 2019b). But the performance of this method is shown to be
60 sensitive to the channel LSE uncertainties, implying that accurate LSE of the study area must
61 be known in advance (Jiang and Li, 2008; Wang et al., 2019a). Third, the temperature/emissivity
62 separation (TES) method that can be used to simultaneously separate LST and LSE (Gillespie
63 et al., 1998). However, accurate atmospheric correction is needed before applying this method
64 (Hulley and Hook, 2011; Hulley et al., 2018; Ren et al., 2020). Four, the day/night methods that
65 could also be used to retrieve LST and LSE at the same time (Wan and Li, 1997; Wan, 2008;
66 Wan, 2014). By using two geographically matched images and assuming the LSE is unchanged
67 in the two images, this method requires only the shape information of atmospheric profiles (Wan

68 and Li, 1997). But the assumption may not be true because of the spatial misregistration or
69 variations in land surface conditions (Gao et al., 2014; Wan and Li, 1997).

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

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

101 **2. Data**

102 **2.1 AIRS data**

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

112 specifications of the AIRS instrument are listed in Table 1.

113 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

114 **2.2 MODIS LST product for cross validation**

115 Because it is difficult to obtain the ground-measured LST globally, the well-validated and
116 commonly used MODIS LST product (MYD11_L2) was introduced as a reference for
117 intercomparison (Wan et al., 2015). The MYD11_L2 product was generated at a spatial
118 resolution of 1 km from the generalized split-window algorithm (GSW), which requires prior-
119 known LSE information (Wan and Dozier, 1996). Similar to other SW methods, the
120 performance of the GSW algorithm is sensitive to uncertainties in the emissivity difference
121 between the two adjacent channels. Currently, the LSE of MODIS bands 31 and 32 are estimated
122 by the classification-based method (Snyder et al., 1998). This method firstly classifies each
123 pixel according to the input data in Land Cover (MCDLC1KM) and dynamic and seasonal
124 factors. For different land cover types, the appropriate kernel models were applied to determine

125 the bidirectional reflectance distribution function models (BRDF), and then to obtain the
126 hemispherical-directional reflectance, $\rho(\theta)$. Finally, the emissivity could be computed by
127 $1-\rho(\theta)$ (Snyder et al., 1998). However, the misclassification of the land cover type or
128 inhomogeneity inside a single MODIS pixel still may decrease the LST retrieval accuracy.

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

139 In addition to the MYD11_L2 product, a newer MODIS LST product, MYD21_L2, was
140 also available currently (Hulley, 2015). The MODIS MYD21_L2 product was generated using
141 the TES method which was originally designed for the Advanced Spaceborne Thermal
142 Emission Reflection Radiometer (ASTER) instrument (Hulley and Hook, 2011). Comparing
143 with the GSW method used in the MYD11_L2 product, the LST and LSE were retrieved
144 simultaneously in the MYD21_L2 product. However, accurate atmospheric correction was
145 firstly required before the LST retrieval (Hulley and Hook, 2011). The auxiliary atmospheric

146 products (Hulley et al., 2016) and the Water Vapor Scaling (WVS) Method (Tonooka, 2005)
147 therefore were used to remove the atmospheric effect. The uncertainty of the MYD21_L2 LST
148 product was announced to be <1.5 K (Hulley et al., 2016; Hulley et al., 2012).

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

156 **2.3 In-situ LST measurements for direct validation**

157 To provide a direct evaluation of the performance of the extended SW-TES method, the T-
158 based validation method was used in this study to compare the derived LST with the in-situ
159 LST measured at the satellite overpass in a simple manner. The T-based method requires the
160 ground LST to be measured accurately at the satellite passing time and is adequate for
161 representing the LST at the satellite pixel scale (Li et al., 2013a). Therefore, collection of in-
162 situ LST measurements is quite difficult, and the available datasets are very limited. The
163 SURFRAD sites built in 1993 provide continuous, long-term, accurate, and quality-controlled
164 measurements of broadband hemispherical upwelling and downwelling longwave radiation (3-
165 50 μm) at seven different locations in the United States (Augustine et al., 2000). After applying
166 the Stefan–Boltzmann law, in-situ LST could be obtained every 3 minutes (before 2009) or

167 every 1 minute (after 2009) from the measured longwave radiation (Duan et al., 2019a). The
 168 SURFRAD measurements are currently widely used to evaluate the uncertainties of various
 169 satellite LST products, such as ASTER (Wang and Liang, 2009), AATSR (Ghent et al., 2017),
 170 MODIS (Duan et al., 2019a; Li et al., 2014; Wang and Liang, 2009), and VIIRS (Guillevic et
 171 al., 2014; Liu et al., 2015). In this study, the SURFRAD measurements were chosen to provide
 172 a direct validation of the extended SW-TES method described in Section 3. Relevant
 173 information on the seven SURFRAD sites is listed in Table 2.

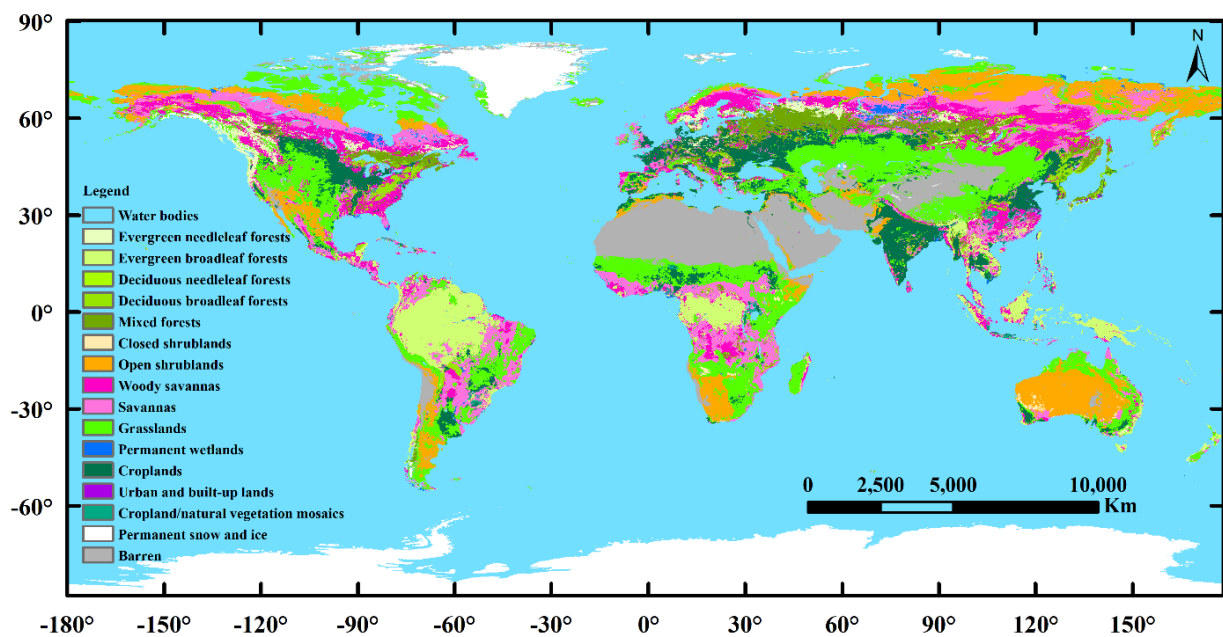
174 Table 2. Detailed information of the seven SURFRAD sites.

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

175 **2.4 MODIS land cover product**

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

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



186
 187 Fig. 1. Global land cover types according to the IGBP layer in MCD12C1 C6 product for the year 2018.

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

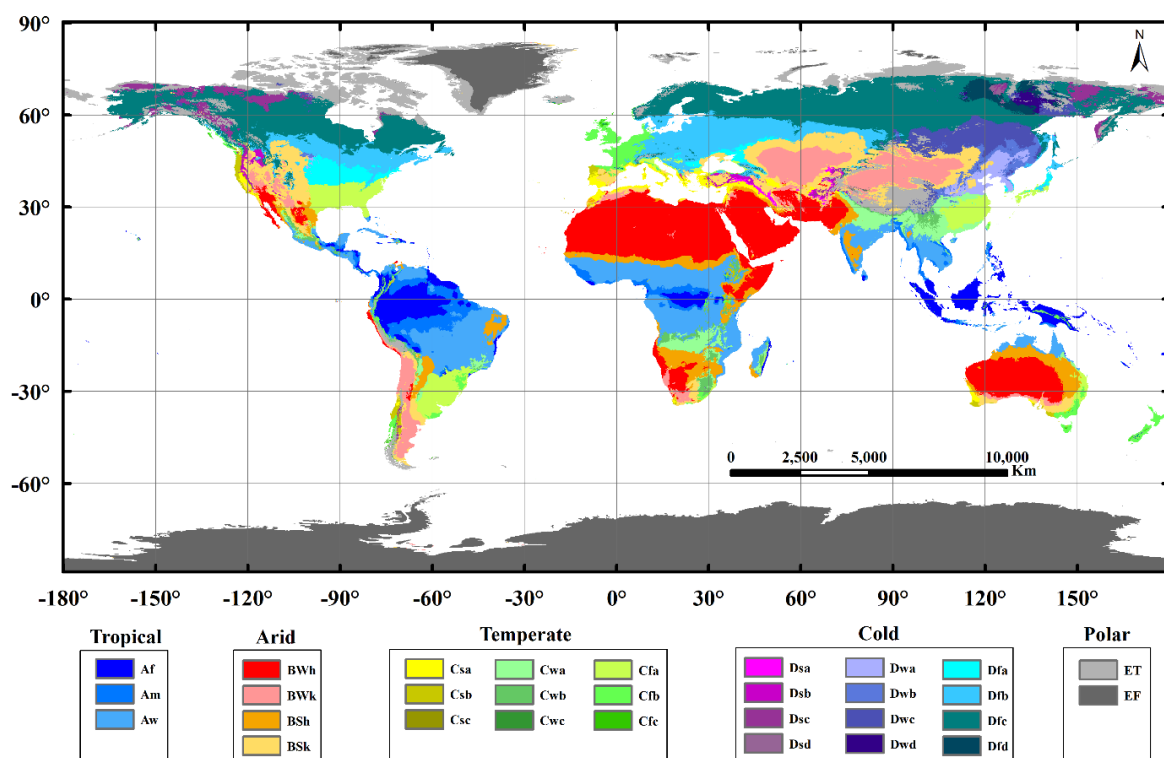
195 Table 3. Legends of the 17 original MODIS IGBP land cover types and the merged land cover
 196 types used in this study.

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

197 **2.5 Global climate classification map**

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

199 TES method, a Köppen-Geiger climate classification map was introduced (Beck et al., 2018).
 200 According to previous studies (Kriticos et al., 2012; Peel et al., 2007), the Köppen-Geiger
 201 system classifies the global climate into five main climate types (i.e., tropical, arid, temperate,
 202 cold, and polar) and 30 sub-types (Table 4). The map is generated according to the air
 203 temperature and precipitation data of the world using a threshold-based method (Beck et al.,
 204 2018). The Köppen-Geiger climate classification map used in this study was derived at a spatial
 205 resolution of 1-km by combining climatic air temperature and precipitation data from multiple
 206 independent sources, as shown in Fig. 2 (Beck et al., 2018, 2020). The spatial resolution was
 207 unprecedented compared to previous versions (Beck et al., 2018). Moreover, the topographic
 208 effects were also corrected explicitly in this version, which could provide more accurate results
 209 in mountainous areas (Beck et al., 2018).



210
 211 Fig. 2. Global Köppen-Geiger climate classification map (referring to Table 4 for abbreviations of each sub-
 212 classes).

213 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
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

214 **3. Methodology**

215 **3.1 Extension of the SW-TES method to different local viewing zenith angles**

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

230 In this study, the SW-like equation was optimized following previous studies to suppress
231 the angular effects during ground brightness temperature retrieval (Wan and Dozier, 1996).
232 Specifically, the coefficients in Eq. (1) were obtained as a function of the local viewing zenith
233 angle, rather than constants (Niclos et al., 2011; Pérez-Planells et al., 2021). Similar to the
234 traditional SW method, these coefficients could be determined by regression analysis of the
235 simulation data (Wan and Dozier, 1996). First, the simulated satellite TIR measurements for

236 seven local viewing zenith angles ($1 / \cos(\theta_v) = 1.00, 1.15, 1.30, 1.45, 1.60, 1.75,$ and 1.90 , with
 237 $\theta_v = 0.00^\circ, 29.59^\circ, 39.72^\circ, 46.40^\circ, 51.32^\circ, 55.15^\circ,$ and 58.24° , respectively) were generated.
 238 Then, the ground brightness temperature at different local viewing zenith angles could be
 239 acquired feasibly using this angle-dependent SW-like equation [Eq. (1)].

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

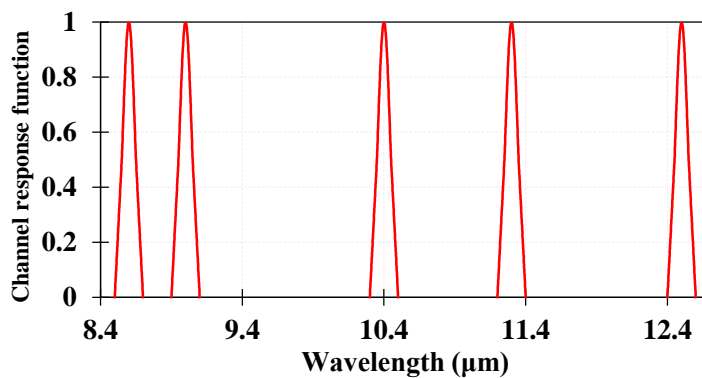
240 where T_{gi} is the ground brightness temperature of channel i . $A_0(\theta_v)$, $A_1(\theta_v)$, $A_2(\theta_v)$, and
 241 $A_3(\theta_v)$ are the regression coefficients that vary with the local viewing zenith angle θ_v . Here, i
 242 and j represent the channels. T_i and T_j are the brightness temperatures of channels i and j ,
 243 respectively.

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

246 **3.2 Simulation dataset**

247 To generate the simulation dataset for determining the coefficients of the SW-like equation
 248 [Eq. (1)], the five suggested channel response functions in the study by Zheng et al. (2019) were
 249 introduced. According to their algorithm sensitivity analysis, the performance of the SW-TES
 250 method decreased as the channel width (full width at half maximum, FWHM) and channel noise
 251 (noise equivalent differential temperature, NE Δ T) increased (Zheng et al., 2019). Based on
 252 numerical simulations, the channel FWHM and NE Δ T of possible future sensors should be
 253 respectively limited within $0.1 \mu\text{m}$ and 0.1 K if the LST retrieval accuracy of 1.0 K is required.
 254 For those sensors with channel FWHM $<0.6 \mu\text{m}$, the LST could be retrieved with RMSE <1.2
 255 K and $<1.5 \text{ K}$ provided that the channel NE Δ T is $<0.1 \text{ K}$ and $<0.2 \text{ K}$, respectively (Zheng et al.,

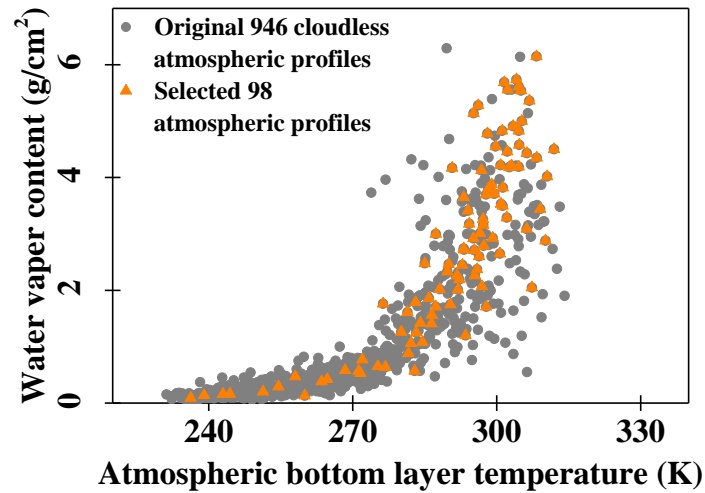
256 2019). In this study, considering that the spectral resolution of AIRS is ultra-fine, the ideal
257 channel FWHM of $0.1 \mu\text{m}$ was therefore used in the following calculations (Fig. 3). Together
258 with different ground-measured LSE spectra, atmospheric profiles, LST inputs, and sensor
259 viewing angles, the satellite-observed channel brightness temperature under various imaging
260 conditions could be simulated using the MODTRAN (MODerate resolution atmospheric
261 TRANsmission) computer code, version 5.2.2 (Berk et al., 2004; Berk et al., 2006).



262
263 Fig. 3. Channel response functions of the five required channels.

264 In this study, the TIGR (The Institute for Genomic Research,
265 <http://ara.abct.lmd.polytechnique.fr/>) database (Chevallier et al., 1998) was introduced as the
266 MODTRAN inputs to generate simulation dataset of various atmospheric conditions (Chen et
267 al., 2017; Ren et al., 2020; Wang et al., 2019b). After quality check and cloudy-sky removal,
268 946 atmospheric profiles remained which were considered to be redundant because many of
269 them were similar to each other (Fig. 4). Therefore, it is necessary to select different
270 atmospheric profiles to make the total atmospheric water vapor content (WVC) and the bottom
271 layer temperature (T_0) uniformly distributed from dry to moist and from cold to warm,
272 respectively. Following these criteria, a subset with a moderate sample number of 98 was
273 selected and believed to be representative enough for determining the coefficients of the SW-

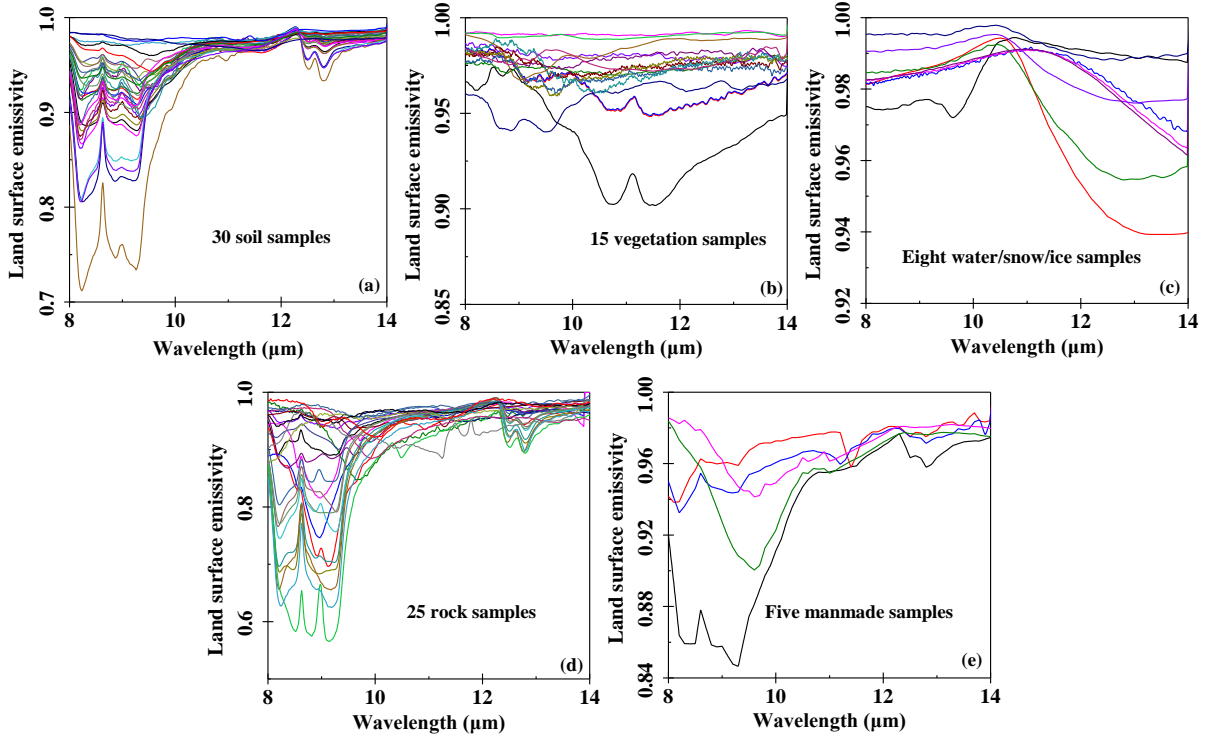
274 like equation. As shown in Fig. 4, the selected 98 atmospheric profiles have varying
275 characteristics with the water vapor content ranging from 0.09 g/cm² to 6.15 g/cm² and the
276 bottom layer temperature (T_0) ranging from 236.25 K to 311.95 K.



277
278 Fig. 4. Scatter plot of bottom layer temperature and total atmospheric water vapor content of the original 946
279 atmospheric profiles (represented by circles) and the selected 98 ones (represented by triangles).

280 The input LST was then assigned to vary with an interval of 5 K from $T_0 - 5$ K to $T_0 + 15$
281 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
282 used in the study of Zheng et al. (2019) was refined to be more representative at the global scale.
283 First, the soil samples with similar characteristics were removed. Second, the vegetation
284 samples were expanded to cover typical dry and fresh vegetations. Third, the typical rock
285 samples were included to cover the three rock types (i.e. igneous, sedimentary, and
286 metamorphic). Forth, the manmade samples were introduced to include typical paving,
287 constructing, and roofing materials. In addition, the number of each kind of emissivity sample
288 was determined referring to the study of Ren et al. (2020). Finally, a new dataset of 83 LSE
289 spectra, containing 30 soil, 15 vegetation, 8 water/snow/ice, 25 rock, and 5 manmade samples
290 were selected from the ECOSTRESS library (<https://speclib.jpl.nasa.gov/>) to represent different

291 Earth surface coverages globally (Fig. 5).



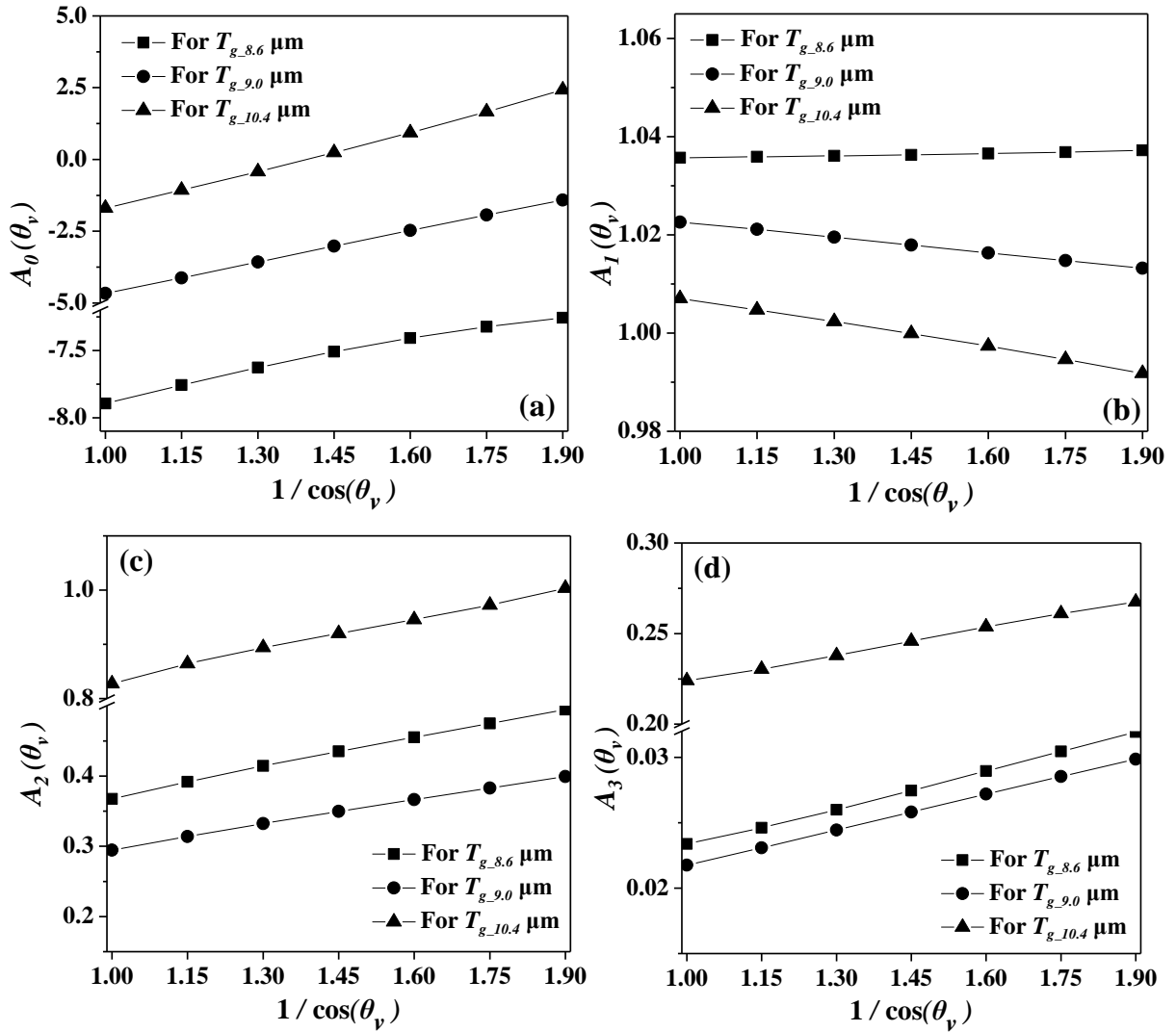
294 Fig. 5. LSE spectra of the (a) 30 soil samples, (b) 15 vegetation samples, (c) 8 water/snow/ice samples, (d)
295 25 rock samples, and (e) 5 man-made samples.

296 Finally, under each of the seven local viewing zenith angles (0.00° , 29.59° , 39.72° , 46.40° ,
297 51.32° , 55.15° , 58.24°), a dataset containing 37,848 cases could be simulated using the
298 parameters described above.

299 3.3 LST retrieval with different local viewing zenith angle

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

305 method for its application to global TIR measurements with different viewing angles.



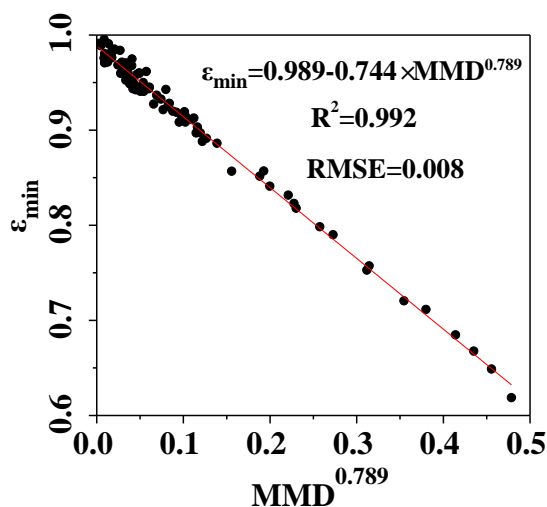
306

307

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

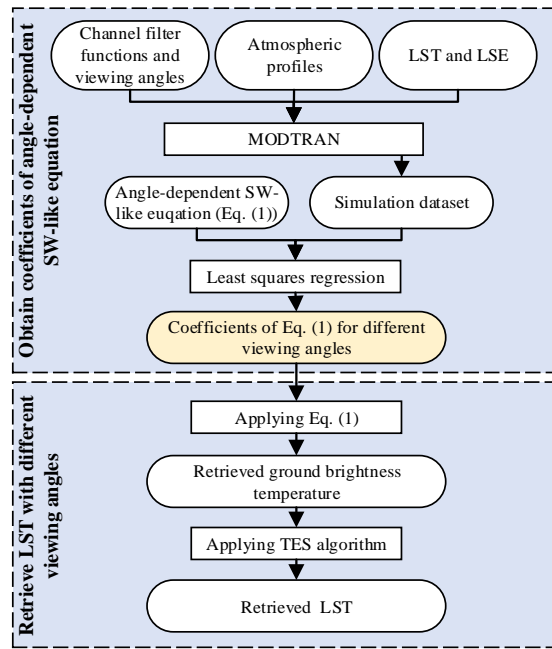
310 Using the coefficients obtained above, the ground brightness temperature of the three
 311 channels centered at 8.6, 9.0, and 10.4 μm could be obtained based on Eq. (1). Then, the TES
 312 method was introduced to separate the LST and LSE. This method combines three modules
 313 (Gillespie et al., 1998): the NEM algorithm, used to provide a first guess of the temperature and
 314 the emissivity spectrum; the Ratio algorithm, used to preserve shape of the emissivity spectrum;
 315 and the MMD (maximum-minimum difference) algorithm, used to find the accurate

316 temperature and emissivity by relating the minimum emissivity (ϵ_{\min}) to the spectra contrast
317 (MMD). Therefore, an empirical relationship between ϵ_{\min} and MMD is very important and
318 required in the TES method (Hulley and Hook, 2011). On the basis of previously selected 83
319 emissivity spectra, this $\epsilon_{\min} \sim MMD$ relationship was adjusted for the spectra response
320 functions of the three relevant channels with an uncertainty of about 0.008 (Fig. 7).



321
322 Fig. 7. The empirical relationship between ϵ_{\min} and MMD , based on the selected 83 emissivity spectra
323 samples.

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



326

327

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

328

Using the simulation dataset above, the uncertainty of the extended SW-TES method could

329

be accessed. Firstly, the simulated satellite observations under each of the seven local viewing

330

zenith angles were separately inputted into the extended SW-TES method. The LST retrievals

331

then were compared with the MODTRAN LST inputs used for generating the simulated dataset

332

under each viewing zenith angle. As shown in Fig. 9a, the root-mean-square error (RMSE) of

333

retrieved LST increased by about 0.4 K from 0.94 K to 1.37 K, as the secant local viewing

334

zenith angle increased from 1.0 to 1.9. When the simulations and the relevant LST retrievals

335

regarding the seven local viewing zenith angles were considered together, the overall LST

336

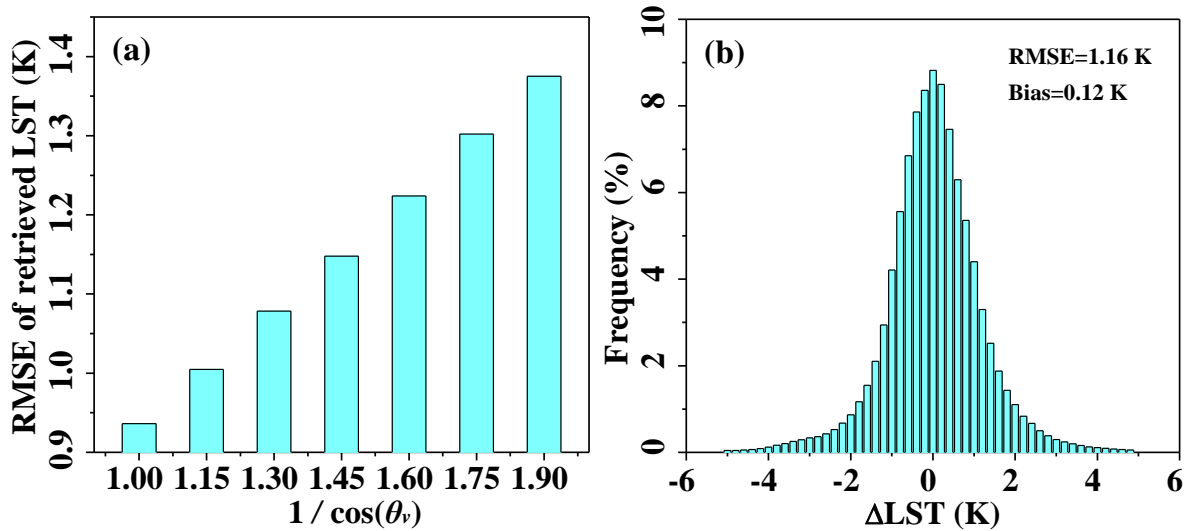
retrieval uncertainty was approximately 1.16 K, as shown in Fig. 9b. Although the increase of

337

local viewing zenith angle affects the performance of the extended SW-TES method, the overall

338

LST retrieval accuracy was still acceptable according to the simulation results.



339
 340 Fig. 9. Land surface temperature (LST) retrieval accuracy on the basis of simulation dataset. (a) RMSE of
 341 retrieved LST as a function of the secant local viewing zenith angle. (b) Residual histograms of retrieved LST
 342 when considering the simulations of the seven viewing zenith angles all together.

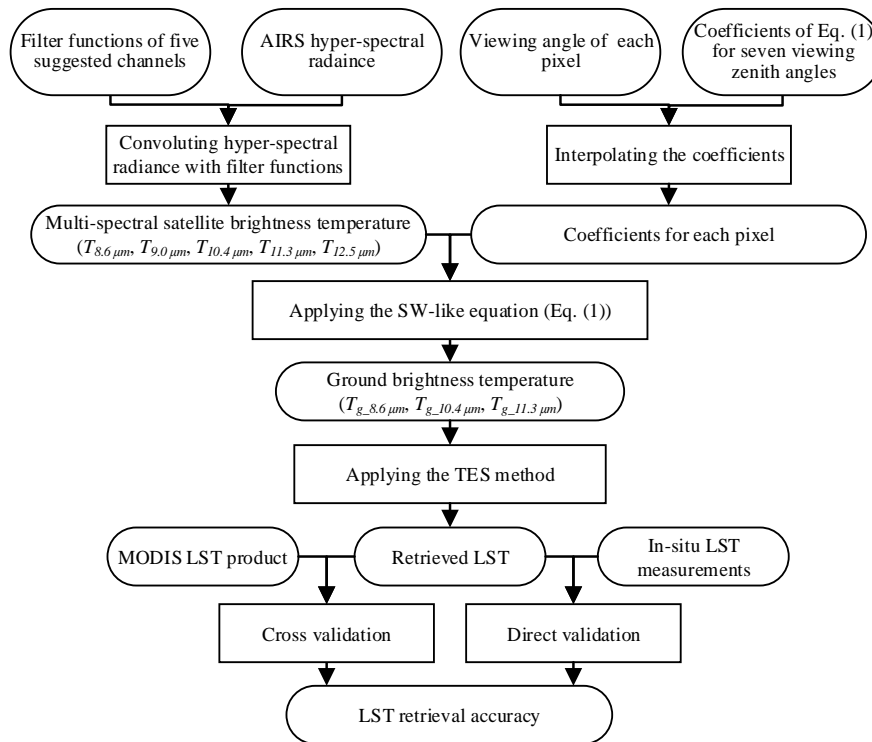
343 4. Application to real satellite data

344 4.1 LST retrieval from real AIRS data

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

352 To apply the angle-dependent SW-like equation [Eq. (1)] to real satellite data, a look-up
 353 table of the coefficients for different viewing angles is required. Because the coefficients in Eq.
 354 (1) varied smoothly with the viewing zenith angle, according to Fig. 6, it was sufficient to

355 establish this look-up table using the coefficients at the seven viewing zenith angles described
 356 in Section 3.3. The coefficients at any other viewing angle could thus be linearly interpolated
 357 according to this look-up table. Then, the ground brightness temperature of each pixel could be
 358 easily obtained. Consequently, LST could be retrieved without knowing any prior knowledge
 359 after applying the TES algorithm. Finally, the LST retrieval accuracy was assessed through
 360 cross-validation with the MODIS LST product and direct validation with the in-situ LST
 361 measurements. The detailed processing flow for applying the extended SW-TES method to real
 362 AIRS data is shown in Fig. 10.



363
 364 Fig. 10. The flow diagram of applying the extended SW-TES method to real AIRS data.

365 4.2 Intercomparison with MYD11 LST product

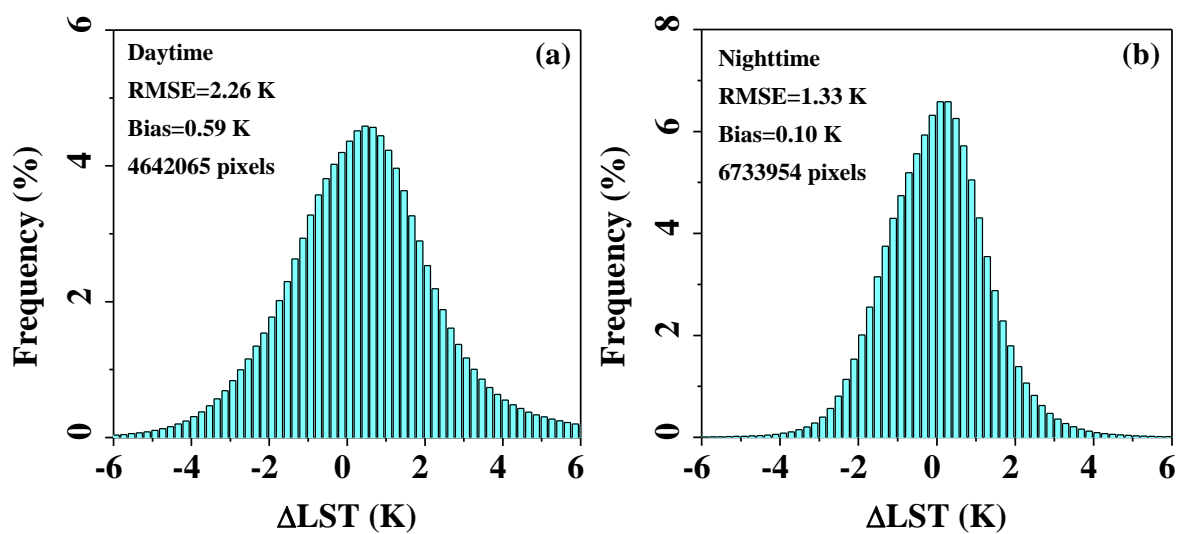
366 4.2.1 Intercomparison residuals for daytime and nighttime observations

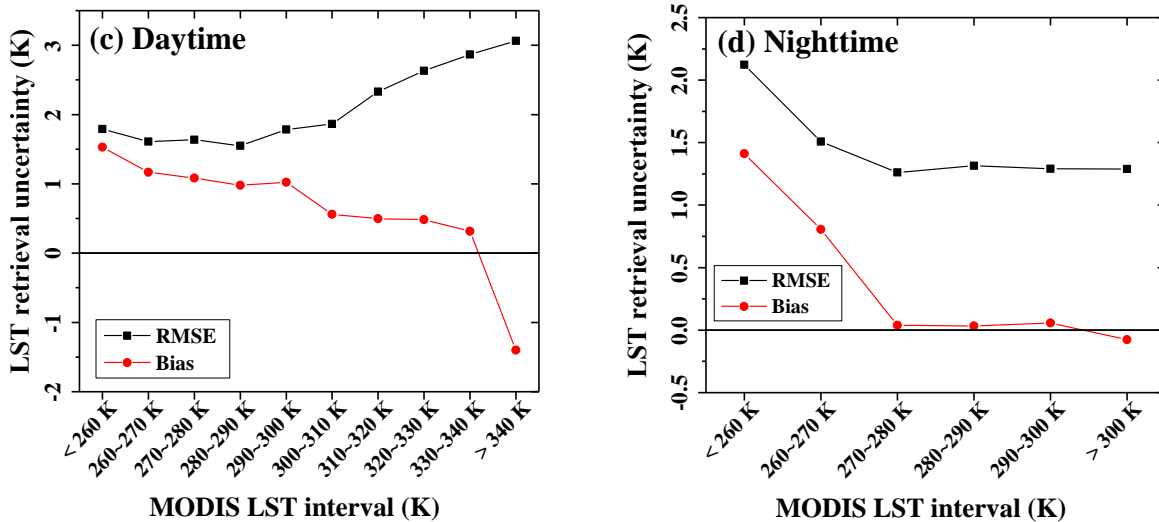
367 Following the processing flow shown in Fig. 10, the extended SW-TES method was

368 implemented to retrieve LST globally from the real AIRS data. Then, the well-validated
369 MODIS MYD11_L2 LST product was considered as a reference to evaluate the global LST
370 retrieval results. Based on the quality control layers distributed along with the MYD11_L2 LST
371 product, only cloudless pixels with announced LST retrieval accuracy which were better than
372 1.5 K were retained. Moreover, considering that the pixel size of MODIS is around 1.0 km,
373 while that of AIRS is 13.5 km, the valid MODIS LST pixels were first aggregated to match the
374 AIRS spatial resolution using the area-weighted pixel aggregation method (Gao et al., 2012;
375 Qian et al., 2013). For example, all the MODIS LST pixels inside or overlapped with the
376 corresponding AIRS pixel footprint were weighted and aggregated according to their
377 overlapping areas. Please note that the pixel size varies along with the view angles for both
378 instruments. Therefore, the actual coordinates of the four vertexes of each AIRS and MODIS
379 pixel were calculated to determine their accurate overlapping areas. In addition, for each AIRS
380 pixel, the corresponding MODIS LST product was aggregated only if all the involved MODIS
381 pixels had valid LST values and the standard deviation was <1.5 K. We applied these two
382 criteria to ensure the representativeness of the aggregated MODIS LST on the AIRS pixel scale
383 and to guarantee the quality and homogeneity of the selected pixel pairs for intercomparison.

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

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





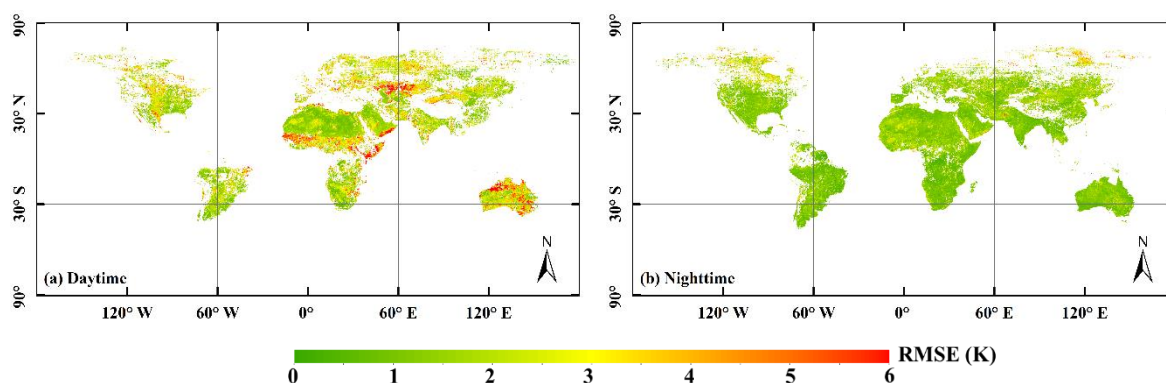
402
403
404

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

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

414 According to the global RMSE distribution map for daytime observations (Fig. 12a), a large
415 difference between the retrieved LST and MODIS LST products was found in some places,
416 such as the regions in the east of the Caspian Sea, south of the Arabian Peninsula, north and
417 southeast of Australia, and central Africa. However, during the nighttime, almost all pixels had
418 relatively smaller LST retrieval RMSE and the magnitude generally did not vary with

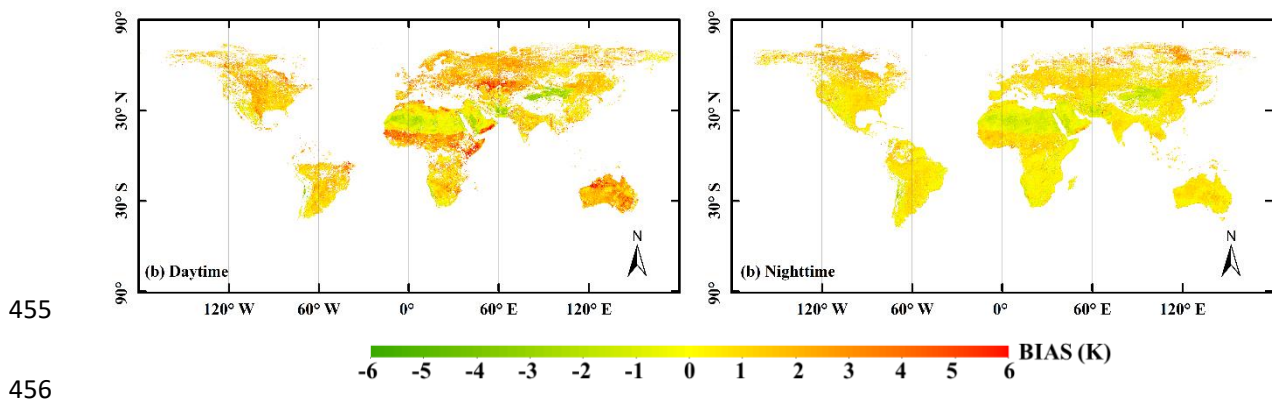
419 geolocation (Fig. 12b). The spatial continuity of the global RMSE distribution map for
 420 nighttime observations was also better than that of daytime observations. According to the
 421 MODIS land cover product for 2018 (Fig. 1), the land surface of the above-mentioned regions
 422 was much more likely dominated by sparse shrublands and grasslands. Therefore, the severe
 423 surface thermal heterogeneity was considered as the most probable reason for the large
 424 discrepancies in these regions during the daytime.



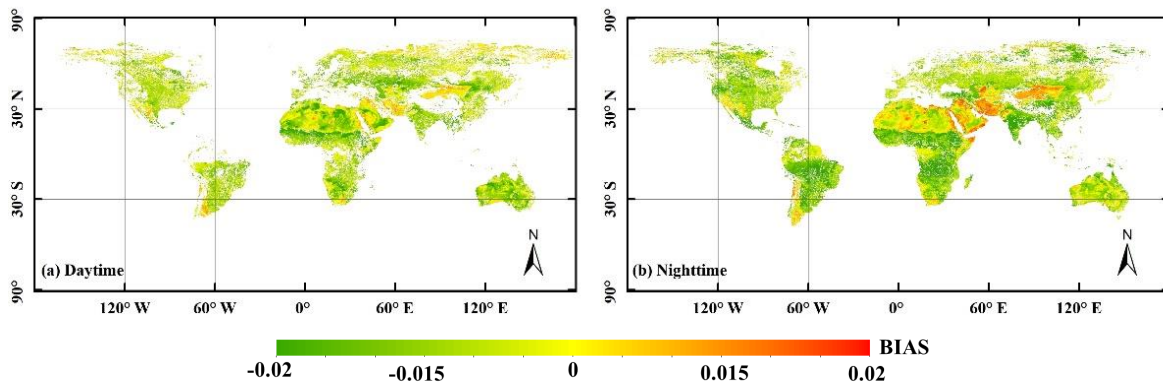
425
 426
 427 Fig. 12. RMSE distribution map of the retrieved land surface temperature (LST) intercomparing with the
 428 aggregated MYD11_L2 LST product for global observations acquired during (a) the daytime and (b) the
 429 nighttime.

430 Fig. 13 shows that the distribution pattern of the global bias coincided with the RMSE
 431 distribution maps. The results showed that the retrieved LST was slightly underestimated
 432 compared with the aggregated MODIS LST product in desert areas (such as the regions of
 433 Sahara, Taklimakan, and Arabian) and overestimated in the rest areas for both daytime and
 434 nighttime observations. To investigate the possible reason that caused the LST underestimation
 435 in these areas, the retrieved LSE was compared with those used in the MODIS MYD11_L2 LST
 436 product. However, the variation of the surface emissivity spectra around $8.6 \mu\text{m}$ is quite large
 437 and the LSE for MODIS channel 29 is not provided by the MYD11_L2 product. In addition,

438 none of the five channels used in this study overlap with MODIS channel 32 when evaluated
 439 using their channel response functions. Therefore, we went on to compare the retrieved LSE for
 440 the 10.4 μm channel and the LSE from MODIS channel 31. We corrected for the potential
 441 impact of differences in the channel response functions by fitting the retrieved LSE for 10.4 μm
 442 channel to that of MODIS channel 31 using linear regression. The adaption coefficients were
 443 obtained using the emissivity dataset used in this study. Finally, we used this data to produce
 444 the bias distribution maps for the retrieved and adapted LSE values for the 10.4 μm channel
 445 using the aggregated MYD11_L2 LSE for MODIS band 31 as a reference (Fig. 14). Results
 446 showed that the retrieved LSE in some regions of the Sahara, Arabian, and the Taklimakan was
 447 a bit larger than the corresponding MODIS LSE product during both daytime and nighttime.
 448 But in other regions, it was generally smaller than the MODIS LSE. This may be one of the
 449 reasons that the retrieved LST was slightly underestimated in some desert areas comparing with
 450 the MODIS MYD11_L2 LST product while overestimated in the rest areas. Moreover, it should
 451 be noticed that the LST retrieval accuracy regarding to bare soil pixels in the C6 MYD11_L2
 452 product was generally lower than that of other land cover types, the absolute uncertainty of
 453 which was validated to be around ± 2.0 K according to previous studies (Duan et al., 2017; Duan
 454 et al., 2018; Wan, 2014). This may also influence the comparison results.



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



460
461
462 Fig. 14. Bias distribution map of the retrieved land surface emissivity (LSE) for the 10.4 μm channel
463 intercomparing with the aggregated MYD11_L2 LSE for MODIS band 31 (the former minus the latter) for
464 global observations acquired during (a) the daytime and (b) the nighttime. Please note that the retrieved LSE
465 for the 10.4 μm channel has been adapted to the response function of MODIS channel 31 by linear regression.

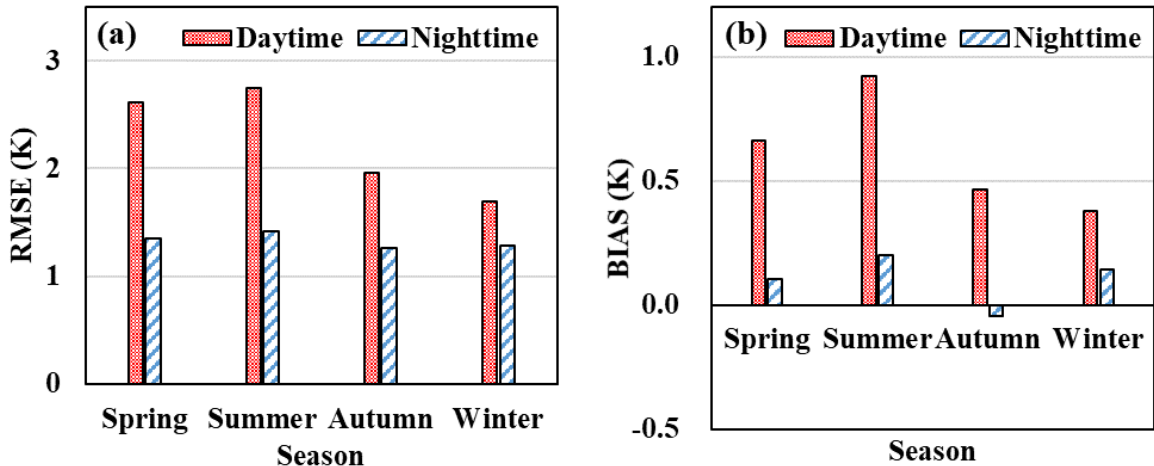
466 These results indicated that the discrepancy between the retrieved LST and MODIS
467 MYD11_L2 LST product was much smaller for nighttime observations than for daytime
468 observations. During the nighttime, because of the absence of solar loading, the LST is
469 considered to be more spatially homogeneous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

473 4.2.2 Intercomparison residuals for different seasons

474 To evaluate the influence of seasons on the LST retrieval results, the discrepancy between

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

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

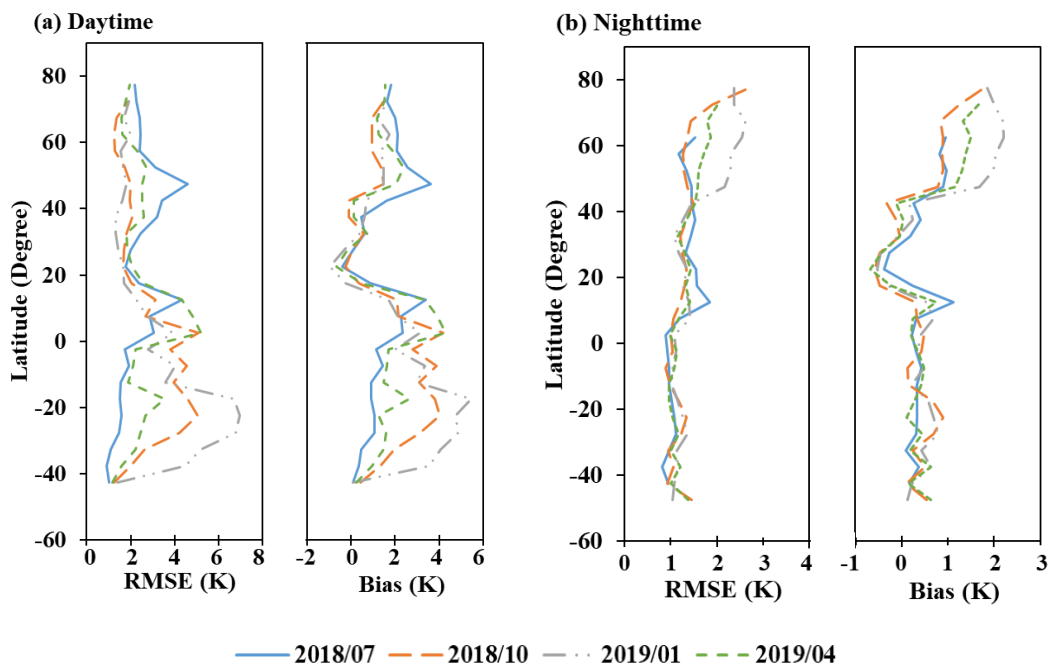


496
 497 Fig. 15. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST) intercomparing with the
 498 MYD11_L2 LST product (the former minus the latter) for different seasons at the global scale. The spring,
 499 summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019,
 500 respectively, for the northern hemisphere, while refer to results in October 2018, January 2019, April 2019,
 501 and July 2018, respectively, for the southern hemisphere. Therefore, the results for these four months were
 502 rearranged accordingly for statistical analysis.

503 In addition, variations in the RMSE and bias with latitude were investigated for different
 504 months (Fig. 16). Statistics were made for a latitude interval of five degrees. During the daytime,
 505 with a decrease in latitude from 80 °N to 20 °N, the RMSE and bias for the observations in July
 506 2018 first increased and then decreased. When the latitude approached the equator, the RMSE
 507 and bias increased again, but gradually decreased as the latitude continually decreased to the
 508 South Pole. As for the observations in January 2019, the varying pattern of RMSE and bias was
 509 reversed as the latitude increased from the South Pole to the North Pole. The results of the
 510 spring and autumn were generally between summer and winter. However, during the nighttime,
 511 the LST retrieval RMSE of the four months increased gradually with increasing latitude from
 512 the South Pole to the North Pole, and no other obvious variations were found. But there was a

513 decrease in the biases with respect the latitude between 20 °N and 40 °N for both daytime and
 514 nighttime observations. This is believed to be relevant to desert regions, as shown in Fig. 13.

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

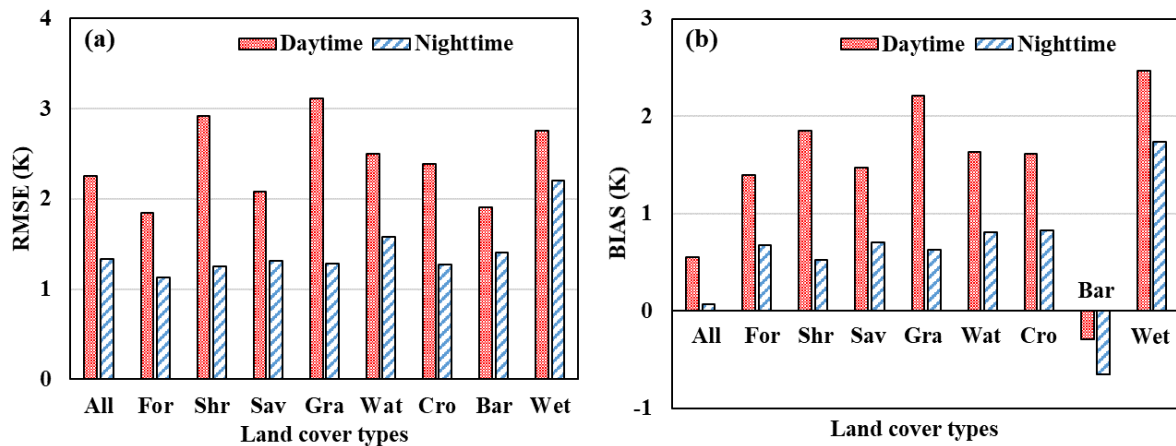


526 Fig. 16. RMSE and bias of the retrieved land surface temperature (LST) intercomparing with the MYD11_L2
 527 LST product as functions of latitude (the former minus the latter) during (a) the daytime, (b) nighttime.

528 **4.2.3 Intercomparison residuals for different land cover types**

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

538 During the daytime, it was shown that the RMSE of LST retrievals was <2.5 K for different
539 land cover types, except for shrublands and grasslands, with a magnitude of approximately 3.0
540 K (Fig. 17a). Compared with the daytime results, the RMSE was <1.4 K for all land cover types
541 during the nighttime, except for the water and wetlands (Fig. 17a). Nevertheless, the LST
542 retrieval RMSE of nighttime observations over the water and wetlands were about 1.6 K and
543 2.2 K respectively, which were still lower than most daytime results. Moreover, as indicated by
544 the statistics of bias, the retrieved LST was underestimated compared with the MODIS
545 MYD11_L2 LST product for barren and overestimated for the remaining land cover types (Fig.
546 17b).

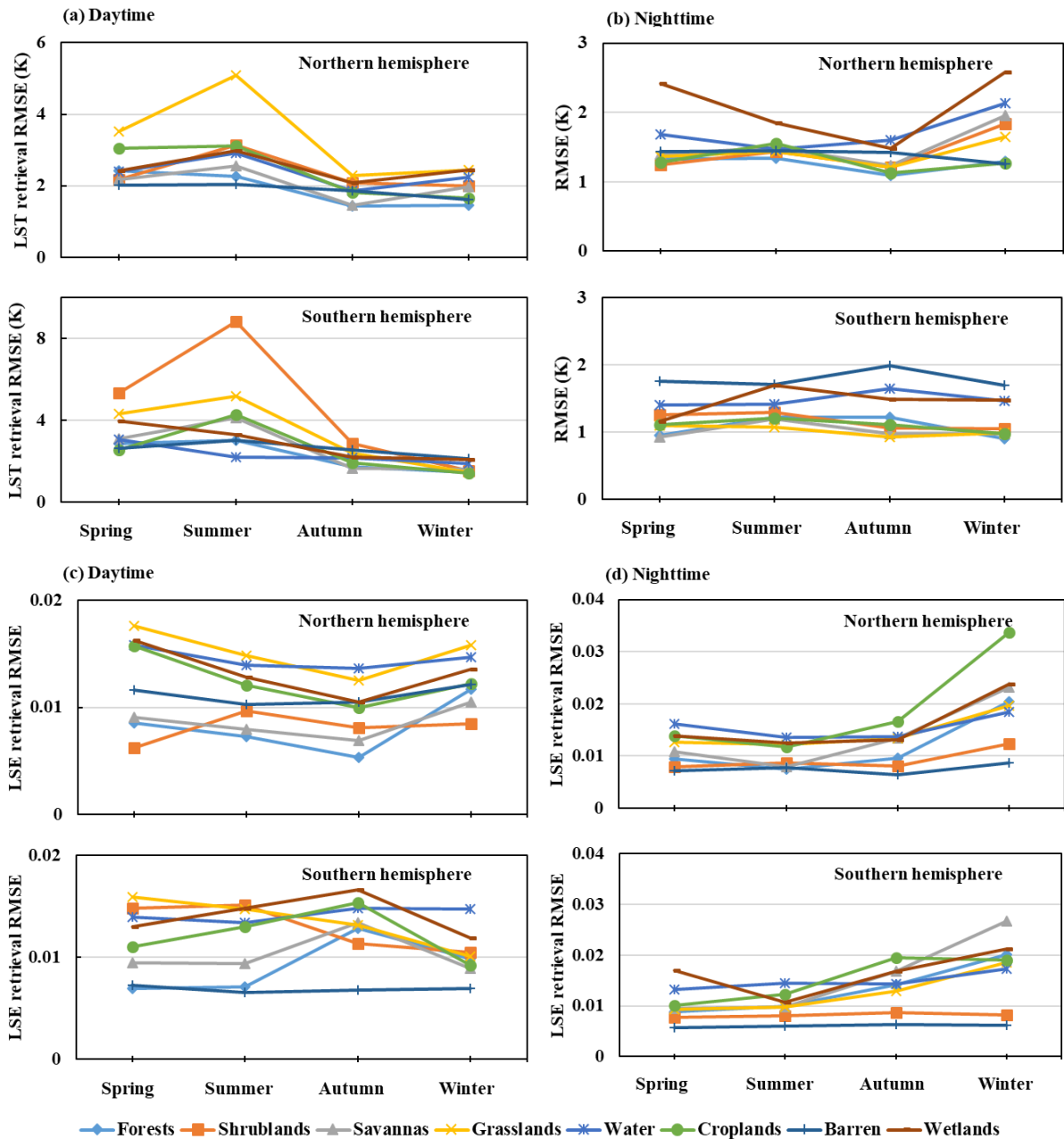


547
 548 Fig. 17. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD11_L2 LST
 549 product (the former minus the latter) for different land cover types. Observations acquired in the four months
 550 were merged to generate the statistics. The meaning of abbreviations follows: All=All land cover types
 551 together, For=Forests, Shr=Shrublands, Sav=Savannas, Gra=Grasslands, Wat=Water, Cro=Croplands,
 552 Bar=Barren, Wet=Wetlands.

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

564 MODIS LST product may also increase. Second, the LSE was retrieved simultaneously with
565 the LST in the extended SW-TES method, and it was assigned according to the land cover types.
566 For heterogeneous land surfaces, the LSE may be generated differently using these two
567 strategies. This may also introduce additional biases in the comparison results. As shown in Fig.
568 18, during the daytime of summer, the RMSE between the retrieved LSE for the 10.4 μm
569 channel and the aggregated LSE for the MODIS band 31 for grasslands is the largest in the
570 northern hemisphere. In the southern hemisphere, the retrieved LSE for the 10.4 μm channel
571 for shrublands, grasslands, and wetlands are generally more different from the aggregated LSE
572 for the MODIS band 31 than other land cover types. However, the LSE retrieval RMSE using
573 the MODIS product as a reference is <0.02 , which generally could not lead to an LST retrieval
574 discrepancy of >4.0 K. Therefore, the third reason may be that the atmospheric correction using
575 the SW-like equation (1) was failed under these conditions.

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



586

587

588

589 Fig. 18. RMSE of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product

590 as a function of season for different land cover types separately for observations during (a) the daytime, and

591 (b) the nighttime; and the RMSE between the retrieved land surface emissivity (LSE) for the 10.4 μ m channel

592 and the aggregated MYD11_L2 LSE for MODIS band 31 as a function of season for different land cover

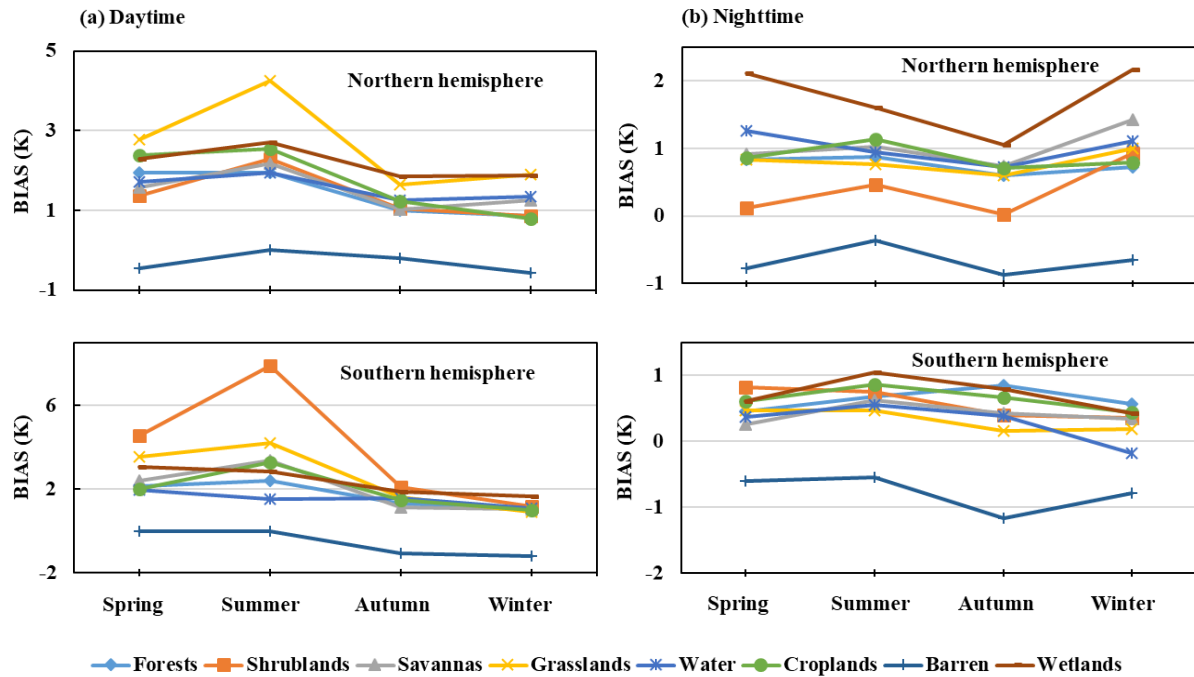
593 types separately for observations during (c) the daytime, and (d) the nighttime. Please note that the retrieved

594 LSE for 10.4 μ m channel has been adapted to the response function of MODIS channel 31 by linear regression.

595 The spring, summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January

596 2019, respectively, for the northern hemisphere, and refer to results in October 2018, January 2019, April
597 2019, and July 2018, respectively, for the southern hemisphere.

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



612
 613
 614 Fig. 19. Bias of retrieved land surface temperature (LST) intercomparing with MYD11_L2 LST product (the
 615 former minus the latter) as a function of season for different land cover types separately. The statistics were
 616 processed for observations acquired in the (a) daytime, and (b) nighttime. The spring, summer, autumn, and
 617 winter refer to results in April 2019, July 2018, October 2018, and January 2019, respectively, for the northern
 618 hemisphere, and refer to results in October 2018, January 2019, April 2019, and July 2018, respectively, for
 619 the southern hemisphere.

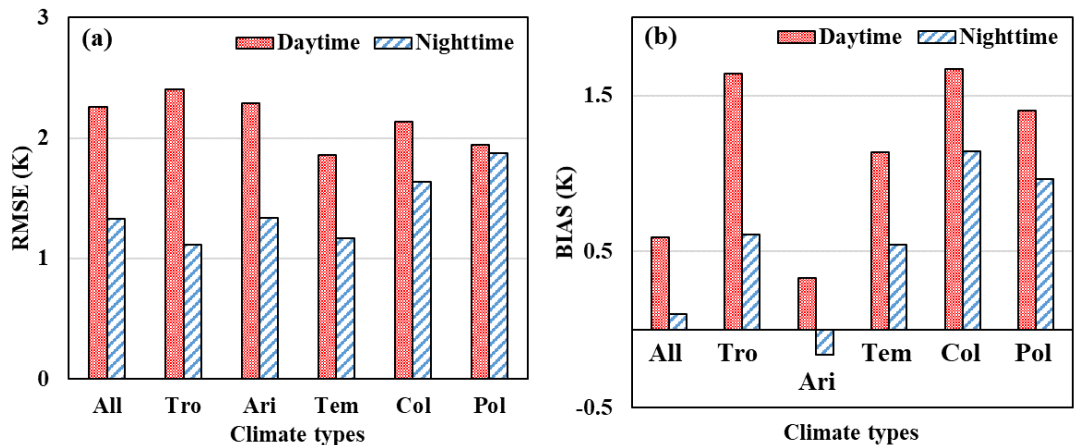
620 Overall, these results suggested that the strong solar loading during the daytime could have
 621 a considerable influence on the comparison results. The relatively smaller RMSE and bias for
 622 nighttime observations than the daytime observations implied that thermal heterogeneity was
 623 likely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

625 4.2.4 Intercomparison residuals for different climate types

626 Climate type may also impact the performance of the extended SW-TES method. Statistics

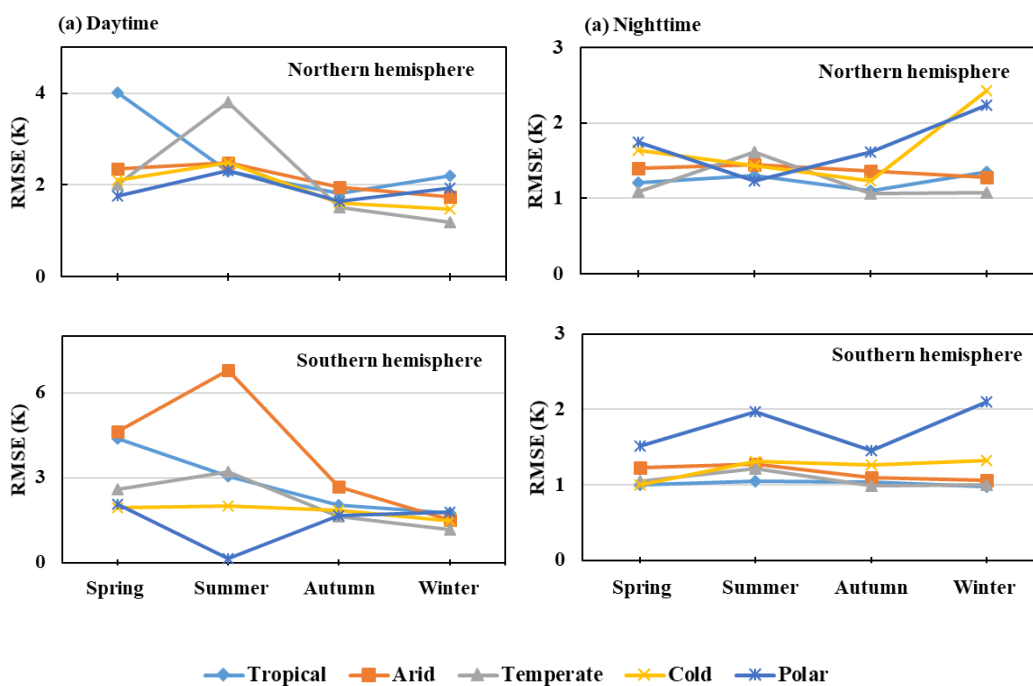
627 of the discrepancy between retrieved LST and MODIS MYD11_L2 LST products were
 628 processed separately for different climate types. In this section, the RMSE and bias of retrieved
 629 LST with MODIS MYD11_L2 LST product as a reference were calculated for each of the five
 630 main climate types classified by the Köppen-Geiger system (Fig. 2 and Table 4).

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



639 Fig. 20. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD11_L2 LST
 640 product (the former minus the latter) for different climate types. Observations acquired in the four months
 641 were merged to generate the statistics. The meaning of abbreviations follows: All=All climate types together,
 642 Tro=Tropical, Ari=Arid, Tem=Temperate, Col=Cold, Pol=Polar.
 643

644 Similar to land cover types, variations in the LST retrieval RMSE for each climate type as
 645 a function of season were also investigated, as shown in Fig. 21. For daytime observations (Fig.
 646 21a), it was observed that the LST retrieval RMSE varied much more considerably during
 647 spring and summer in both the hemispheres. During autumn and winter, the RMSE of the
 648 retrieved LST was almost the same referring to the MODIS MYD11_L2 LST product. This was
 649 perhaps caused by the severe thermal heterogeneity of the land surface during the daytime of
 650 hot seasons. The largest value was observed in spring with respect to the tropical climate in
 651 northern hemisphere (≈ 4.0 K) and the arid climate in southern hemisphere (≈ 6.8 K). For all
 652 nighttime observations (Fig. 21b), the RMSE of the retrieved LST were < 2.0 K, except for cold
 653 and polar climates during the winter. During the night, the highest values were observed in
 654 winter with respect to the cold climate in northern hemisphere (≈ 2.4 K) and the polar climate
 655 in southern hemisphere (≈ 2.1 K). Obviously, the LST retrieval RMSE during nighttime
 656 coincided with the MODIS MYD11_L2 LST product better than that during the daytime.

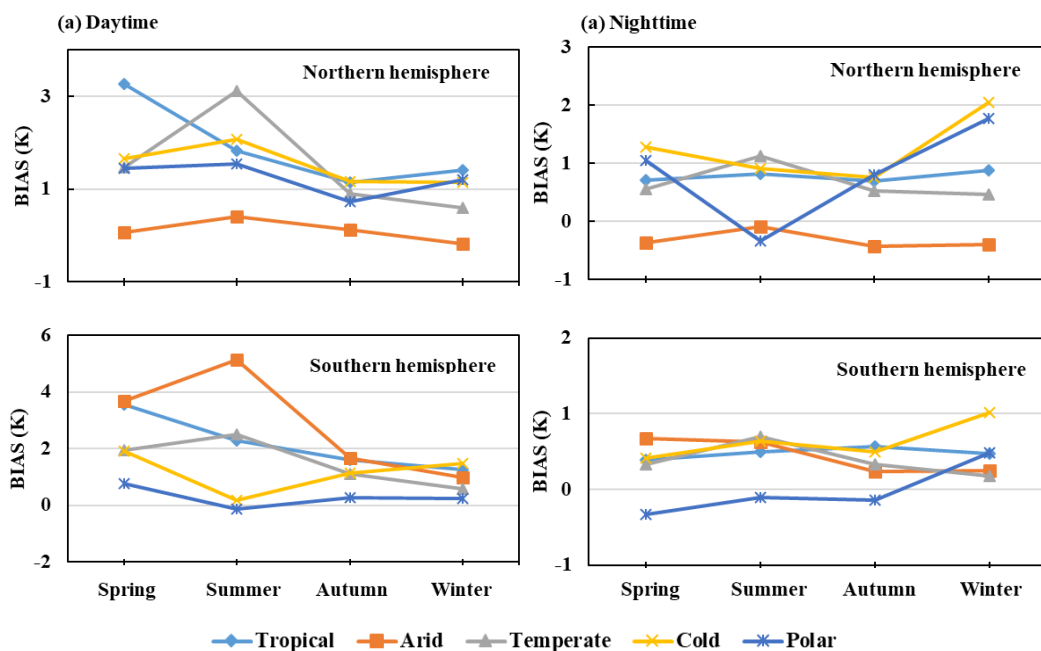


657

658

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

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



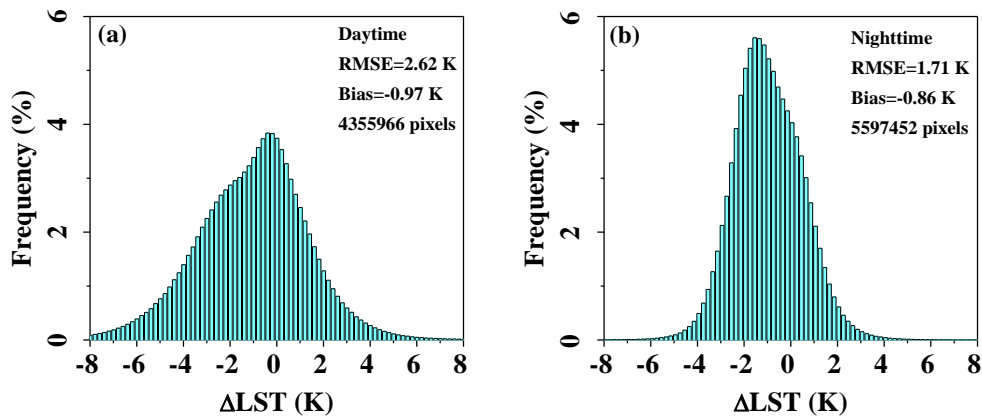
673
 674

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

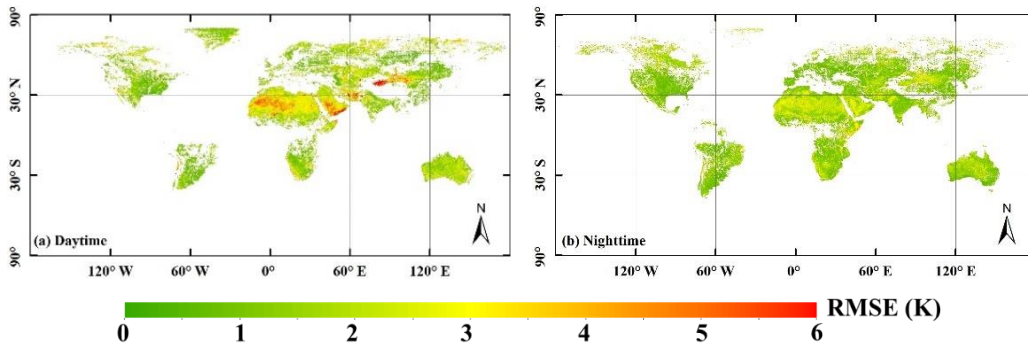
681 **4.3 Intercomparison with MYD21 LST product**

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

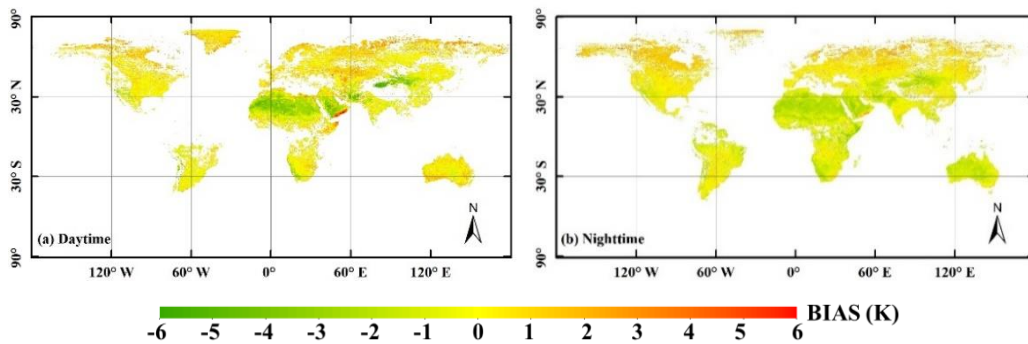
696 to be underestimated (Fig. 25). One reason for this may be that the LSE for these regions was
 697 slightly overestimated (Fig. 26).



698
 699 Fig. 23. Residual histograms of retrieved land surface temperature (LST) intercomparing with MYD21_L2
 700 LST product (the former minus the latter) for global observations acquired during (a) the daytime and (b) the
 701 nighttime.

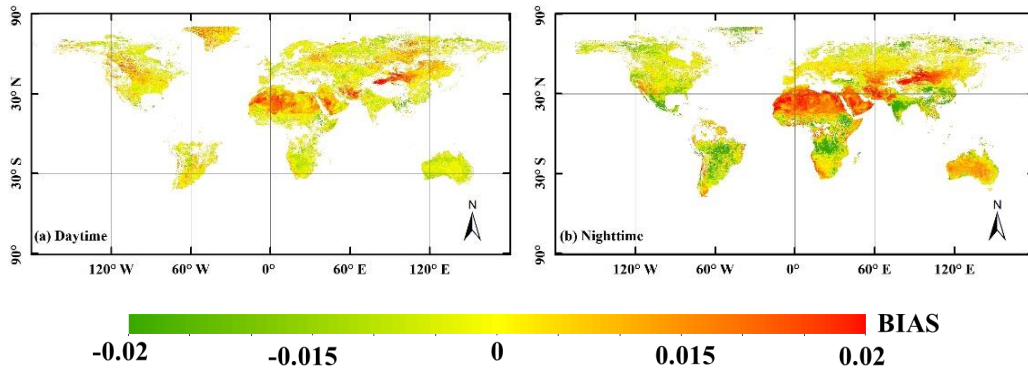


702
 703
 704 Fig. 24. RMSE distribution map of the retrieved land surface temperature (LST) intercomparing with the
 705 aggregated MYD21_L2 LST product for global observations acquired during (a) the daytime and (b) the
 706 nighttime.



707
 708

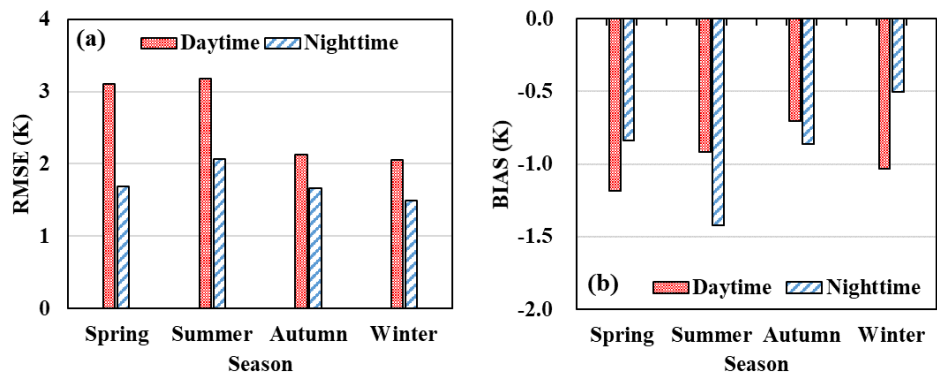
709 Fig. 25. Bias distribution map of the retrieved land surface temperature (LST) intercomparing with the
 710 aggregated MYD21_L2 LST product (the former minus the latter) for global observations acquired during
 711 (a) the daytime and (b) the nighttime.



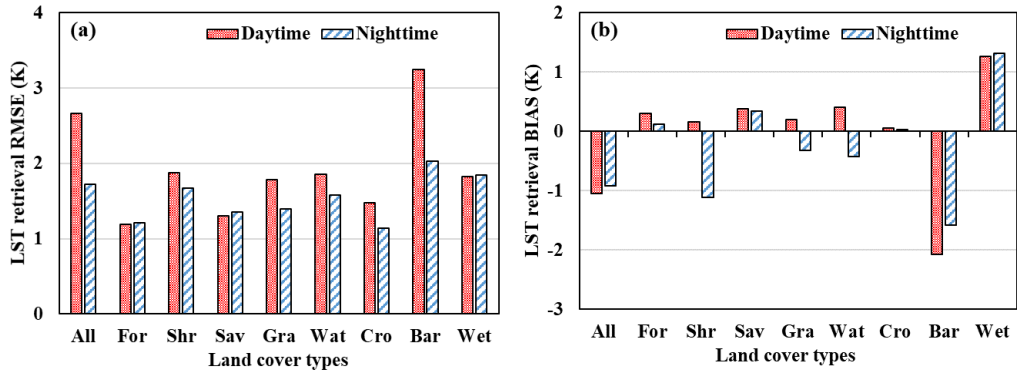
714 Fig. 26. Bias distribution map of the retrieved land surface emissivity (LSE) for the 10.4 μm channel
 715 intercomparing with the aggregated MYD21_L2 LSE for MODIS band 31 (the former minus the latter) for
 716 global observations acquired during (a) the daytime and (b) the nighttime. Please note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

718 The RMSE for the retrieved LST for each season was almost the same as those calculated
 719 using the MYD11_L2 LST product as a reference (Fig. 27a). However, all the biases were
 720 negative (Fig. 27b), suggesting that the LST retrievals were all smaller than the MYD21_L2
 721 LST products for each of the different seasons. Fig. 28 showed that the largest discrepancy
 722 appears in the land cover type values for barren data with an RMSE of >3.0 K and a bias of $<-$
 723 2.0 K. The other land cover types had an RMSE of <2.0 K using the MYD21_L2 LST as a
 724 reference (Fig. 28) while the RMSE for the LSE retrievals in the 10.4 μm channel when
 725 evaluated against the aggregated MYD21_L2 LSE for MODIS band 31 was <0.02 (Fig. 29).
 726 Additionally, for those regions with an “Arid” climate type, the RMSE and bias values for the
 727 retrieved LST were >2.5 K (Fig. 30a) and <-1.0 K (Fig. 30b) respectively, which was

728 significantly poorer than the results for the other climate types. In fact, barren land cover was
 729 generally shown to overlap with the arid climate type. Therefore, the results shown in Fig. 30
 730 agreed with those of Fig. 28. Moreover, a comparison of the LST retrieval biases when using
 731 the MYD21_L2 LST as a reference with those results shown in section 4.2 once again showed
 732 that the retrieved LST values fell between these two MODIS LST products under most
 733 conditions.

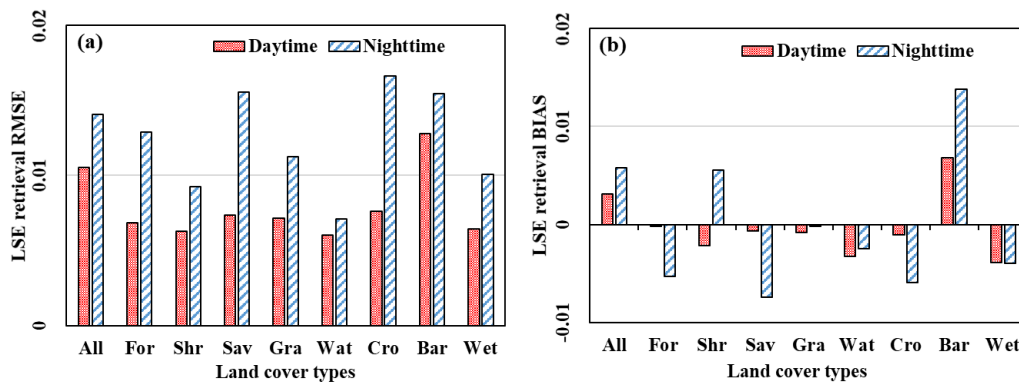


734
 735 Fig. 27. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST) intercomparing with the
 736 MYD21_L2 LST product (the former minus the latter) for different seasons at the global scale. The spring,
 737 summer, autumn, and winter refer to results in April 2019, July 2018, October 2018, and January 2019,
 738 respectively, for the northern hemisphere, while refer to results in October 2018, January 2019, April 2019,
 739 and July 2018, respectively, for the southern hemisphere. Therefore, the results for these four months were
 740 rearranged accordingly for statistical analysis.

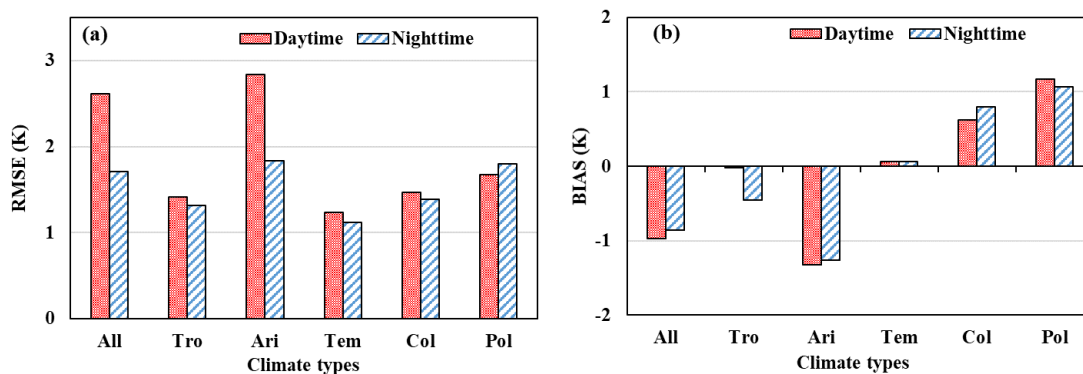


741
 45

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



747
 748 Fig. 29. (a) RMSE and (b) Bias between retrieved land surface emissivity (LSE) for the 10.4 μm channel and
 749 the aggregated MYD21_L2 LSE for MODIS band 31 (the former minus the latter) for different land cover
 750 types. Observations acquired in the four months were merged to generate the statistics. The meaning of
 751 abbreviations follows: All=All land cover types together, For=Forests, Shr=Shrublands, Sav=Savannas,
 752 Gra=Grasslands, Wat=Water, Cro=Croplands, Bar=Barren, Wet=Wetlands.



753
 754 Fig. 30. (a) RMSE and (b) Bias between the retrieved land surface temperature (LST) and MYD21_L2 LST
 755 product (the former minus the latter) for different climate types. Observations acquired in the four months

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

758 **4.4 Validation with in-situ LST measurements**

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

$$764 \quad LST = \left[\frac{R_{\uparrow} - (1 - \varepsilon_b)R_{\downarrow}}{\sigma\varepsilon_b} \right]^{\frac{1}{4}} \quad (2)$$

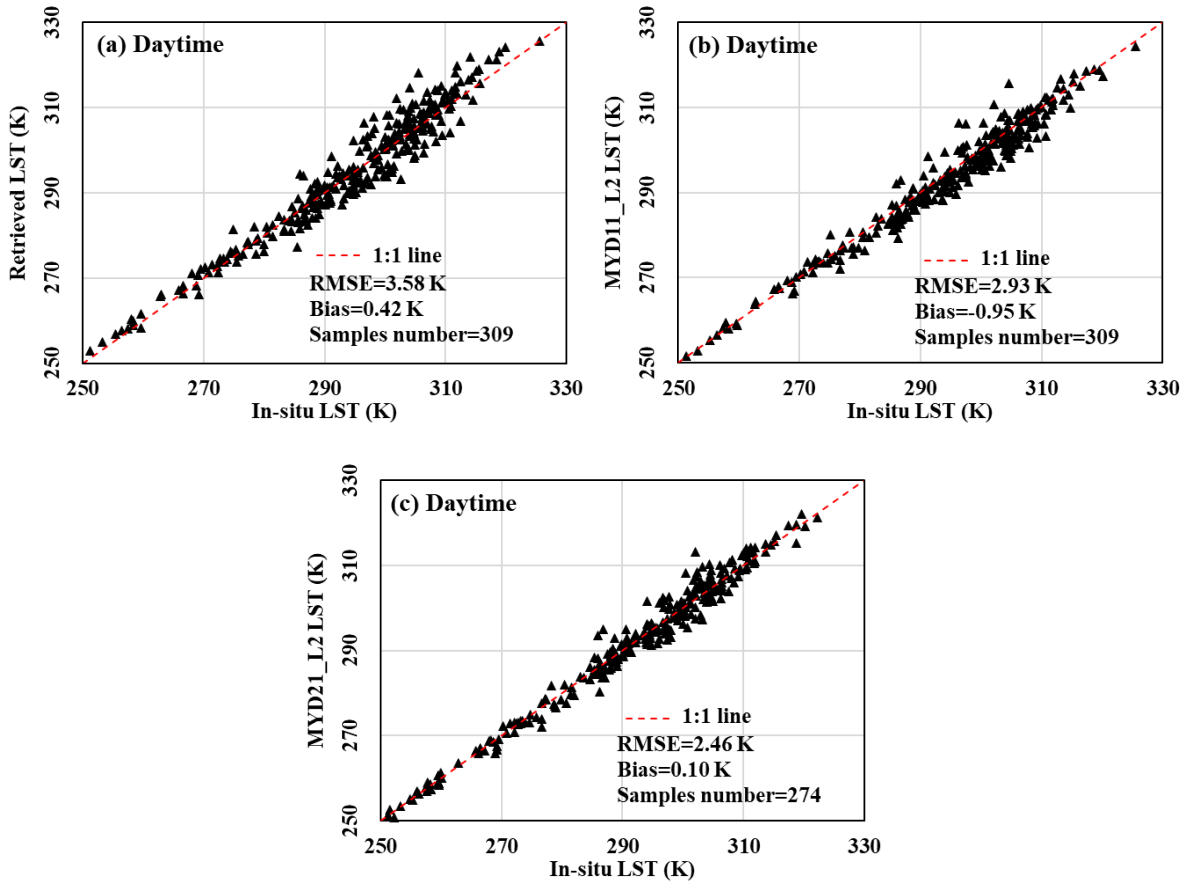
765 where R_{\uparrow} and R_{\downarrow} are respectively the upwelling and downwelling infrared radiation (W/m^2),
766 ε_b is the surface broadband emissivity, σ is the Stephan-Boltzmann constant (5.67×10^{-8}
767 $\text{W}/\text{m}^2/\text{K}^4$).

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

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

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

788 Taking the in-situ LST measurements from the seven SURFRAD sites in the year 2018 and
789 2019 as ground truth, the validation results showed that the LST could be retrieved with an
790 RMSE of 3.58 K during the daytime (Fig. 31a). As a comparison, the uncertainties of MODIS
791 MYD11_L2 and MYD21_L2 LST products were also evaluated and found to have the RMSE
792 of 2.93 K (Fig. 31b) and 2.46 K (Fig. 31c), respectively. The bias of retrieved LST was around
793 0.42 K , which was larger than that of the MODIS MYD11_L2 LST product with the bias of
794 about -0.95 K (Fig. 31b). Therefore, the retrieved LST was generally larger than the MODIS
795 MYD11_L2 LST product, as described in the previous sections.



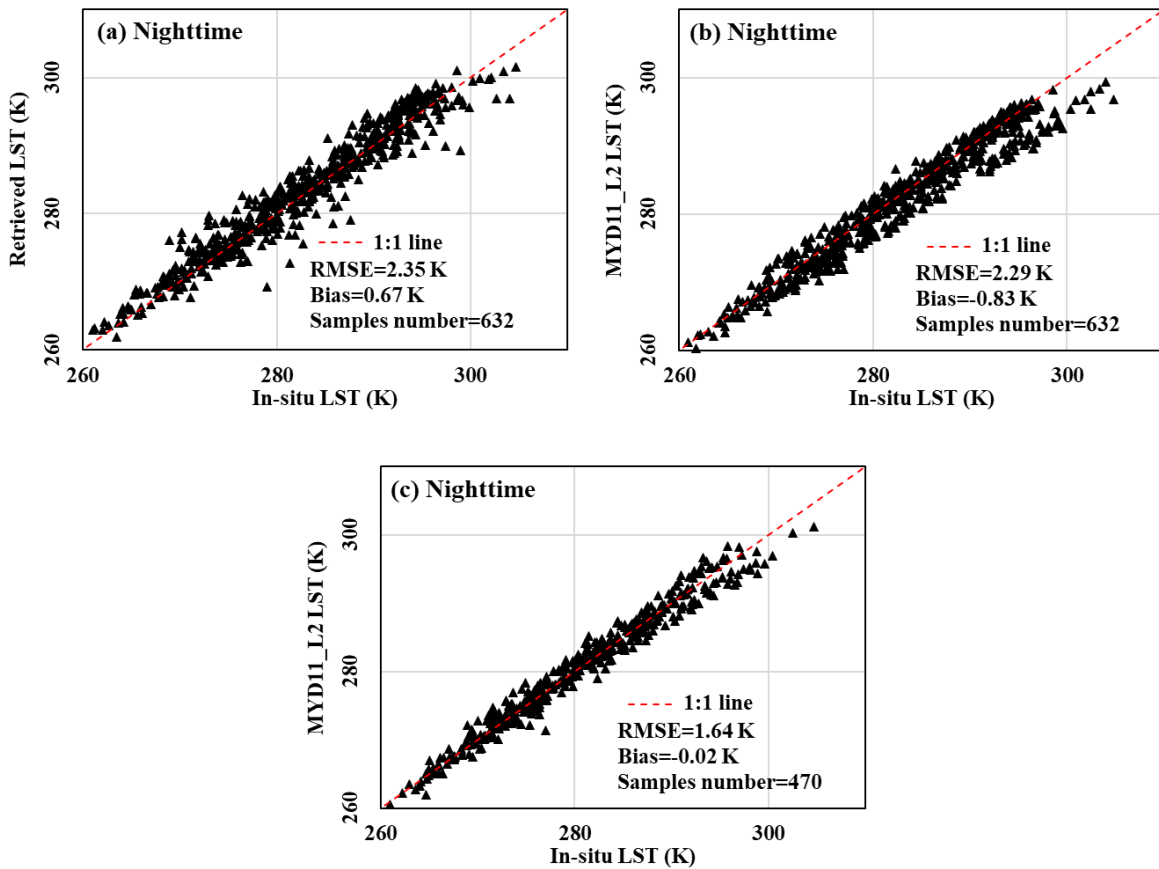
796

797

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

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

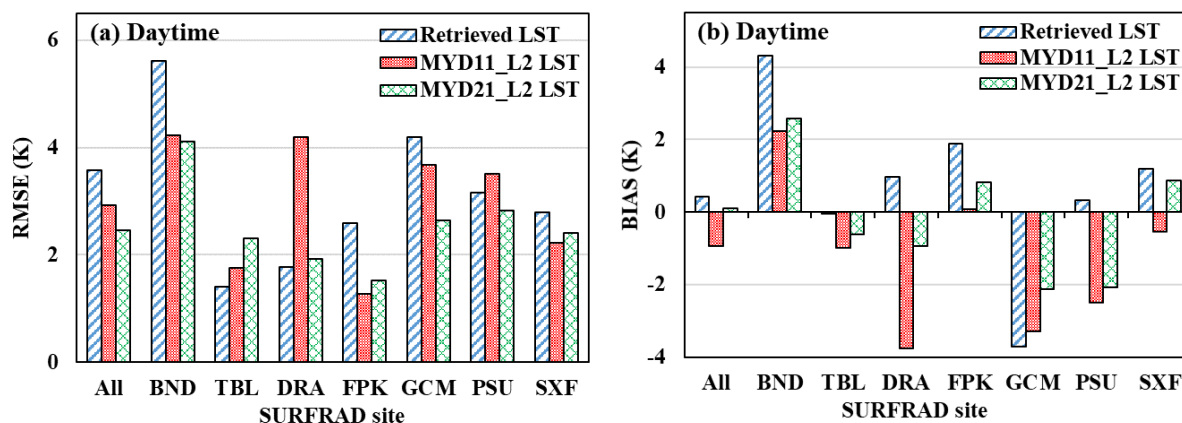
809 larger than the MODIS MYD11_L2 LST product under most conditions.



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

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

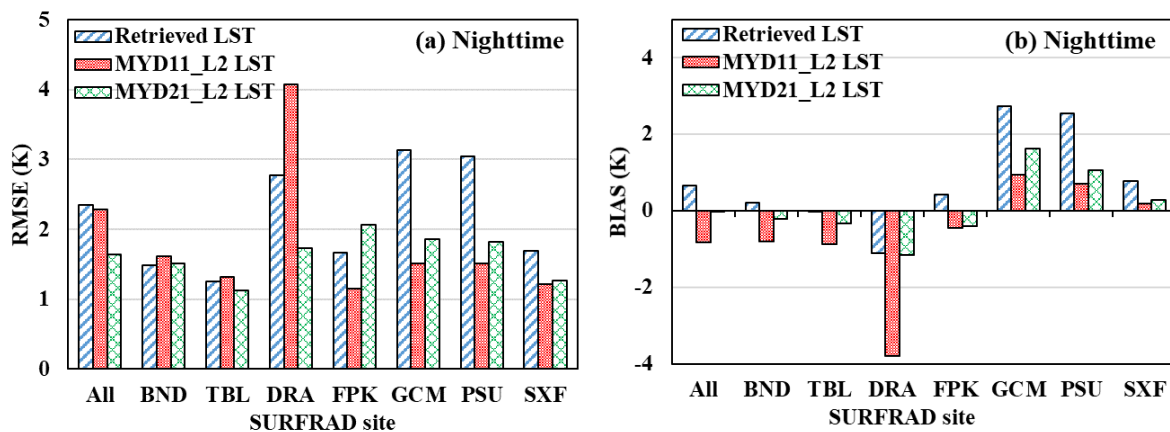
822 Table Mountain site (grass) and Fort Peck (grass) sites produced an RMSE of <2.0 K. For the
 823 MYD21_L2 LST product, the RMSEs at the Desert Rock (open shrub) and Fort Peck (grass)
 824 sites were <2.0 K. All the observations at the other SURFRAD sites (dominated by cropland)
 825 showed that both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



830
 831 Fig. 33. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST), the MYD11_L2 LST product,
 832 and the MYD21_L2 LST product taking in-situ LST from the seven SURFRAD sites as references (retrieved
 833 LST, MYD11_LST, and MYD21_LST minus the in-situ LST). Observations acquired in the daytime of year
 834 2018 and 2019 were merged to generate the statistics for each site separately. The meaning of abbreviations
 835 follows: All=All sites together, BND= Bondville, TBL= Table Mountain, DRA= Desert Rock, FPK= Fort
 836 Peck, GCM= Goodwin Creek, PSU= Penn State, SXF=Sioux Falls.

837 During the nighttime, the LST could be retrieved with an RMSE of <2.0 K except at the
 838 Desert Rock (open shrub), Goodwin Creek (cropland/natural vegetation), and the Penn State
 839 (cropland/natural vegetation) sites (Fig. 34a). But for the MYD11_L2 LST product, the RMSE

840 was over 2.0 K only at the Desert Rock (open shrub) site. For the MYD21_L2 LST product, the
 841 RMSEs at all sites were <2.0 K except the Fort Peck site (grass) which has the RMSE of 2.07
 842 K. Results also showed that the MYD11_L2 LST was greatly underestimated at the Desert Rock
 843 site (with the bias recorded at around -4.0 K), while the extended SW-TES method
 844 overestimated the retrieved LST at the Goodwin Creek (cropland/natural vegetation), and the
 845 Penn State (cropland/natural vegetation) sites with the bias being recorded at around 2.5 K (Fig.
 846 34b).



847 Fig. 34. (a) RMSE and (b) Bias of the retrieved land surface temperature (LST), the MYD11_L2 LST product,
 848 and the MYD21_L2 LST product taking in-situ LST from the seven SURFRAD sites as references (retrieved
 849 LST, MYD11_LST, and MYD21_LST minus the in-situ LST). Observations acquired in the nighttime of year
 850 2018 and 2019 were merged to generate the statistics for each site separately. The meaning of abbreviations
 851 follows: All=All sites together, BND= Bondville, TBL= Table Mountain, DRA= Desert Rock, FPK= Fort
 852 Peck, GCM= Goodwin Creek, PSU= Penn State, SXF=Sioux Falls.

854 5. Discussions

855 5.1 Uncertainties in the cross-validation results when referencing the MODIS LST 856 products

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

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

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

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

900 **5.2 Issues associated with the T-based validation method using ground measurements**
901 **from SURFRAD sites**

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

913 First, the accuracy of the in-situ LST values derived from the pyrgeometer readings may
914 be limited. For example, Guillevic et al. (2012) indicated that the instrumental uncertainty alone
915 leads to an error of about 1.0 K in the derived LSTs from the SURFRAD sites. In addition,
916 when the derived LST for the grassland sites were compared with those from the TIR
917 radiometers, the authors noted an average difference in readings of up to 2.0 K (Krishnan et al.,
918 2020). Gerace et al. (2020) also announced that the standard deviation in the differences
919 between the LST derived data from the SURFRAD sites and those retrieved from the
920 spaceborne TIR instruments was around 2.0 K. Since this 2.0 K residual error consistently exists,
921 the SURFRAD measurements are likely to be particularly challenging in T-based validations

922 requiring high degrees of accuracy. This error is likely the result of reflected solar radiation as
923 the pyrgeometers are sensitive to radiation from 3 to 50 μm (Hulley et al., 2021).

924 Second, the differences in the spatial representativeness of the ground sites and satellite
925 pixels are also a challenge for T-based validation using SURFRAD measurements especially
926 during the daytime. Previous study has demonstrated that the shadow issues, phenological
927 changes in the crop/grassland sites, and differences in the surface geology surrounding the
928 Desert rock site lead to more heterogeneous surfaces (Guillevic et al., 2014; Malakar et al.,
929 2018; Wang and Liang, 2009), making the measurements from the SURFRAD sites less reliable
930 at the satellite pixel scale. Malakar et al. (2018) showed that the Sioux Falls, Penn State, and
931 Goodwin Greek sites are more suitable for LST validation when using sensors with a spatial
932 resolution of >100 m than the other four SURFRAD sites.

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

944 above were also included in the comparisons. In addition, the uncertainty in our results is
945 comparable with that of the MODIS LST products evaluated using the same in-situ
946 measurements as their reference.

947 **6. Conclusions**

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

960 Taking the MODIS MYD11_L2 LST product as the reference, we found that the LST could
961 be retrieved with RMSE <2.5 K for global daytime observations using the extended SW-TES
962 method. For nighttime observations, the RMSE decreased to <1.5 K because of the absence of
963 solar loading. The generally positive biases indicated that the retrieved LST was overestimated
964 compared with the MODIS MYD11_L2 LST product. It should also be noted that the

965 uncertainties of the MYD11_L2 LST product are approximately 1.0 K (Duan et al., 2018; Duan
966 et al., 2019a; Wan, 2014), which may have also contributed to the discrepancies between the
967 retrieved LST and MYD11_L2 LST product. Therefore, it can be concluded that the extended
968 SW-TES method could be used to retrieve the global LST from the suggested multichannel
969 satellite observations with acceptable accuracy.

970 Separate intercomparison results by season showed that for both hemispheres, the largest
971 RMSE of retrieved LST during daytime occurred in summer, followed by spring, autumn, and
972 winter. The RMSE exceeded 4.0 K, while the bias exceeded 3.0 K based on the results of
973 summer and spring in the southern hemisphere. In contrast, LST retrievals coincided with the
974 MYD11_L2 LST product much better in the northern hemisphere with RMSE <2.5 K and bias
975 <1.0 K. For nighttime observations, no obvious relations were found between the LST retrieval
976 RMSE and the seasons. In addition, the RMSE was shown to be <1.5 K in both the hemispheres
977 while the bias was <0.7 K, thus indicating that the discrepancy between the retrieved LST and
978 MYD11_L2 LST product was much smaller than the corresponding daytime results. Moreover,
979 the variations in RMSE and bias as a function of latitude for daytime observations showed that
980 the LST retrievals during the local cold seasons generally coincided with the MYD11_L2 LST
981 product better than during the local hot seasons. However, for nighttime observations, the
982 results for each latitude during different seasons were almost the same and increased gradually
983 with increasing latitude from the South Pole to the North Pole.

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

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

997 As for the residual statistics made for different climate types, results showed that the RMSE
998 of retrieved LST decreased gradually as the climate transited from “tropical” to “polar” for
999 daytime observations. However, for nighttime observations, the variation pattern reversed, that
1000 is, the LST retrieval RMSE increased gradually as the climate transited from “tropical” to
1001 “polar.” Moreover, attention should be paid to daytime observations regarding tropical, arid,
1002 and temperate climates during spring and summer. The magnitude of RMSE under these
1003 circumstances could reach 4.0 K in the northern hemisphere and 6.8 K in the southern
1004 hemisphere. A possible reason for this may be that the thermal surroundings of the land surface
1005 would be more heterogeneous.

1006 In addition to the comparison results with the MYD11_L2 LST product, the retrieved LST

1007 was also compared with the MYD21_L2 LST product. Results showed that the LST retrieval
1008 RMSE was larger by about 0.4 K than that when using MYD11_L2 LST product as a reference
1009 for both daytime and nighttime observations at the global scale. Moreover, the biases showed
1010 that the retrieved LST was generally underestimated with magnitude of about -1.0 K at the
1011 global scale when taking the MYD21_L2 LST product as a reference, while was overestimated
1012 when taking the MYD11_L2 LST product as a reference. This indicated that the LST retrieval
1013 results using the extended SW-TES method fell in between these two MODIS LST products for
1014 most circumstances globally.

1015 In addition to the cross-validation using the MODIS LST product, the extended SW-TES
1016 method was validated based on in-situ LST measurements. It was demonstrated that the LST
1017 could be retrieved with an RMSE of 3.58 K for daytime observations and 2.35 K for nighttime
1018 observations. As a comparison, the RMSE of the MODIS MYD11_L2 LST product comparing
1019 with the same in-situ LST measurements was around 2.93 K for daytime observations and 2.29
1020 K during the nighttime. Additionally, the RMSE of the MODIS MYD21_L2 LST product
1021 comparing with the in-situ LST was around 2.46 K for daytime observations and 1.64 K during
1022 the nighttime. Therefore, it was concluded that the extended SW-TES method could be used to
1023 retrieve LST with an accuracy close to that of the MODIS LST product. Compared with
1024 previous studies, the extended SW-TES method had the advantages of the method proposed by
1025 Zheng et al. (2019), which does not require any prior knowledge of the atmosphere and LSE
1026 during LST retrieval. In addition, this method could be used to retrieve LST from global TIR
1027 measurements at different viewing angles.

1028 Nevertheless, additional efforts are required to further improve LST retrieval accuracy. For
1029 example, because the biases of retrieved LST were found to vary with the land cover and climate
1030 types, and were different for daytime and nighttime observations, obtaining the coefficients of
1031 the SW-like equation separately according to the land cover type, climate type, and
1032 day/nighttime may improve the performance of the proposed method. Moreover, the
1033 coefficients of the SW-like equation can also be parameterized as a function of the atmospheric
1034 water vapor content to increase the ground brightness temperature retrieval accuracy under
1035 humid atmospheric conditions (Niclos et al., 2011; Pérez-Planells et al., 2021). Detailed
1036 investigations are planned in one of our succeeding works. Besides, the angle effect of surface
1037 emissivity maybe significant for TIR measurements with large viewing angles. Therefore, the
1038 empirical relationship between minimum emissivity and emissivity contrast could also be
1039 adapted for large viewing angles to improve the LST and LSE separation accuracy.

1040 **Acknowledgements**

1041 This work was supported by the National Natural Science Foundation of China (Grant No.
1042 41921001) and also by the China Scholarship Council (CSC). Thanks would be given to Dr.
1043 Xiaoyue GONG for her help in the English language editing and in the collection of satellite
1044 data. Thanks would also be given to the anonymous reviewers for their valuable comments,
1045 which have been very useful in improving the manuscript quality.

1046 **References**

1047 Anderson, M.C., Norman, J.M., Kustas, W.P., Houborg, R., Starks, P.J., Agam, N., 2008. A thermal-based
1048 remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from
1049 field to regional scales. *Remote Sens. Environ.* 112, 4227-4241.
1050 Augustine, J.A., DeLuisi, J.J., Long, C.N., 2000. Surfrad—a national surface radiation budget network for

1051 atmospheric research. *Bull. Am. Meteorol. Soc.* 81, 2341-2357.

1052 Augustine, J.A., Hodges, G.B., Cornwall, C.R., Michalsky, J.J., Medina, C.I., 2005. An update on surfrad—
 1053 the gcos surface radiation budget network for the continental united states. *J. Atmos. Ocean. Technol.*
 1054 22, 1460-1472.

1055 Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and
 1056 future koppen-geiger climate classification maps at 1-km resolution. *Sci. Data.* 5, 1-12.

1057 Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2020. Publisher
 1058 correction: Present and future koppen-geiger climate classification maps at 1-km resolution. *Sci. Data.*
 1059 7, 1-2.

1060 Becker, F., 1987. The impact of spectral emissivity on the measurement of land surface temperature from a
 1061 satellite. *Int. J. Remote Sens.* 8, 1509-1522.

1062 Becker, F., Li, Z.-L., 1990. Towards a local split window method over land surfaces. *Int. J. Remote Sens.* 11,
 1063 369-393.

1064 Berk, A., Cooley, T.W., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-
 1065 Golden, S.M., Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Lewis, P.E., 2004.
 1066 Modtran5: A reformulated atmospheric band model with auxiliary species and practical multiple
 1067 scattering options. In, *Proc. SPIE 5571, Remote Sensing of Clouds and the Atmosphere IX* (pp. 78-85).
 1068 Maspalomas, Canary Islands, Spain: SPIE.

1069 Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M.,
 1070 Chetwynd, J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E.,
 1071 Shettle, E.P., 2006. Modtran5: 2006 update. In, *SPIE 6233, Algorithms and Technologies for*
 1072 *Multispectral, Hyperspectral, and Ultraspectral Imagery XII* (pp. 1-8). Orlando (Kissimmee), Florida,
 1073 United States: SPIE.

1074 Chen, Y., Duan, S.-B., Ren, H., Labeled, J., Li, Z.-L., 2017. Algorithm development for land surface
 1075 temperature retrieval: Application to chinese gaofen-5 data. *Remote Sens.* 9, 1-19.

1076 Chevallier, F., Ch éruy, F., Scott, N.A., Ch édin, A., 1998. A neural network approach for a fast and accurate
 1077 computation of a longwave radiative budget. *J. Appl. Meteorol.* 37, 1385-1397.

1078 Coll, C., Caselles, V., 1997. A split-window algorithm for land surface temperature from advanced very high
 1079 resolution radiometer data: Validation and algorithm comparison. *J. Geophys. Res.* 102, 16697-16713.

1080 Coll, C., Wan, Z., Galve, J.M., 2009. Temperature-based and radiance-based validations of the v5 modis land
 1081 surface temperature product. *J. Geophys. Res.* 114, 1-15.

1082 Duan, S.-B., Li, Z.-L., Cheng, J., Leng, P., 2017. Cross-satellite comparison of operational land surface
 1083 temperature products derived from modis and ASTER data over bare soil surfaces. *ISPRS J.*
 1084 *Photogramm. Remote Sens.* 126, 1-10.

1085 Duan, S.-B., Li, Z.-L., Wu, H., Leng, P., Gao, M., Wang, C., 2018. Radiance-based validation of land surface
 1086 temperature products derived from collection 6 modis thermal infrared data. *Int. J. Appl. Earth Obs.*
 1087 *Geoinf.* 70, 84-92.

1088 Duan, S.-B., Li, Z.-L., Li, H., G ötsche, F.-M., Wu, H., Zhao, W., Leng, P., Zhang, X., Coll, C., 2019a.
 1089 Validation of collection 6 modis land surface temperature product using in situ measurements. *Remote*
 1090 *Sens. Environ.* 225, 16-29.

1091 Duan, S.-B., Li, Z.-L., Wang, C., Zhang, S., Tang, B.-H., Leng, P., Gao, M.-F., 2019b. Land-surface
 1092 temperature retrieval from landsat 8 single-channel thermal infrared data in combination with ncep
 1093 reanalysis data and ASTER ged product. *Int. J. Remote Sens.* 40, 1763-1778.

1094 Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. Modis

1095 collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote*
1096 *Sens. Environ.* 114, 168-182.

1097 Gao, C., Jiang, X., Wu, H., Tang, B., Li, Z., Li, Z.-L., 2012. Comparison of land surface temperatures from
1098 msg-2/seviri and terra/modis. *J. Appl. Remote Sens.* 6, 1-16.

1099 Gao, C., Li, Z.-L., Qiu, S., Tang, B., Wu, H., Jiang, X., 2014. An improved algorithm for retrieving land
1100 surface emissivity and temperature from msg-2/seviri data. *IEEE Trans. Geosci. Remote Sens.* 52, 3175-
1101 3191.

1102 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.,
1103 2017. Land surface temperature retrieval from fy-3c/virr data and its cross-validation with terra/modis.
1104 *IEEE J-Stars.* 10, 4944-4953.

1105 Gerace, A., Kleynhans, T., Eon, R., Montanaro, M., 2020. Towards an operational, split window-derived
1106 surface temperature product for the thermal infrared sensors onboard landsat 8 and 9. *Remote Sens.* 12,
1107 1-15.

1108 Ghent, D.J., Corlett, G.K., Göttsche, F.M., Remedios, J.J., 2017. Global land surface temperature from the
1109 along-track scanning radiometers. *J. Geophys. Res. Atmos.* 122, 12167-12193.

1110 Gillespie, A., Rokugawa, S., Matsunaga, T., Cothorn, J.S., Hook, S., Kahle, A.B., 1998. A temperature and
1111 emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer
1112 (ASTER) images. *IEEE Trans. Geosci. Remote Sens.* 36, 1113-1126.

1113 Guillevic, P.C., Privette, J.L., Coudert, B., Palecki, M.A., Demarty, J., Ottlé C., Augustine, J.A., 2012. Land
1114 surface temperature product validation using noaa's surface climate observation networks—scaling
1115 methodology for the visible infrared imager radiometer suite (viirs). *Remote Sens. Environ.* 124, 282-
1116 298.

1117 Guillevic, P.C., Biard, J.C., Hulley, G.C., Privette, J.L., Hook, S.J., Olioso, A., Göttsche, F.M., Radocinski,
1118 R., Román, M.O., Yu, Y., Csiszar, I., 2014. Validation of land surface temperature products derived from
1119 the visible infrared imaging radiometer suite (viirs) using ground-based and heritage satellite
1120 measurements. *Remote Sens. Environ.* 154, 19-37.

1121 Guo, J., Ren, H., Zheng, Y., Lu, S., Dong, J., 2020. Evaluation of land surface temperature retrieval from
1122 landsat 8/tirs images before and after stray light correction using the surfrad dataset. *Remote Sens.* 12,
1123 1-21.

1124 Heidinger, A.K., Laszlo, I., Molling, C.C., Tarpley, D., 2013. Using surfrad to verify the noaa single-channel
1125 land surface temperature algorithm. *J. Atmos. Ocean. Technol.* 30, 2868-2884.

1126 [dataset] Hulley, G. (2015), Myd21 modis/aqua land surface temperature/3-band emissivity 5-min l2 1km
1127 v006. U.S.A, Accessed: [Aug. 12, 2021], 10.5067/MODIS/MOD11_L2.006.

1128 Hulley, G., Malakar, N., Freepartner, R. (2016). Moderate resolution imaging spectroradiometer (modis) land
1129 surface temperature and emissivity product (mxd21) algorithm theoretical basis document collection-6.
1130 Pasadena, California. J.P. Laboratory.

1131 Hulley, G.C., Hook, S.J., 2011. Generating consistent land surface temperature and emissivity products
1132 between ASTER and modis data for earth science research. *IEEE Trans. Geosci. Remote Sens.* 49, 1304-
1133 1315.

1134 Hulley, G.C., Hughes, C.G., Hook, S.J., 2012. Quantifying uncertainties in land surface temperature and
1135 emissivity retrievals from ASTER and modis thermal infrared data. *J. Geophys. Res. Atmos.* 117, 1-18.

1136 Hulley, G.C., Malakar, N.K., Islam, T., Freepartner, R.J., 2018. Nasa's modis and viirs land surface
1137 temperature and emissivity products: A long-term and consistent earth system data record. *IEEE J-Stars.*
1138 11, 522-535.

- 1139 Hulley, G.C., Gottsche, F.M., Rivera, G., Hook, S.J., Freepartner, R.J., Martin, M.A., Cawse-Nicholson, K.,
 1140 Johnson, W.R., 2021. Validation and quality assessment of the ecostress level-2 land surface temperature
 1141 and emissivity product. *IEEE Trans. Geosci. Remote Sens. Early Access*, 1-23.
- 1142 Jacob, F., Petitcolin, F.o., Schmugge, T., Vermote, É., French, A., Ogawa, K., 2004. Comparison of land
 1143 surface emissivity and radiometric temperature derived from modis and ASTER sensors. *Remote Sens.*
 1144 *Environ.* 90, 137-152.
- 1145 Jiang, G.-M., Li, Z.-L., 2008. Split - window algorithm for land surface temperature estimation from msg1 -
 1146 sevir data. *Int. J. Remote Sens.* 29, 6067-6074.
- 1147 Jimenez-Munoz, J.C., Sobrino, J.A., 2003. A generalized single-channel method for retrieving land surface
 1148 temperature from remote sensing data. *J. Geophys. Res.* 108, 1-9.
- 1149 Jimenez-Munoz, J.C., Cristobal, J., Sobrino, J.A., Soria, G., Ninyerola, M., Pons, X., Pons, X., 2009. Revision
 1150 of the single-channel algorithm for land surface temperature retrieval from landsat thermal-infrared data.
 1151 *IEEE Trans. Geosci. Remote Sens.* 47, 339-349.
- 1152 Kappas, M., Phan, T.N., 2018. Application of modis land surface temperature data: A systematic literature
 1153 review and analysis. *J. Appl. Remote Sens.* 12, 1-20.
- 1154 Krishnan, P., Meyers, T.P., Hook, S.J., Heuer, M., Senn, D., Dumas, E.J., 2020. Intercomparison of in situ
 1155 sensors for ground-based land surface temperature measurements. *Sensors (Basel)* 20, 1-26.
- 1156 Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J., Scott, J.K., 2012. Climond: Global
 1157 high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods Ecol.*
 1158 *Evol.* 3, 53-64.
- 1159 Li, H., Li, R., Yang, Y., Cao, B., Bian, Z., Hu, T., Du, Y., Sun, L., Liu, Q., 2021. Temperature-based and
 1160 radiance-based validation of the collection 6 myd11 and myd21 land surface temperature products over
 1161 barren surfaces in northwestern china. *IEEE Trans. Geosci. Remote Sens.* 59, 1794-1807.
- 1162 Li, S., Yu, Y., Sun, D., Tarpley, D., Zhan, X., Chiu, L., 2014. Evaluation of 10 year aqua/modis land surface
 1163 temperature with surfrad observations. *Int. J. Remote Sens.* 35, 830-856.
- 1164 Li, Z.-L., Tang, R., Wan, Z., Bi, Y., Zhou, C., Tang, B., Yan, G., Zhang, X., 2009. A review of current
 1165 methodologies for regional evapotranspiration estimation from remotely sensed data. *Sensors (Basel)* 9,
 1166 3801-3853.
- 1167 Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., Sobrino, J.A., 2013a. Satellite-derived
 1168 land surface temperature: Current status and perspectives. *Remote Sens. Environ.* 131, 14-37.
- 1169 Li, Z.-L., Wu, H., Wang, N., Qiu, S., Sobrino, J.A., Wan, Z., Tang, B.-H., Yan, G., 2013b. Land surface
 1170 emissivity retrieval from satellite data. *Int. J. Remote Sens.* 34, 3084-3127.
- 1171 Liu, Y., Yu, Y., Yu, P., Göttsche, F., Trigo, I., 2015. Quality assessment of s-npp viirs land surface temperature
 1172 product. *Remote Sens.* 7, 12215-12241.
- 1173 Malakar, N.K., Hulley, G.C., Hook, S.J., Laraby, K., Cook, M., Schott, J.R., 2018. An operational land surface
 1174 temperature product for landsat thermal data: Methodology and validation. *IEEE Trans. Geosci. Remote*
 1175 *Sens.* 56, 5717-5735.
- 1176 McMillin, L.M., 1975. Estimation of sea surface temperatures from two infrared window measurements with
 1177 different absorption. *J. Geophys. Res.* 80, 5113-5117.
- 1178 Niclos, R., Galve, J.M., Valiente, J.A., Estrela, M.J., Coll, C., 2011. Accuracy assessment of land surface
 1179 temperature retrievals from msg2-sevir data. *Remote Sens. Environ.* 115, 2126-2140.
- 1180 O'Connor, B., Bojinski, S., Rössli, C., Schaepman, M.E., 2020. Monitoring global changes in biodiversity
 1181 and climate essential as ecological crisis intensifies. *Ecol. Inform.* 55, 1-8.
- 1182 Ottlé C., Vidal-Madjar, D., 1992. Estimation of land surface temperature with noaa9 data. *Remote Sens.*

1183 Environ. 40, 27-41.

1184 Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the köppen-geiger climate
1185 classification. *Hydrol. Earth Syst. Sci.* 11, 1633-1644.

1186 Páez-Planells, L., Ničlòs, R., Puchades, J., Coll, C., Göttsche, F.-M., Valiente, J.A., Valor, E., Galve, J.M.,
1187 2021. Validation of sentinel-3 slstr land surface temperature retrieved by the operational product and
1188 comparison with explicitly emissivity-dependent algorithms. *Remote Sens.* 13, 1-25.

1189 Pinker, R.T., Ma, Y., Chen, W., Hulley, G., Borbas, E., Islam, T., Hain, C., Cawse-Nicholson, K., Hook, S.,
1190 Basara, J., 2019. Towards a unified and coherent land surface temperature earth system data record from
1191 geostationary satellites. *Remote Sens.* 11, 1-23.

1192 Qian, Y.-G., Li, Z.-L., Nerry, F., 2013. Evaluation of land surface temperature and emissivities retrieved from
1193 msg/seviri data with modis land surface temperature and emissivity products. *Int. J. Remote Sens.* 34,
1194 3140-3152.

1195 Qin, Z., Dall'Olmo, G., Karnieli, A., Berliner, P., 2001a. Derivation of split window algorithm and its
1196 sensitivity analysis for retrieving land surface temperature from noaa-advanced very high resolution
1197 radiometer data. *J. Geophys. Res.-Atmos.* 106, 22655-22670.

1198 Qin, Z., Karnieli, A., Berliner, P., 2001b. A mono-window algorithm for retrieving land surface temperature
1199 from landsat tm data and its application to the israel-egypt border region. *Int. J. Remote Sens.* 22, 3719-
1200 3746.

1201 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
1202 algorithm for land surface temperature retrieval from multiple-band thermal infrared image without
1203 atmospheric and emissivity data inputs. *Int. J. Digit. Earth.* 13, 1430-1453.

1204 Sánchez-Aparicio, M., Andrés-Anaya, P., Del Pozo, S., Lagüela, S., 2020. Retrieving land surface
1205 temperature from satellite imagery with a novel combined strategy. *Remote Sens.* 12, 1-18.

1206 Sattari, F., Hashim, M., 2014. A breife review of land surface temperature retrieval methods from thermal
1207 satellite sensors. *Middle East J. Sci. Res.* 22, 757-768.

1208 Snyder, W.C., Wan, Z., Zhang, Y., Feng, Y.Z., 1998. Classification-based emissivity for land surface
1209 temperature measurement from space. *Int. J. Remote Sens.* 19, 2753-2774.

1210 Sobrino, J.A., Coll, C., Caselles, V., 1991. Atmospheric correction for land surface temperature using noaa-
1211 11 avhrr channels 4 and 5. *Remote Sens. Environ.* 38, 19-34.

1212 Sobrino, J.A., Del Frate, F., Drusch, M., Jimenez-Munoz, J.C., Manunta, P., Regan, A., 2016. Review of
1213 thermal infrared applications and requirements for future high-resolution sensors. *IEEE Trans. Geosci.*
1214 *Remote Sens.* 54, 2963-2972.

1215 Sulla-Menashe, D., Friedl, M.A. (2018). User guide to collection 6 modis land cover (mcd12q1 and mcd12c1)
1216 product. South Dakota USA. N.E.L.P. DAAC.

1217 Sulla-Menashe, D., Gray, J.M., Abercrombie, S.P., Friedl, M.A., 2019. Hierarchical mapping of annual global
1218 land cover 2001 to present: The modis collection 6 land cover product. *Remote Sens. Environ.* 222, 183-
1219 194.

1220 Tomlinson, C.J., Chapman, L., Thornes, J.E., Baker, C., 2011. Remote sensing land surface temperature for
1221 meteorology and climatology: A review. *Meteorol. Appl.* 18, 296-306.

1222 Tonooka, H., 2005. Accurate atmospheric correction of ASTER thermal infrared imagery using the wvs
1223 method. *IEEE Trans. Geosci. Remote Sens.* 43, 2778-2792.

1224 Townshend, J.R.G., Justice, C.O., Skole, D., Malingreau, J.P., Cihlar, J., Teillet, P., Sadowski, F., Ruttenberg,
1225 S., 1994. The 1 km resolution global data set: Needs of the international geosphere biosphere programme.
1226 *Int. J. Remote Sens.* 15, 3417-3441.

1227 Trigo, I.F., Monteiro, I.T., Olesen, F., Kabsch, E., 2008. An assessment of remotely sensed land surface
1228 temperature. *J. Geophys. Res.* 113, 1-12.

1229 Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from
1230 space. *IEEE Trans. Geosci. Remote Sens.* 34, 892-905.

1231 Wan, Z., Li, Z.-L., 1997. A physics-based algorithm for retrieving land-surface emissivity and temperature
1232 from eos/modis data. *IEEE Trans. Geosci. Remote Sens.* 35, 980-996.

1233 Wan, Z., 2008. New refinements and validation of the modis land-surface temperature/emissivity products.
1234 *Remote Sens. Environ.* 112, 59-74.

1235 Wan, Z., Li, Z.-L., 2008. Radiance - based validation of the v5 modis land - surface temperature product.
1236 *Int. J. Remote Sens.* 29, 5373-5395.

1237 Wan, Z., 2014. New refinements and validation of the collection-6 modis land-surface temperature/emissivity
1238 product. *Remote Sens. Environ.* 140, 36-45.

1239 [dataset] Wan, Z., Hook, S., Hulley, G. (2015), Mod11_l2 modis/terra land surface temperature/emissivity 5-
1240 min l2 swath 1km v006. U.S.A, Accessed: [Feb. 5, 2019], 10.5067/MODIS/MOD11_L2.006.

1241 Wang, H., Mao, K., Mu, F., Shi, J., Yang, J., Li, Z., Qin, Z., 2019a. A split window algorithm for retrieving
1242 land surface temperature from fy-3d mersi-2 data. *Remote Sens.* 11, 1-25.

1243 Wang, K., Liang, S., 2009. Evaluation of ASTER and modis land surface temperature and emissivity products
1244 using long-term surface longwave radiation observations at surfrad sites. *Remote Sens. Environ.* 113,
1245 1556-1565.

1246 Wang, M., He, G., Zhang, Z., Wang, G., Wang, Z., Yin, R., Cui, S., Wu, Z., Cao, X., 2019b. A radiance-based
1247 split-window algorithm for land surface temperature retrieval: Theory and application to modis data. *Int.*
1248 *J. Appl. Earth Obs. Geoinf.* 76, 204-217.

1249 Xing, Z., Li, Z.-L., Duan, S.-B., Liu, X., Zheng, X., Leng, P., Gao, M., Zhang, X., Shang, G., 2021. Estimation
1250 of daily mean land surface temperature at global scale using pairs of daytime and nighttime modis
1251 instantaneous observations. *ISPRS J. Photogramm. Remote Sens.* 178, 51-67.

1252 Yao, R., Wang, L., Huang, X., Gong, W., Xia, X., 2019. Greening in rural areas increases the surface urban
1253 heat island intensity. *Geophys. Res. Lett.* 46, 2204-2212.

1254 Yao, R., Wang, L., Wang, S., Wang, L., Wei, J., Li, J., Yu, D., 2020. A detailed comparison of myd11 and
1255 myd21 land surface temperature products in mainland china. *Int. J. Digit. Earth.* 13, 1-17.

1256 Zheng, X., Li, Z.-L., Nerry, F., Zhang, X., 2019. A new thermal infrared channel configuration for accurate
1257 land surface temperature retrieval from satellite data. *Remote Sens. Environ.* 231, 1-12.

1258