

# A stepwise unmixing model to address the scale gap issue present in downscaling of geostationary meteorological satellite surface temperature images



Fei Xu <sup>a</sup>, Xiaolin Zhu <sup>a,\*</sup>, Jin Chen <sup>b</sup>, Wenfeng Zhan <sup>c</sup>

<sup>a</sup> Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

<sup>b</sup> State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, China

<sup>c</sup> Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, International Institute for Earth System Science, Nanjing University, Nanjing 210023, China

## ARTICLE INFO

Edited by Marie Weiss

**Keywords:**

Geostationary satellite  
Downscaling  
Temperature unmixing  
GOES-R  
Landsat

## ABSTRACT

Land surface temperature (LST) is a critical parameter that drives the response of a variety of ecosystems to environmental and climatic changes. The geostationary satellite brings unique opportunities to monitor the LST at a hemispheric scale with temporal resolutions of up to 5 min. However, the ultra-coarse spatial resolutions ranging from 2 km to 5 km limit its application at local spatial scales. Downscaling the geostationary satellite LST image with the high-resolution low-Earth-orbit satellite images is a cost-effective way to circumvent this dilemma. Yet, the big gap between the observation scales of these satellite data poses a challenge for accurate downscaling. To address this problem, we proposed a stepwise temperature unmixing (TUM) model called 'UnmixGO', which downscale hourly LST images of Geostationary Operational Environmental Satellites (GOES-R) from 2 km to 100 m resolution. The spatially adaptive endmembers and the constrained solution space of the TUM model keep the errors in downscaled LSTs from being over-amplified in the stepwise data treatment. We validated the algorithm in six experimental areas and at five flux tower sites across the contiguous United States, revealing that UnmixGO outperformed conventional methods in accuracy by 0.49 K and 1.11 K on average for downscaling the simulated and real GOES-R LST images, respectively. Furthermore, the technical framework employed by UnmixGO is compatible with multi-source satellite images, enhancing the added value of our study in a future where various remote sensing data is increasingly accessible.

## 1. Introduction

Land surface temperature (LST) is a critical measure of the energy exchange between the land and atmosphere, and it changes rapidly with variations in solar incidence angle (Li et al., 2013; Wang et al., 2021a). Temporally dense LST data are therefore valuable to model prompt responses of ecosystems to environmental and climatic changes (Bindhu et al., 2013; Carlson et al., 2021; Enku and Melesse, 2014; Singh et al., 2017). Spaceborne remote sensing has been recognized as the most efficient way to obtain this information, with geostationary meteorological satellites offering the unique capability to observe surface thermal conditions with a temporal resolution of up to 5 min. However, geostationary satellite LST (Geo-LST) images have ultra-coarse spatial resolutions down to 5 km, which restricts the ability of this data to reveal

thermal variations in regions with heterogeneous surfaces.

Downscaling the Geo-LST image has proven to be an effective and low-cost solution to this dilemma. Existing studies generally performed two branches of technical routes, including (i) fusing Geo-LST images with polar-orbiting satellite LST (Pol-LST) images and (ii) downscaling Geo-LST images with polar-orbiting satellite surface reflectance (Pol-SR) images. Two representative methods have been widely used accordingly. The first methodology operates a spatiotemporal data fusion model, which simulates LST changes at high spatial resolutions by summing the temporally dense observations of Geo-LST image pixels in a weighted manner, and then adds the simulated LST variations to the Pol-LST images with spatial resolutions up to 100 m (Huang et al., 2013; Li et al., 2023a; Quan et al., 2018; Weng and Fu, 2014; Wu et al., 2015). The second approach maps different kernels (e.g., normalized difference

\* Corresponding author.

E-mail address: [xiaolin.zhu@polyu.edu.hk](mailto:xiaolin.zhu@polyu.edu.hk) (X. Zhu).

vegetation index and normalized difference built-up index) calculated from surface reflectance data to temporal LSTs. Tools to establish such correspondences include linear regression (Agam et al., 2007; Bechtel et al., 2012; Li et al., 2019; Wang et al., 2020a; Zakšek and Oštir, 2012) and geographically weighted regression (Duan and Li, 2016; Wang et al., 2021b; Zhang et al., 2020) models, as well as machine learning models (e.g., support vector machine and random forest) and deep learning models (Chen et al., 2022; Kumar, 2022; Moosavi et al., 2015; Wu et al., 2022a; Yang et al., 2009; Yin et al., 2020; Yuan et al., 2020). Notably, these kernel regression models are typically trained at a low spatial resolution (e.g., 5 km) but are implemented at a spatial resolution up to 30 m based on the assumption that the dependence between kernels and LST is scale invariant in local space (Lezama Valdes et al., 2021; Sismanidis et al., 2016; Weng and Fu, 2014).

Yet, the ‘scale gap issue’, which refers to the considerable difference in spatial resolution between input and output images, might be the primary factor limiting the performance of existing techniques in downscaling Geo-LST images. Practically, the observation of a Pol-LST image pixel (100 m resolution) might hardly represent the thermal status of the broader view observed by a Geo-LST image pixel (2 to 5 km resolution), even though the spatial coverage of the two pixels is mutually inclusive. This mismatch in observational scales could lead to low comparability of surface information recorded by Pol-LST image and Geo-LST image. Moreover, implementing geometric registration of satellite images can be difficult because the texture information in Pol-LST images is almost indiscernible in the Geo-LST image. These derivative problems ultimately undermine the accuracy of downscaled LSTs (Zhou et al., 2021).

Various solutions have been proposed to address the scale gap issue. Studies using kernel regression models tried to select spectral indices as kernels that are insensitive to changes in the satellite observation scale for the same land cover type (Agam et al., 2007; Bonafoni, 2016; Wu et al., 2022b). However, this operation may not be applicable to regions with heterogeneous surfaces. On the other hand, studies using spatio-temporal data fusion models made use of LST image with moderate spatial resolution (e.g., 1 km resolution) to bridge the scale gap between Geo-LST image and the Pol-LST image with spatial resolutions up to 100 m (Wu et al., 2015; Quan et al., 2018). However, acquisition times for the LST images being fused are typically inconsistent and can be 15 to 30 min apart, which introduces the temporal variation of LST as an additional source of error into the fused results (Ma et al., 2022).

Additionally, Temperature UnMixing technique (TUM, Mitraka et al., 2015; Zhan et al., 2010) is another alternative to address the scale gap issue. TUM uses a temperature mixing model to decompose the thermal observation of a pixel into elemental temperatures, or ‘end-members’, of a set of pixels with higher spatial resolution (Collins et al., 2001). By doing so, TUM bridges the LSTs observed across spatial scales physically and has proven robust in various studies (Dennison et al., 2006; Pu and Bonafoni, 2023). Also, Pol-SR and Geo-LST images do not necessarily have to be acquired at the same time for TUM. The main reason is that the endmembers determined from the Pol-SR image represent true thermal compositions within the Geo-LST pixel’s field of view, as long as the study area has not experienced surface changes between the two image acquisition times.

Despite its potential, TUM methodology has rarely been used to downscale Geo-LST images in past studies due to the large-scale gap between input and output images, which undermines its applicability. For example, when downscaling a Geo-LST image pixel from 2 km to 100 m resolution, a temperature mixing equation with up to 400 end-members (i.e., unknown variables) is supposed to be solved. However, solving this model implies building a system containing the same number of temperature mixing equations as the number of endmembers, which might be infeasible in practice. Moreover, conventional TUM-based image downscaling methods determine the endmembers by clustering the pixels in a high spatial resolution Pol-SR image, and the number of clusters (corresponding to the number of endmembers) is

often set manually (Li et al., 2023b; Zhan et al., 2013). This perhaps limits the richness of thermal information in the downsampled LST image. Furthermore, the types of endmembers being identified and the collinearity between endmembers have been shown to affect the accuracy of the solution for temperature mixing models (Chen et al., 2011). Existing studies, however, rarely constrained the solution space of endmembers (Deng and Wu, 2013; Mitraka et al., 2015; Pu and Bonafoni, 2021) or set a broad range of solutions such as making the solved endmembers to be positive (Li et al., 2023c; Wang et al., 2020b).

To address the limitations mentioned above, this paper presents a stepwise temperature unmixing model (UnmixGO - ‘Unmixing-based downscaling of GOES-R LST image’) to downscale the LST images acquired by NOAA’s latest generation of geostationary meteorological satellites, GOES-R, to produce hourly surface temperature images at the 100 m resolution. Our proposed UnmixGO addresses the scale gap issue by dividing the downscaling problem with a large-scale gap into three subset problems with smaller scale gaps. Also, UnmixGO features two innovations over the conventional TUM-based downscaling methods including (i) determining spatially adaptive endmember types for each Geo-LST image pixel and (ii) effectively constraining the solution space of the temperature mixing model. Furthermore, the downscaling framework employed by UnmixGO is compatible with multi-source satellite images to downscale the Geo-LST image. To test the generic nature and robustness of the newly proposed method, the high spatio-temporal LST images were generated in selected areas across contiguous United States and were cross-referenced to Landsat LST products and in-situ LSTs.

## 2. Materials

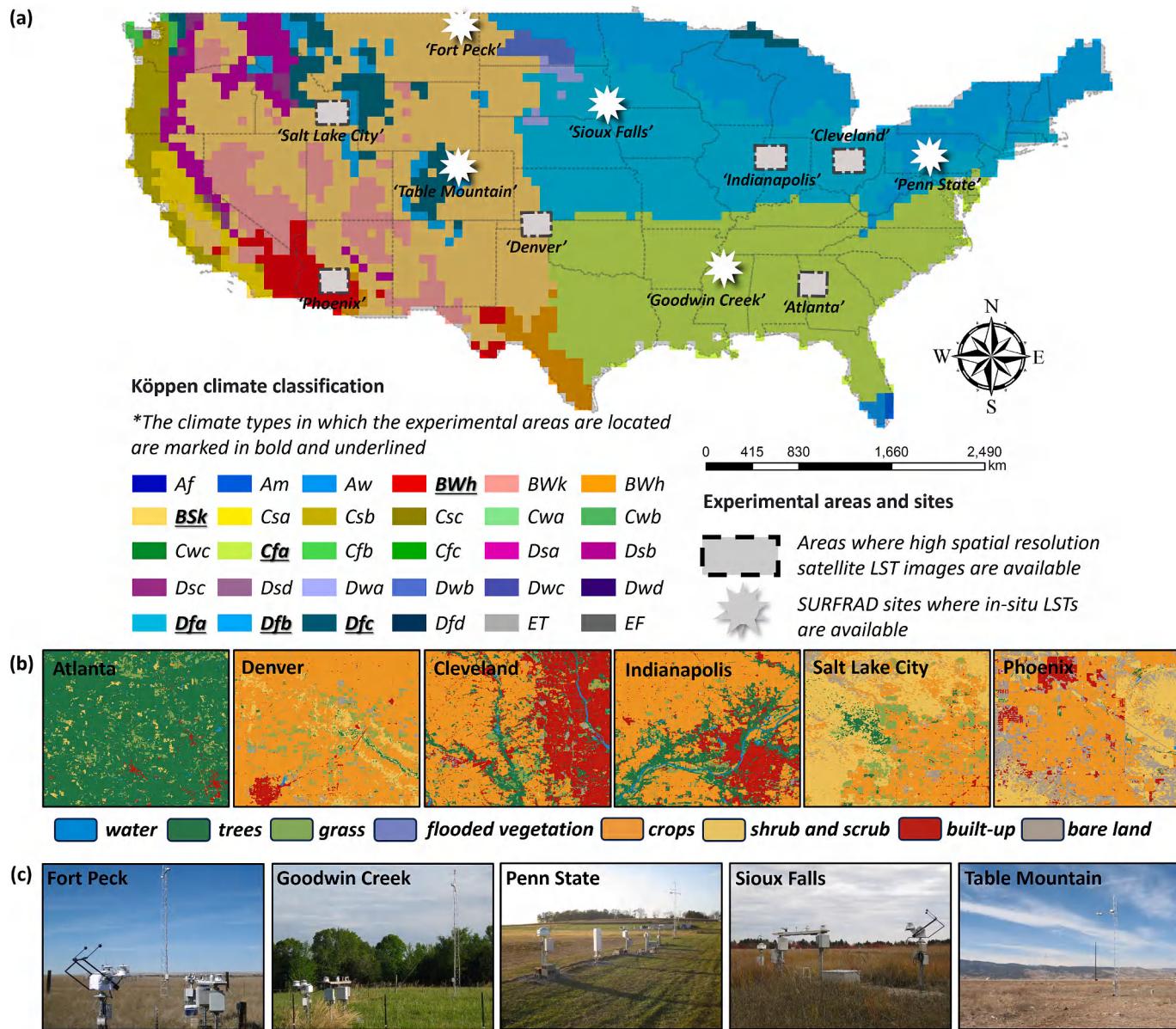
### 2.1. Study areas

As shown in Fig. 1, we performed experiments throughout the contiguous United States (CONUS), in the states of Arizona, Colorado, Georgia, Indiana, Mississippi, Montana, Ohio, Pennsylvania, South Dakota, Uta, and Wyoming, and in climate zones including BWh (arid, desert, cold), Cfa (temperate, no dry season, hot summer), Dfa (cold, no dry season, hot summer), and BSk (arid, steppe, cold), according to the Köppen-Geiger climate classification (Beck et al., 2018). The surface types in study areas include croplands, deserts, forests, grasslands, lakes, rivers, and settlements, providing a complex picture of surface thermal conditions.

The criteria for selecting experimental areas include: (i) the areas should be in a wide range of climate zones to allow for a diversity of phenological land cover changes; (ii) land surfaces should be spatio-temporally heterogeneous to evaluate the robustness of the LST image downscaling methods; and (iii) data availability in experimental areas should satisfy the requirements for performing the tested LST image downscaling methods. Thus, we selected two groups of experimental areas to validate the LST downscaling algorithm.

The first set of experimental areas includes ‘Atlanta’, ‘Cleveland’, ‘Denver’, ‘Indianapolis’, ‘Phoenix’, and ‘Salt Lake City’ where we used the simulated GOES-R LST images as input to the LST downscaling algorithm and used the Landsat LST image as reference data (introduced in Subsection 4.1). These areas were named after their neighboring cities, which does not mean that they are the cities’ administrative areas. All the image data used in these areas were summarized in Table 1 and were presented in Fig. A1.

The second set of experimental areas refers to the five SURFace RADiation budget network (SURFRAD, Augustine et al., 2005) sites, namely ‘Fort Peck’, ‘Goodwin Creek’, ‘Penn State’, ‘Sioux Falls’, and ‘Table Mountain’. We acquired real GOES-R LST images covering these sites as input to the LST downscaling algorithm and used in-situ LSTs as reference data (introduced in Subsection 4.1). All the image data used at SURFRAD sites were summarized in Table 2 and were presented in Fig. A2.



**Fig. 1.** (a) The map demonstrating spatial distributions of experimental areas and SURFRAD sites as well as climate zones for CONUS in 2010. Detailed explanations of climate types can be referred to Beck et al., 2018; (b) Dynamic World land cover maps (Brown et al., 2022) of experimental areas for studied periods of time; (c) On-site photos of SURFRAD sites obtained from the NOAA Global Monitoring Laboratory (<https://gml.noaa.gov/grad/surfrad/>).

## 2.2. Satellite images

Our study used four types of satellite image data to perform LST downscaling: (i) GOES-R LST image with 2 km spatial resolution and hourly temporal resolution; (ii) MODIS LST image with 1 km spatial resolution and up to four data records per day; (iii) Landsat LST image with 100 m spatial resolution and 16-day temporal resolution; (iv) Harmonized Landsat Sentinel-2 (HLS) surface reflectance image with 30 m spatial resolution and an average of 2.5-day temporal resolution. The HLS reflectance image holds spectral information ranging from visible to near-infrared wavelengths.

Specifically, the GOES-R LST images were obtained from the NOAA Comprehensive Large Array-data Stewardship System (<https://www.av1.class.noaa.gov/saa/products/catSearch>), MODIS and HLS image data were obtained from the NASA Earth Science Data Systems (<https://search.earthdata.nasa.gov>), and seasonal Landsat LST images were acquired from the USGS Earth Explorer (<https://earthexplorer.usgs.gov>). Cloud-free images were downloaded for experiments. Specifically, the

GOES-R LST image and the MODIS LST image have the same acquisition date,  $day_{goes}$ , for each experimental area (and site). The acquisition date of the HLS image is close to  $day_{goes}$  and two seasonal Landsat LST images were taken from the nearest dates before and after  $day_{goes}$ , respectively. All the satellite images that were used in this study are shown in Tables 1 and 2.

MODIS and GOES-R LST images were geometrically calibrated with the 30 m resolution HLS reflectance image as a reference prior to all experimental operations. Taking calibrating the GOES-R LST image as an example, we firstly upscaled the HLS image to 2 km resolution to make its pixel size consistent with the GOES-R LST image; second, we selected ground control points (GCPs) from the GOES-R LST image using the upscaled HLS image as reference. The selected GCPs were the pixels that represent unique terrain features which are discernible at 2 km spatial resolution, such as lakefronts, intersections of rivers, and junctions of urban and rural areas; third, we conducted geometric corrections to the GOES-R LST image with the selected GCPs in QGIS software ([www.qgis.org](http://www.qgis.org)). By doing so, the map projection of the GOES-R LST

**Table 1**

The image data used in experimental areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City\*.

Experimental area	Satellite images	Date (DD/MM/YYYY) and time
Atlanta	Landsat LST	10/04/2022, 12:12 <b>19/10/2022, 12:13</b>
		04/03/2023, 12:12
	MODIS LST	19/10/2022, 14:10
Cleveland	HLS surface reflectance	21/10/2022
	Landsat LST	20/04/2022, 12:10 <b>21/06/2022, 12:10</b>
		11/10/2022, 12:10
Denver	MODIS LST	21/06/2022, 11:15
	HLS surface reflectance	18/06/2022
	Landsat LST	05/08/2022, 11:20 <b>01/11/2022, 11:20</b>
Indianapolis		19/12/2022, 11:20
	MODIS LST	01/11/2022, 11:15
	HLS surface reflectance	29/10/2022
Phoenix	Landsat LST	30/05/2021, 12:29 <b>06/11/2021, 12:29</b>
		18/06/2022, 12:29
	MODIS LST	06/11/2021, 17:30
Salt Lake City	HLS surface reflectance	27/10/2021
	Landsat LST	19/02/2022, 11:04 <b>25/10/2022, 11:04</b>
		19/04/2023, 11:04
	MODIS LST	25/10/2022, 10:45
	HLS surface reflectance	24/10/2022
	Landsat LST	11/07/2022, 12:14 <b>05/10/2022, 12:14</b>
		07/10/2022, 12:14
	MODIS LST	05/10/2022, 14:30
	HLS surface reflectance	23/08/2022

\* The Landsat LST image of the date in bold was used to simulate the 2 km resolution GOES-R LST image as well as being used as reference data for the downscaled LST image.

image would become consistent with the HLS image. Note that the geometric error of the corrected image would be kept less than one pixel size. Identical operations were conducted on MODIS LST image, except for HLS images being upscaled to the 1 km resolution.

Notably, we employed an ‘image aggregation’ approach to upscale the reflectance and LST images with high spatial resolution (referred to as fine image) to a low spatial resolution (referred to as coarse image). The approach takes the average pixel value of all fine image pixels located within the coarse pixel’s observational footprint as the value of the upscaled pixel. This data resampling approach can be expressed mathematically by a linear point spread function (PSF), which assigns equal weights to the fine image pixels to contribute to the upscaled reflectance value. Note that nonlinear PSFs can model the conversion between fine and coarse images more accurately, owing to accounting for additional energy source (e.g., multiple scattering between ground objects) that contributes to satellite imaging (Somers et al., 2009). However, determining an appropriate PSF in practice might be hard because the form of the PSF can be influenced by varying factors such as topography, land cover type, and sensor performance (Valenzuela et al., 2023). Alternatively, the linear PSF is a robust simplification of the satellite imaging mechanism, as widely used in experiments for fusing both reflectance and LST images (Dong et al., 2020; Gevaert and García-Haro, 2015; Wang et al., 2016; Zhao et al., 2018; Zhukov et al., 1999).

### 2.3. In-situ LST data

We used in-situ LSTs as a set of independent validation datasets, which were measured from SURFRAD sites. The SURFRAD stations, located on diverse types of surfaces (e.g., grassland and forest) across the CONUS, measure surface upwelling and downwelling radiation with a temporal resolution of less than three minutes. The radiometer mounted on the observation towers typically has an effective field of view of 30 to

**Table 2**

The image data used in experiments conducted at SURFRAD sites of Fort Peck, Goodwin Creek, Penn State, Sioux Falls, Table Mountain.

Experimental site	Datasets	Date (DD/MM/YYYY) and time
Fort Peck, Montana	GOES-R LST	08/08/2022, {00:00 to 24:00}
	Landsat LST	30/09/2022, 10:47 04/11/2022, 10:47
	MODIS LST	04/11/2022, 05:00
Goodwin Creek, Mississippi	HLS surface reflectance	04/08/2022
	In-situ LST	08/08/2022, {00:00 to 24:00}
	GOES-R LST	09/30/2022, {00:00 to 24:00}
Penn State, Pennsylvania	Landsat LST	23/02/2022, 10:31 04/01/2023, 10:31
	MODIS LST	09/30/2022, 10:30
	HLS surface reflectance	02/10/2022
Sioux Falls, South Dakota	In-situ LST	09/30/2022, {00:00 to 24:00}
	GOES-R LST	09/11/2020, {00:00 to 24:00}
	Landsat LST	22/09/2020, 10:52 22/10/2022, 10:52
Table Mountain, Colorado	MODIS LST	09/11/2020, 02:30
	HLS surface reflectance	29/10/2022
	In-situ LST	09/11/2020, {00:00 to 24:00}
	GOES-R LST	21/07/2022, {00:00 to 18:00}
	Landsat LST	29/04/2021, 11:11 30/08/2022, 11:11
	MODIS LST	21/07/2022, 10:25
	HLS surface reflectance	21/07/2022
	In-situ LST	21/07/2022, {00:00 to 18:00}
	GOES-R LST	21/06/2021, {00:00 to 9:00, 21:00}
	Landsat LST	29/09/2020, 11:38 29/10/2022, 11:38
	MODIS LST	21/06/2021, 11:40
	HLS surface reflectance	22/06/2021
	In-situ LST	21/06/2021, {00:00 to 9:00, 21:00}

45 m in diameter (Wang and Liang, 2009) and the measurements were calibrated regularly by NOAA maintenance team following the standards set by the Baseline Surface Radiation Network. The in-situ LSTs were derived from a surface energy balance model:

$$T_s = \left[ \frac{L_{\uparrow} - (1 - \varepsilon_b) \bullet L_{\downarrow}}{\varepsilon_b \bullet \sigma} \right]^{1/4} \quad (1)$$

where  $L_{\uparrow}$  and  $L_{\downarrow}$  represent the upwelling and downwelling longwave radiation, respectively;  $\varepsilon_b$  represents broad band emissivity over the entire infrared spectrum;  $\sigma$  is the Stefan-Boltzmann’s constant ( $5.67 \times 10^{-8} \text{ Wm}^{-2} \text{ K}^{-4}$ ). Practically,  $\varepsilon_b$  was estimated using an empirical model based on MODIS narrow-band emissivity products with center wavelengths at  $8.5 \mu\text{m}$ ,  $11.04 \mu\text{m}$ , and  $12.02 \mu\text{m}$  (Wang and Liang, 2009):

$$\varepsilon_b = 0.2122 \bullet \varepsilon_{8.5\mu\text{m}} + 0.3859 \bullet \varepsilon_{11.04\mu\text{m}} + 0.4029 \bullet \varepsilon_{12.02\mu\text{m}} \quad (2)$$

where  $\varepsilon_{8.5\mu\text{m}}$ ,  $\varepsilon_{11.04\mu\text{m}}$  and  $\varepsilon_{12.02\mu\text{m}}$  are emissivity products for the 29th, 31st, and 32nd bands of MODIS emissivity products (i.e., MOD11L2 and MYD11L2).

We derived in-situ LST values from 0:00 to 24:00 local time at an hourly interval within the acquisition day of the GOES-R LST image. The raw data were resampled to generate the LST value at the time of interest. The sampling window of raw LST values aligns with the period during which the GOES-R thermal imager integrates the surface emitted energy. For instance, knowing that GOES-R Advanced Baseline Imager (ABI) takes 0:01 to 0:03 AM to generate an observation for 0:01 AM, we averaged the LSTs derived from Eq. (1) for the same period as the resampled value.

### 3. Methods

#### 3.1. Briefing of temperature unmixing model

The Temperature UnMixing model (TUM, Collins et al., 2001; Denison et al., 2006) is an adaptation of subpixel spectral analysis (Adams et al., 1986) from reflectance images to thermal images. Specifically, the TUM model is used to disaggregate the temperature recorded by a low spatial resolution pixel (referred to as the ‘mixed pixel’) into temperatures of more than one surface material (endmembers) that occupy the mixed pixel in space. The LST of a mixed pixel can be expressed as (Keshava and Mustard, 2002):

$$LST_{mix} = F \times LST_{endmember} + \epsilon \quad (3)$$

in which  $LST_{mix}$  denotes the temperature recorded by the mixed pixel,  $LST_{endmember}$  and  $F$  represent the endmembers and their area fractions occupying the mixed pixel’s spatial coverage, respectively, and  $\epsilon$  represents the thermal information that cannot be modeled with the determined  $LST_{endmember}$  and  $F$ . If we knew the endmember types at a higher spatial resolution and their area fractions, we could mathematically derive the endmembers from Eq. (3).

#### 3.2. UnmixGO - unmixing-based downscaling of geostationary satellite LST image

In view of the explicit mechanism of TUM for bridging LSTs of distinct observational scales, we developed a novel temperature unmixing model to downscale the GOES-R LST image (called ‘UnmixGO’) from 2 km to 100 m resolution. To address the limitations of the conventional TUM methods, as discussed earlier in *Introduction*, UnmixGO features the following innovations (i) making the temperature unmixing model operational in downscaling the Geo-LST image by decomposing a ‘large scale gap’ downscaling problem into subset problems with smaller scale gaps, (ii) defining endmembers as the effective temperatures of sub-scenes that make up the field-of-view of the mixed pixel and determining the spatially adaptive endmember types for each mixed pixel, and (iii) constraining the solution space of

the temperature unmixing model effectively using the predictions of kernel regression model.

The overall framework of UnmixGO consists of three sequential downscaling tasks to address the scale gap issue (illustrated in Fig. 2). The first task is to downscale the GOES-R LST image from 2 km to 1 km resolution; the second one is to downscale the result of the previous step from 1 km to 500 m resolution; and the third one is downscaling the result of the second task from 500 m to 100 m resolution. In performing each task, UnmixGO carries out the following operations to downscale each mixed pixel through temperature unmixing: (i) determining endmember types (Subsection 3.2.1), (ii) establishing a set of LST mixing equations as well as solving the endmembers (Subsection 3.2.2), and (iii) constraining the solution space of the established LST mixing equations (Subsection 3.2.3).

#### 3.2.1. Determination of endmember types for each mixed LST pixel

The TUM model theoretically divides the thermal scene observed by a GOES-R LST pixel (TCP, target coarse pixel) into subset scenes represented by high spatial resolution pixels (TFP, target fine pixel) within its spatial coverage. Given that surface reflectance essentially characterizes the land surface type, we defined each TFP with unique reflectance spectral characteristics as an individual endmember type. Thus, determining each TCP’s endmembers is technically identical to categorizing the types of unique reflectance spectra that TFPs have. In this study, the reflectance signatures of TFPs were derived from the HLS reflectance image, which was acquired at a time close to the GOES-R LST image. Furthermore, we upscaled the 30 m resolution HLS image to 1 km, 500 m, and 100 m resolution, respectively, before conducting the three consecutive downscaling tasks (shown in Fig. 2).

We used a spectral matching approach (Xu and Somers, 2021; illustrated in Fig. 3) to categorize the TFPs into different endmember types. Specifically, the absolute spectral difference (ASD, Eq. (4)) between TFPs is used as a basis for determining whether a TFP can be considered an endmember type. If the ASD value of a TFP from other TFPs was greater than the predetermined threshold, we would recognize this TFP as an endmember type for the TCP of interest. Otherwise, the TFPs with ASD values greater than the threshold would be considered to

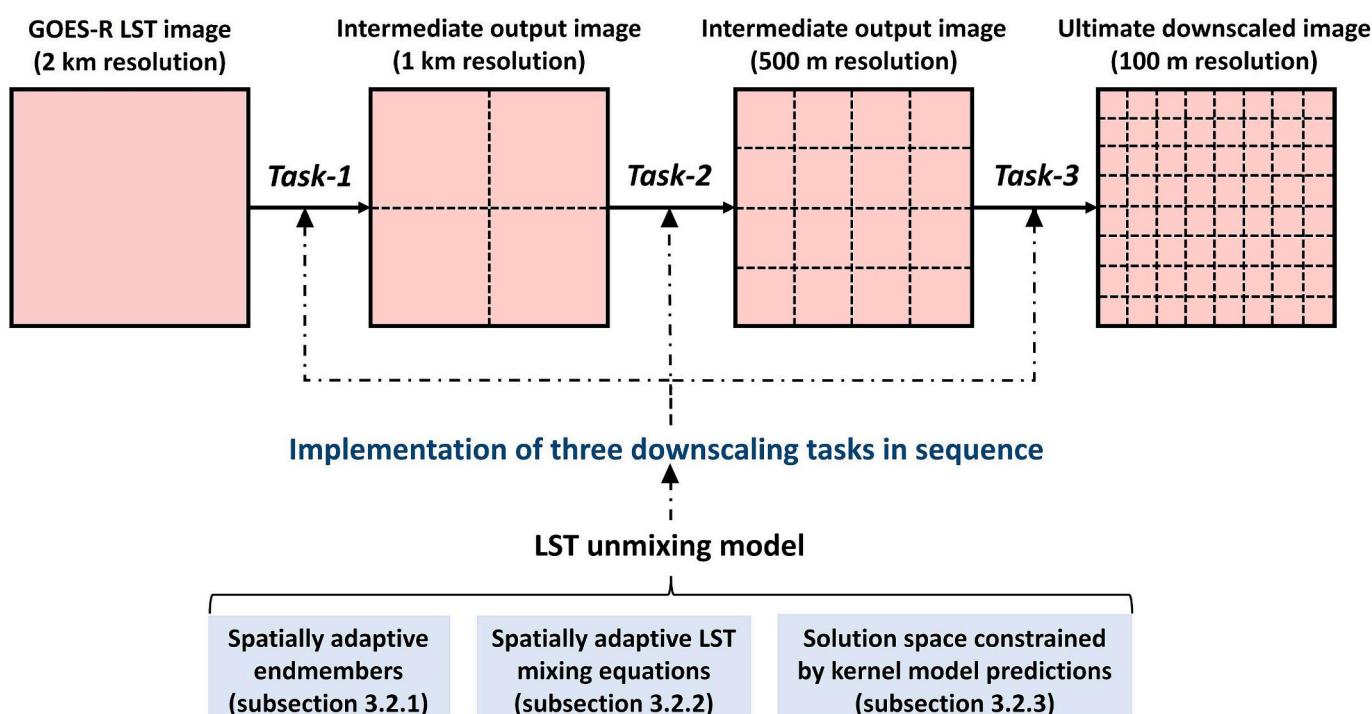
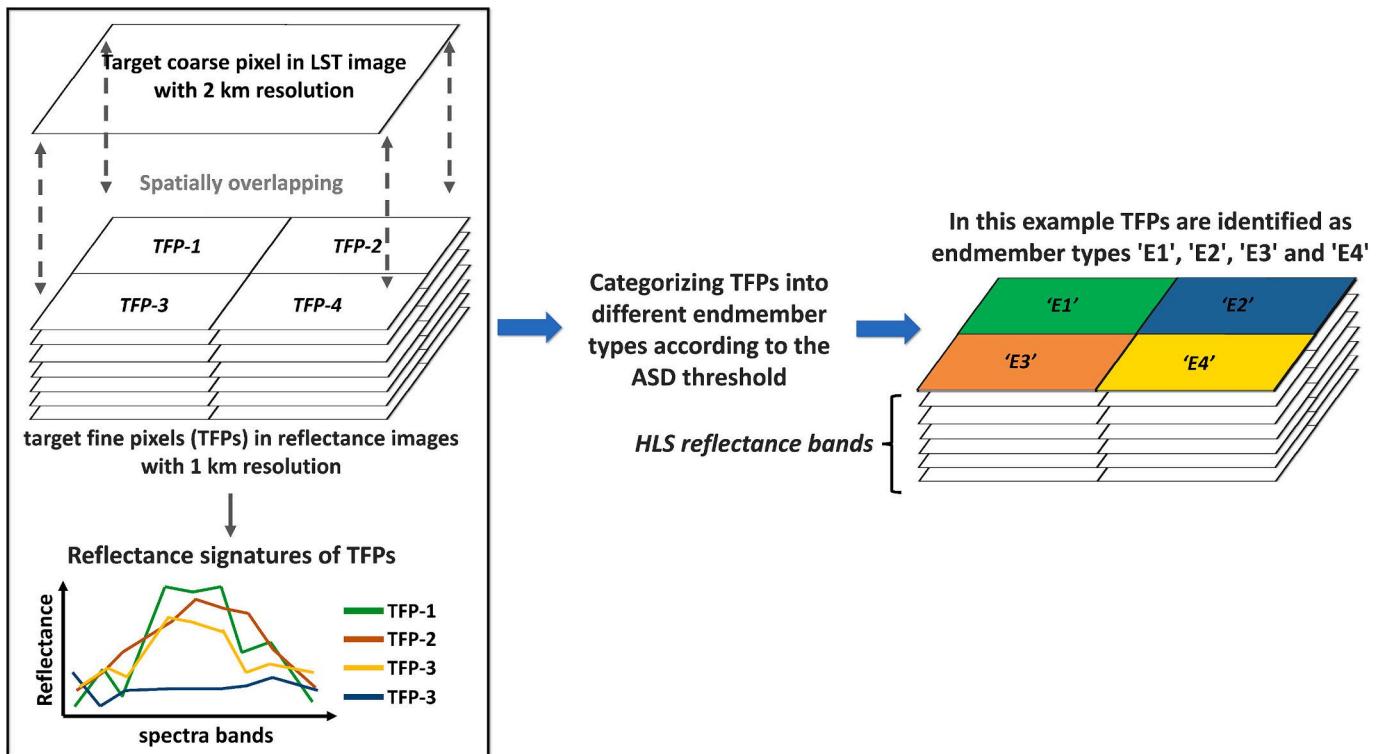


Fig. 2. The overall framework of UnmixGO to downscale GOES-R LST images in three consecutive tasks.



**Fig. 3.** Schematic diagram to determine the combination of endmember types that contribute the LST recorded by the target coarse pixel. This schematic shows an example for the first downscaling task.

belong to the same endmember type. Notably, we normalized the HLS image (Eq. (5)) prior to spectral matching operations to make the threshold comparable across geographic regions. The threshold for determining endmember types was set to 0.05 in normalized reflectance (discussed in Subsection 4.3). It is worth noting that the definition of endmembers here is different from most studies on spectral unmixing in which endmembers were defined as spectral signatures of different land cover types. In this study, endmembers represent thermal signatures of the field of view of TFPs.

$$ASD_{(i,j)} = \frac{\sum_b^{nb} |R_{b,i} - R_{b,j}|}{nb} \quad (4)$$

$$\text{Normalized Image}_b = \frac{Image_b}{maxi_b}, b \in \{1, 2, \dots, nb\} \quad (5)$$

where  $ASD_{(i,j)}$  represents the absolute spectral distance between the  $i$ th TFP and the  $j$ th TFP,  $nb$  represents the number of spectral bands in the reflectance image,  $Image_b$  represents the reflectance image of band  $b$ , and  $maxi_b$  denotes the maximum reflectance value recorded in  $Image_b$ .

By spectrally matching the reflectance signatures of TFPs, we assigned exclusive endmember combinations to each GOES-R LST pixel. As an example, for the first two downscaling tasks of UnmixGO, each TCP could have up to four endmember types ('E1', 'E2', 'E3', and 'E4', as shown in Fig. 3); for the third downscaling task, each TCP could have up to 16 types of endmembers. Particularly, the strict spectral threshold enables the differentiation, as much as possible, of different surface types with similar spectral characteristics. This allows the UnmixGO downscaled LSTs to exhibit spatial distribution characteristics close to real conditions. Yet, we need to point out that the endmember types used in this study were a simplification of the real surface heterogeneity, which would add a part of uncertainty to the downscaled LSTs.

### 3.2.2. Establishment of the set of LST mixing equations

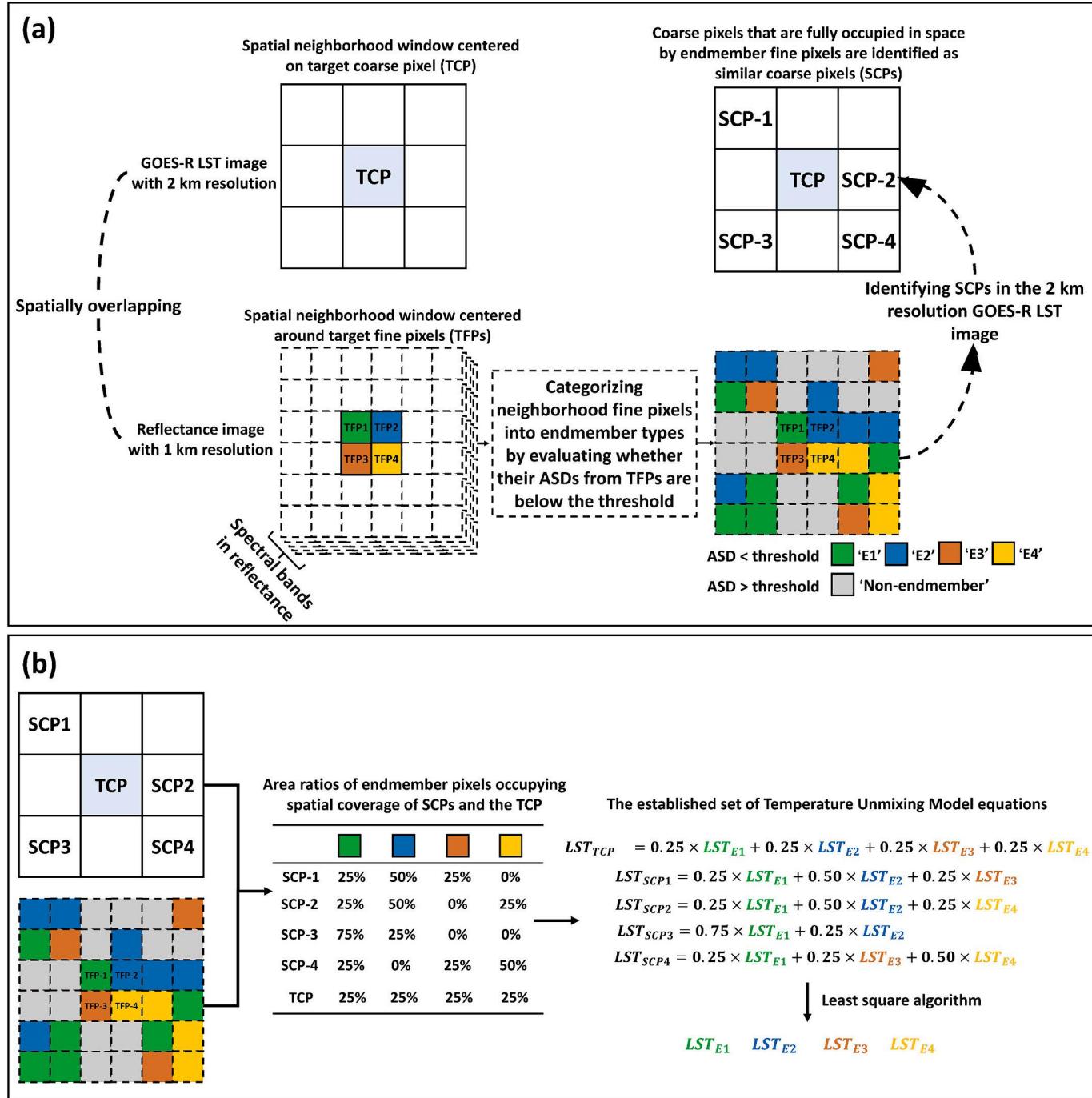
We can express the temperature recorded in each TCP with a

temperature unmixing (TUM, Eq. (3)) model, which comprises the previously defined  $k$  endmember types. Thus, downscaling the land surface temperature of a TCP to those of TFPs is equivalent to solving for endmembers from the TUM model. To this end, we need to obtain at least  $k$  TUM equations that are modeled by the same endmember types to make this problem mathematically solvable. We assumed the existence of coarse pixels (SCPs) in the geographic neighborhood of TCP, and that the surface temperature recorded by these SCPs could also be modeled by the endmembers of the TCP of interest. Determining  $k$  TUM equations is therefore equivalent to searching for  $k$  SCPs.

Taking the first downscaling task as an example (illustrated in Fig. 4), we searched for SCPs through three steps: (i) we created an image window with size  $w$  (i.e., a square window with  $2 \times w + 1$  pixels on the side) centered on the TCP of interest at 2 km resolution. Note that  $w$  was set to a value of 35 in our experiments; (ii) we calculated absolute spectral distances (ASD, Eq. (4)) of each target fine pixel (TFP) to all 1 km resolution pixels within the spatial extent of the created window; (iii) we evaluated whether these 1 km resolution pixels belong to one of endmember types (e.g., 'E1', 'E2', 'E3', and 'E4') with the calculated ASD values. Specifically, a 1 km resolution pixel would be categorized as 'E1' if its ASD value against the TFP representing endmember type 'E1' was below the predefined threshold (0.05). Likewise, the pixel would be categorized as 'unidentified' if its spectral distances from all TFPs were above the threshold. Thus, each 1 km resolution pixel within the created window would be categorized as either an endmember type or a non-endmember type (i.e., 'unidentified'). Eventually, we recognized those 2 km resolution pixels that were spatially overlaid with endmember labeled pixels as SCPs. Meanwhile, the area fractions of endmembers occupying each SCP were obtained through dividing the area of endmember labeled pixels by its spatial coverage.

In an expected case, we obtained  $n$  SCPs, which correspond to  $n$  number of independent TUM equations:

$$LST_{mix}^{n \times 1} = F_{n \times k} \times LST_{endmember}^{k \times 1} + \epsilon_{n \times 1} \quad (6)$$



**Fig. 4.** Schematic diagram for building a system of TUM equations to downscale a target coarse pixel (TCP) of interest: (a) searching for similar coarse pixels (SCPs) within the neighborhood of TCP; (b) building the corresponding set of TUM equations based on the obtained SCPs. This schematic shows an example for the first task that downscale image data from 2 km to 1 km resolution.

$$LST_{mix}^{n \times 1} = \begin{bmatrix} l_{st_{m1}} \\ l_{st_{m2}} \\ \vdots \\ l_{st_{mn}} \end{bmatrix}, F_{n \times k} = \begin{bmatrix} f_{(1,1)}, f_{(1,2)}, \dots, f_{(1,k)} \\ f_{(2,1)}, f_{(2,2)}, \dots, f_{(2,k)} \\ \dots \\ f_{(n,1)}, f_{(n,2)}, \dots, f_{(n,k)} \end{bmatrix}, LST_{endmember}^{k \times 1} = \begin{bmatrix} l_{st_{e1}} \\ l_{st_{e2}} \\ \dots \\ l_{st_{ek}} \end{bmatrix} \quad (7)$$

where  $LST_{mix}^{n \times 1}$  is an array consisting of the LST values of the SCPs we searched for,  $f_{(i,j)}$  represents the area fraction of the  $j$ th endmember occupying the spatial coverage of the  $i$ th SCP, and  $LST_{endmember}^{k \times 1}$  is an array containing  $k$  number of endmembers to be solved for.

Next, we can solve  $LST_{endmember}^{k \times 1}$  from Eq. (7) to perform the

downscaling of the TCP of interest, given that  $LST_{mix}^{n \times 1}$  and  $F_{n \times k}$  are known. When the rank of matrix  $F_{n \times k}$  in Eq. (7) was lower than the number of endmember types, we would progressively raise the threshold to re-evaluate the ASD value and expand the window size  $w$  until sufficient SCPs were searched for. The purpose of doing so is to prevent Eq. (6) from becoming an ill-conditioned system of equations, although such operations may cause a trade-off between the LST modeling accuracy and the functionality of the algorithm. Furthermore, the changes in the spectral matching threshold and the size of the neighborhood window centered on TCP may have an impact on the downscaling accuracy of UnmixGO, which was discussed in experiments introduced in Subsection

#### 4.3.

##### 3.2.3. Constraints on the solution space

The least squares algorithm was used to solve for endmembers from the established TUM equations (Eq. (6)) for each TCP. However, solving the system of equations without constraints on the solution could yield unrealistic results, such as a negative LST value in Kelvins. Therefore, we intended to create a narrowed solution space for Eq. (6) based on the predictions of a regression model driven by kernel variables including surface reflectance and seasonal LSTs, which have proven to have a good dependence on the LST at the time of interest (Agam et al., 2007; Quan et al., 2018; Sismanidis et al., 2016). In this step, we used Random Forest (RF) regression model because of its excellent performance in downscaling satellite LST images and its flexibility in using varied kernel variables (Dong et al., 2020).

There are two feasible approaches to leverage the RF model predictions to constrain the TUM solution space: (i) training the RF model only with the original 2 km resolution LST data and constraining the solution space of the TUM in all three downscaling tasks with the model's predictions; or (ii) in each downscaling task, training the RF model using the high spatial resolution LSTs generated from the previous downscaling task. In this study, we opted for the second approach because of its superior downscaling accuracy as depicted in Fig. A4.

Taking the first downscaling task of UnmixGO as an example, we trained an RF regression model at the 2 km resolution, which creates a mapping between kernels and LSTs observed by the GOES-R satellite. The training samples were divided into two parts. First, visible to near-infrared reflectance and seasonal LSTs from the HLS image and Landsat LST images acquired around the date of interest were taken as samples of kernel variables. Second, GOES-R observed LSTs were used as samples for the response variable of the RF model. In practice, the kernel images (i.e., HLS reflectance image and seasonal Landsat LST images) were aggregated to 2 km resolution. The data recorded by pixels of kernel images and the GOES-R LST image were sampled to form the training samples. Particularly, if the cloud-free MODIS LST image was available on the day of interest, its provided diurnal LST data would be used as an additional kernel for the first downscaling task (i.e., from 2 km to 1 km resolution).

Next, we used the 1 km resolution kernel images as input to the well-trained RF model to estimate endmembers (i.e., the LST values of TFPs). Note that the spatial resolution of the kernel image is 500 m and 100 m when performing the second and third downscaling tasks, respectively. However, the endmembers estimated by the RF model might be inaccurate because the model built at the 2 km spatial resolution may not well account for spatial variations in LST at higher spatial resolutions. Nevertheless, we assumed that these estimates should still be informative for determining true endmembers.

Therefore, we created a buffer centered on the RF model-estimated endmembers in the solution space of TUM equations:

$$\left\{ \begin{array}{l} width = 1.5 \times \delta \\ lst_1 - width \leq e_1 \leq lst_1 + width \\ lst_2 - width \leq e_2 \leq lst_2 + width \\ \dots \\ lst_k - width \leq e_k \leq lst_k + width \end{array} \right. \quad (8)$$

in which  $lst_k$  represents the value (in temperature) of the  $k_{th}$  estimated endmember by the RF model,  $e_k$  represents the solved  $k_{th}$  endmember from Eq. (6),  $\delta$  represents the residual in fitting kernel variables against the GOES-R LST data using RF model, and **width** represents radius of the buffer around the RF model estimated endmembers in solution space. Note that the value of  $k$  can reach up to 4 for the first two downscaling tasks, while the  $k$  value can reach up to 25 for the third one.

Additionally, we imposed another constraint (Eq. (9)) on the TUM

model's solution space based on an assumption of linear mixing of temperature (introduced in Subsection 3.1) for satellite LST images:

$$\left\{ \begin{array}{l} low \leq \sum_{i=1}^k f_i \times e_i \leq high \\ high = LST_{tcp} + width \\ low = LST_{tcp} - width \end{array} \right. \quad (9)$$

where  $LST_{tcp}$  represents the LST recorded by the target coarse pixel (TCP), and  $f_i$  represents the area fraction of the  $i_{th}$  endmember occupying the TCP in space. The other parameters remain the same as in Eq. (8).

## 4. Analysis

### 4.1. Accuracy assessment

We conducted experiments using simulated GOES-R LST images as input to LST downscaling algorithms in areas including Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City. The simulated GOES-R LST images were obtained by upscaling Landsat LST images from 100 m to 2 km resolution. Correspondingly, the original Landsat LST images were used as validation data to evaluate the accuracy of the downscaled Geo-LST images. The reason for using the simulated GOES-R LST image is that GOES-R and Landsat satellites can hardly observe the surface at the same time. NOAA releases hourly GOES-R LST products at full time, while the Landsat observation time is typically around 10:30 am local time. As a result, the observation time of Landsat typically deviates from the acquisition time of real GOES-R LST images by up to 15 to 30 min. On the other hand, we downscaled real GOES-R LST images at SURFRAD sites and used in-situ LSTs to evaluate the accuracy of the downsampled results.

The accuracy of downsampled GOES-R LST images were evaluated using metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), correlation coefficient (CC), and the difference in Robert's edge ( $\Delta Edge$ ), as suggested by our recent study on accuracy assessment of spatiotemporal data fusion algorithms (Zhu et al., 2022). Both the values of RMSE and MAE represent discrepancy between the downsampled LST values and the reference values. Low values of both metrics are thus expected. Additionally, high CC values indicate a good linear correspondence between the values of pixels in the downsampled LST image and the reference image. A  $\Delta Edge$  value close to zero is expected because it indicates good agreement of edge textures in the two images being compared. Above metrics are calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (L_i - T_i)^2}{N}} \quad (10)$$

$$MAE = \frac{\sum_{i=1}^N |L_i - T_i|}{N} \quad (11)$$

$$CC = \frac{\sum_{i=1}^N (L_i - \mu_L)(T_i - \mu_T)}{\sqrt{\sum_{i=1}^N (L_i - \mu_L)^2} \times \sqrt{\sum_{i=1}^N (T_i - \mu_T)^2}} \quad (12)$$

$$\left\{ \begin{array}{l} Edge_{(m,n)} = |L_{(m-1,n-1)} - T_{(m+1,n+1)}| \\ \Delta Edge = \frac{\sum_{i=1}^N |Edge_i^L - Edge_i^T|}{N} \end{array} \right. \quad (13)$$

In Eqs. (10) to (12),  $L_i$  and  $T_i$  represent the value of the  $i_{th}$  pixel in the downsampled LST image and in the reference image, respectively, and  $N$  represents the total number of image pixels. In Eq. (12),  $\mu_L$  and  $\mu_T$  represent the average of LSTs recorded by all pixels in the downsampled image and in the reference image, respectively. In Eq. (13),  $L_{(m,n)}$  and  $T_{(m,n)}$  represent the value of pixel at the  $m_{th}$  row and  $n_{th}$  column in the downsampled image and the reference image, respectively;  $Edge_i^L$  and

$\text{Edge}_i^T$  represent the Robert's Edge values of the  $i$ th pixel of the down-scaled image and of the reference image, respectively.

Moreover, we utilized a Taylor diagram (Fig. 8), which projects the tested LST downscaling algorithms in polar coordinates using RMSE and  $\Delta\text{Edge}$  metrics, to interpret the consistency between the down-scaled image and reference image (Zhu et al., 2022). Specifically, RMSE determines the radius of the marked algorithm from the origin of polar coordinates.  $\Delta\text{Edge}$  determines the angle between the line connecting algorithm's label to the origin and the bottom line of Taylor diagram. An angle of  $90^\circ$  corresponds to a value of 0 for  $\Delta\text{Edge}$ , indicating that edge textures of the down-scaled LST image are perfectly consistent with the reference image. As such, we interpreted those algorithms in the Taylor diagram whose labels were close to the origin of the polar coordinates as performing well.

#### 4.2. Comparing UnmixGO with state-of-the-art algorithms

Three widely used data fusion algorithms for improving the spatial resolution of Geo-LST images were used as benchmarks for comparison with UnmixGO. The algorithms being tested include the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM, Gao et al., 2006), the Spatio-Temporal Integrated Temperature Fusion Model (STITFM, Wu et al., 2015), and BLEnd Spatiotemporal Temperatures (BLEST, Quan et al., 2018). All these methods perform on the principle that changes in surface temperature are spatially autocorrelated, and they model the LST change at high spatial resolution by weighting the LST changes recorded in low spatial resolution pixels. Furthermore, STITFM and BLEST are modified versions of STARFM that use more than two types of satellite LST products to improve modeling accuracy.

The kernel regression model, another widely used methodology for downscaling Geo-LST images, was used as an additional benchmark to evaluate the performance of UnmixGO. Because UnmixGO uses the predictions of the kernel regression model to constrain the solution of the TUM model, we adopt the stepwise downscaling strategy same as UnmixGO, that is, the GOES-R LST image was sequentially down-scaled from 2 km to 1 km, 500 m, and 100 m resolution to guarantee a fair comparison. Also, the kernels fed to the RF model were consistent with the ones used in each task of UnmixGO.

#### 4.3. Factors influencing the performance of UnmixGO

UnmixGO downscales the GOES-R LST image through three consecutive tasks to address the pronounced scale gap issue (i.e., 2 km vs 100 m resolution). However, the downscaling errors generated in each task are naturally passed on to the input data of the successor task. Revealing this error propagation is helpful to understand the capability of TUM model to cope with image downscaling problems with varied scale gap issues. Therefore, we measured the accuracy of UnmixGO for each task in experiments conducted in the areas Atlanta, Cleveland, Denver, Indianapolis, Phoenix and Salt Lake City.

Three empirical parameters can also influence the performance of UnmixGO, including (i) the size of the moving window that is used to search for similar coarse pixels (SCPs), (ii) the threshold for determining endmember types, and (iii) the size of the buffer that constrains the solution space of TUM equations. Moreover, because the HLS image provides UnmixGO with the closest reference to land cover types at high spatial resolution, the use of HLS images with different acquisition times would affect the performance of the algorithm. Accordingly, we carried out four groups of GOES-R LST image downscaling experiments with changing the algorithm parameters: (i) we conducted independent LST downscaling experiments with the size of the moving window changing from 10 to 50; (ii) we progressively increased the spectral matching threshold from 0.01 to 0.1; (iii) we set the size of the solution space buffer to  $0.5\delta, 1.0\delta, \dots, 4.5\delta, 5\delta$ , respectively; (iv) we fed HLS images with varied acquisition times, respectively, to UnmixGO. We recorded the

MAE values of the down-scaled LSTs resulting from each parameter change. Note that we fixed other parameters when testing the sensitivity of the algorithm's performance to a specific one.

In addition, because rapid surface changes could complicate the modeling of high spatial resolution LSTs based on Geo-LST images, the heterogeneity of land cover change in space and time may thus influence the performance of LST downscaling (or fusion) algorithms (Zhu et al., 2010). To evaluate this effect, we derived a heterogeneity index (HI, Eq. (14)) from two seasonal HLS surface reflectance images to represent the spatiotemporal heterogeneity of land cover change over experimental regions. Specifically, we divided each reflectance image into  $n$  grids with an image window of size  $w$ , and then calculated an HI value for each image grid:

$$\text{HI} = \sum_{\text{bandi}}^{nb} \text{stddev} \left( \text{Ref}_{\text{bandi}}^{\text{daybase1}} - \text{Ref}_{\text{bandi}}^{\text{daybase2}} \right) \quad (14)$$

where  $\text{Ref}_{\text{bandi}}^{\text{daybase1}}$  and  $\text{Ref}_{\text{bandi}}^{\text{daybase2}}$  represent the  $i$ th band of the two seasonal HLS reflectance images, respectively, for each image grid;  $\text{stddev}$  represents the calculation of the standard deviation value for the values of pixels within the image grid;  $nb$  is the total number of spectral bands of the reflectance image. Meanwhile, the mean absolute error (MAE) of the down-scaled LST image within each grid area was also calculated. Thus, we could obtain one sample consisting of one MAE value and one HI value from each grid. By performing these calculations for all image grids, we acquired samples of the LST downscaling accuracy (in MAE) for each tested algorithm in response to spatiotemporal heterogeneity of surface changes (HI). The sensitivity of MAE to varied levels of HI values can reveal the robustness of the LST downscaling algorithm when applied to regions where rapid land cover changes occur.

The above-mentioned experiments were performed in areas including Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City. The simulated GOES-R LST image (2 km resolution) and Landsat LST image (100 m resolution) were used as input data of LST downscaling methods and the validation data for the results, respectively.

#### 4.4. Contribution of innovative modules in UnmixGO to the LST downscaling accuracy

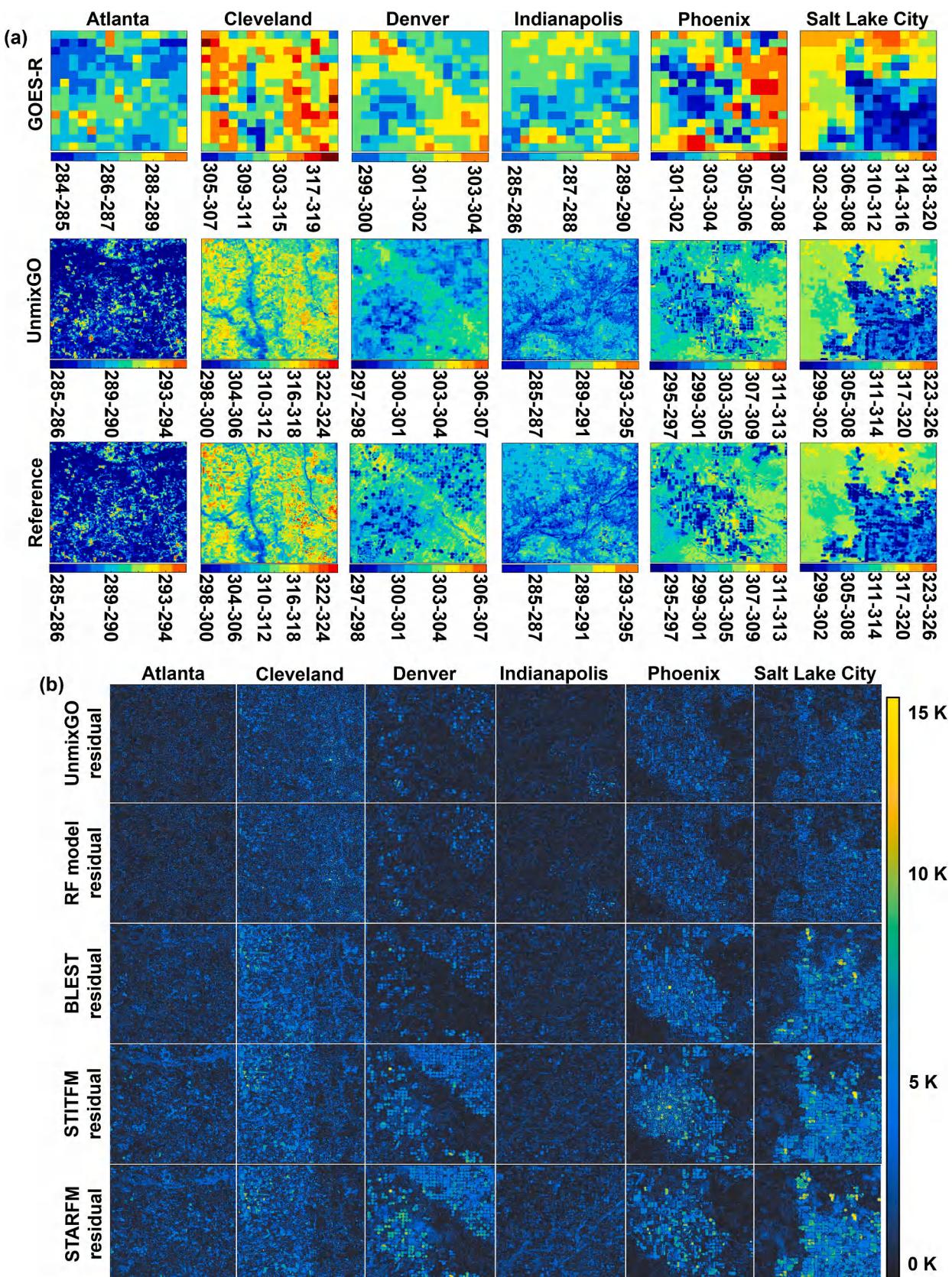
UnmixGO makes innovative modifications to conventional TUM-based image downscaling algorithms. These include a stepwise downscaling strategy to address the scale gap problem, determining spatially adaptive endmember types for each TCP, and constraining the solution space of the TUM model using predictions of the kernel regression model.

We did the ablation experiment to quantify the contribution of these innovations to the TUM methodology. Specifically, we sequentially replaced the new methods in the UnmixGO framework with the conventional approaches and evaluated the resulting accuracy of down-scaled LSTs. First, we down-scaled the GOES-R LST image directly from 2 km to 100 m resolution instead of performing a stepwise unmixing procedure. Second, we determined the endmember types for each TCP using a k-means clustering approach with a  $k$  value of 500. Third, we restricted the solutions of the system of TUM equations to positive values.

## 5. Results and discussion

### 5.1. Experiments with simulated GOES-R LST images

In Fig. 5a, the spatial distribution of UnmixGO down-scaled LSTs exhibits to be generally in agreement with the reference image, except for the results in Denver and Salt Lake City areas showing a bit patchy effect. Moreover, Fig. 5b shows that UnmixGO, Random Forest (RF) model, and BLEST produced LST images with lower bias than STITFM and STARFM. Yet, we noticed that all tested algorithms performed



**Fig. 5.** Visual inspection of downscaled results from experiments conducted in areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City: (a) the images in each row represent the simulated GOES-R LST images (2 km resolution), UnmixGO downscaled images (100 m resolution), and the 100 m resolution Landsat LST images (reference images), respectively. The color bars indicate the surface temperature in Kelvin; (b) the spatial distributions of the absolute bias in the downscaled LST images by UnmixGO, Random Forest (RF), BLEST, STITFM, and STARFM, respectively. The color bars indicate the absolute bias in the downscaled LST images.

relatively poorly over heterogeneous areas. This pronounced error may be due to a violation of the assumptions underlying these algorithms by surface heterogeneity. For instance, UnmixGO could have trouble searching for sufficient similar coarse pixels in heterogeneous areas. The RF regression model established at macrospatial scales might not account for surface temperature variations at high spatial resolution well (Dong et al., 2020). For BLEST, STITFM, and STARFM, the temporal changes in LST observed by low spatial resolution pixels may not adequately represent the changes recorded in high spatial resolution pixels (Zhu et al., 2022).

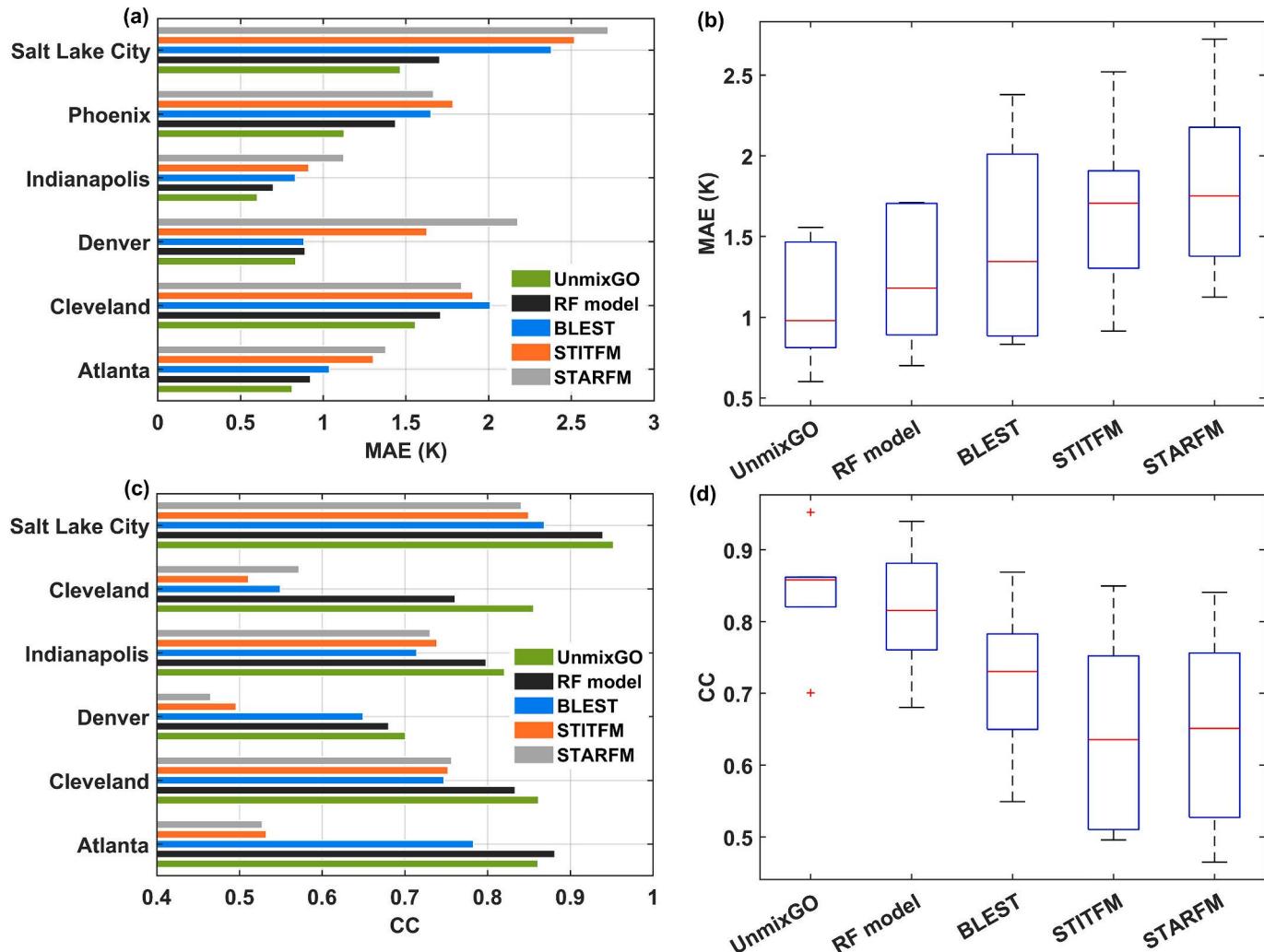
The accuracy assessments (Fig. 6) corroborate these visual interpretations. The downscaled LSTs by UnmixGO exhibit the mean absolute error (MAE) ranging from 0.6 to 1.47 K and the Pearson's correlation coefficient (CC) ranging from 0.7 to 0.95 compared to the reference data. In contrast, other methods yielded LSTs with MAE varying from 0.7 to 2.6 K, and their CC values are generally below 0.9. Additionally, the outcomes of UnmixGO show lower systematic errors across a range of temperatures than other alternatives, as demonstrated in Fig. 7, where the UnmixGO fit line aligns better with the ideal line. These excellent accuracy metrics further point to the consistent image textures of the UnmixGO downscaling results with the reference image, as evidenced by the proximity of the UnmixGO label to the Taylor diagram origin (Fig. 8).

## 5.2. Experiments with in-situ LSTs as validation data

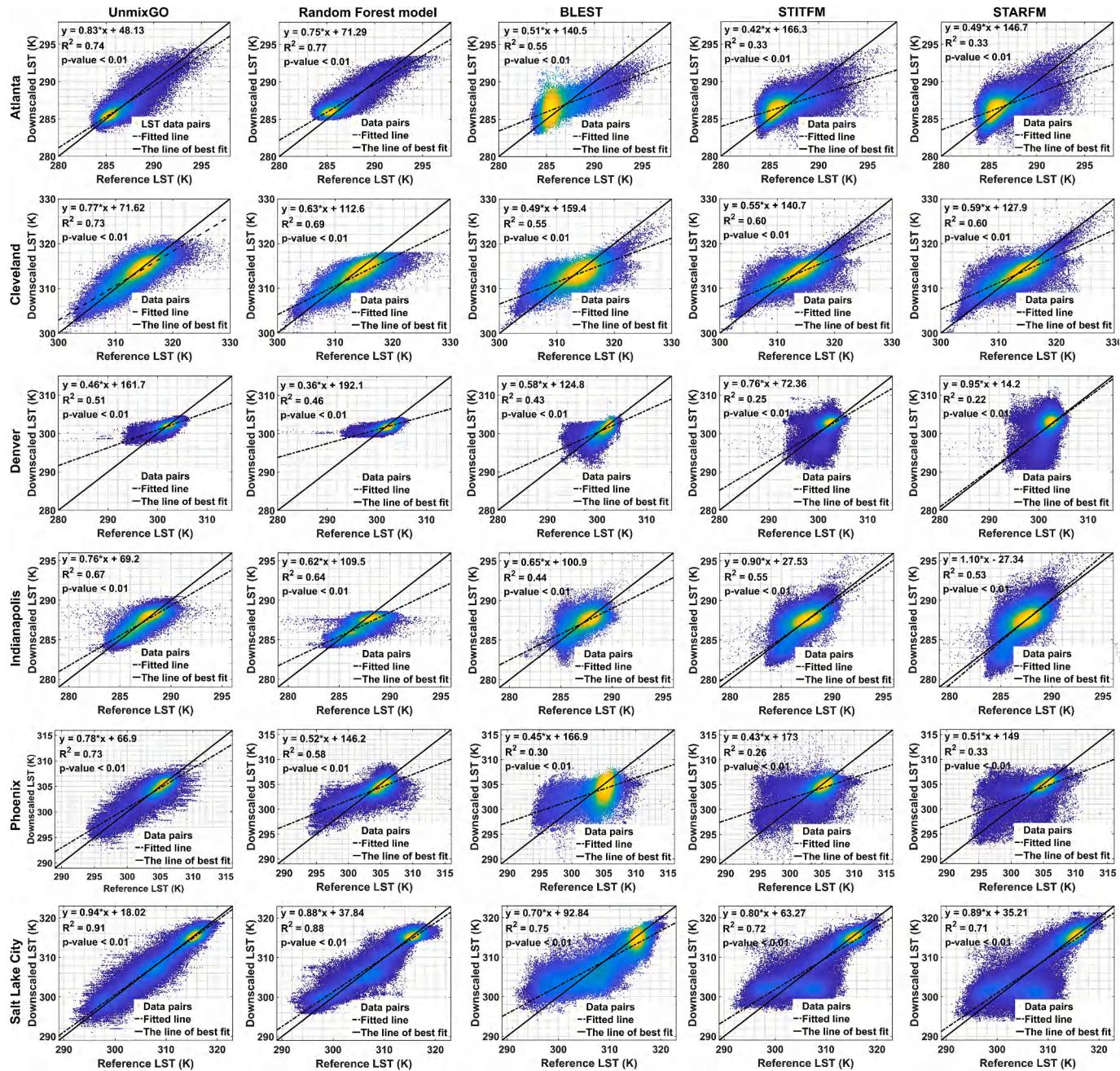
The downscaled LSTs from real GOES-R LST images by all tested methods reveal temporal trends that are generally consistent with the measurements from SURFRAD sites (Fig. 9a to e). However, accuracy metrics of the downscaled results for SURFRAD sites are generally poorer than those for the Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City experimental areas. The average MAE of UnmixGO downscaled results reaches up to 2.98 K, while the values for the RF model, STARFM, STITFM, and BLEST range from 3.21 to 4.91 K (summarized from Fig. 9f).

This high downscaling error may be due to the inconsistency of the observational scales between ground and spaceborne sensors. Specifically, the field-of-view (FOV) of the pyrgeometer installed at SURFRAD sites is a circular area with a diameter of about 30 to 45 m (Wang and Liang, 2009), while the FOV of the downscaled LST image pixel represents a square area of 10,000 m<sup>2</sup>. Moreover, the broadband emissivity used to derive in-situ LSTs holds a 1 km spatial resolution inherited from MODIS narrow band emissivity products (Eq. (2)). Thus, the emissivity used in this study might not accurately represent the thermal conditions of the area observed by pyrgeometers at SURFRAD sites, which consequently causes additional uncertainties in the modeled LSTs.

Fig. 9 illustrates that the RF model and UnmixGO demonstrated



**Fig. 6.** Quantitative evaluation of performances of LST downscaling methods. Figures (a) and (c) demonstrate Mean Absolute Error (MAE) and Pearson's Correlation Coefficient (CC) for the results of five LST downscaling methods in experimental areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City, respectively. Figures (b) and (d) summarize the statistics in figures (a) and (c), respectively. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference.



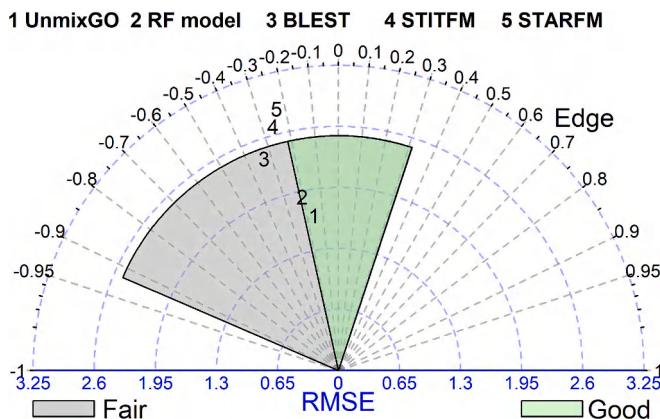
**Fig. 7.** Scatter plots of downscaled LSTs by the tested methods (columns) against reference values for six experimental areas (rows). The black dashed line fits the reference LST value to the downscaled LST value, while the black solid line represents an ideal case where downscaled data is in perfect agreement with the reference data. The  $R^2$  represents the determination coefficient and p-values indicate the probability that the good fitting between the modeled LSTs and the reference ones is due to chance. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference.

comparable accuracy metrics in the downscaling of real GOES-R LST images. The good performance of RF model is attributed primarily to the homogeneity of land surface, which mitigates the effect of the ‘scale gap issue’ on the downscaling accuracy. Specifically, the RF model’s training samples represent a different spatial scale compared to the LSTs it predicts. The SURFRAD sites, at which the experiments were executed, feature an exceptionally uniform surface (Duan et al., 2019), allowing the low spatial resolution LST samples to be representative of the LST variations at finer spatial scales. This principle also explains the substantial downscaling errors of the RF model in heterogeneous areas, as we discussed earlier in Subsection 5.1. Moreover, the RF model outperformed UnmixGO in terms of efficiency and secured the second-

highest downscaling accuracy across all trials, as depicted in Figs. 6 and 9. When considering the trade-off between computational speed and accuracy, the RF model could be a good alternative to UnmixGO in LST downscaling tasks in homogeneous landscapes, such as expansive grassland and forests.

### 5.3. Factors affecting the performance of UnmixGO

The primary source of uncertainty in UnmixGO downscaled LSTs may arise from cumulative errors generated in the three consecutive downscaling tasks. As shown in Fig. 10, the first task contributes 34% to 53% of the downscaling error, the second task contributes 19% to 24%,



**Fig. 8.** The Taylor diagram for evaluating the agreement of downscaled images by all tested methods with reference data in texture information and LST value. The information in this chart was derived from experimental results for the areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference.

and the third task generates 25% to 44% of the error. The underlying principle of this error propagation is that the LST to be downscaled inherits the errors produced in preliminary downscaling tasks, which are then attributed to the solved endmembers through the TUM model (Eq. (3)). These error propagations suggest that caution should be exercised when using UnmixGO to downscale LST images with ultra-coarse spatial resolutions such as the SEVIRI LST image of 3 km spatial resolution (Dumitrescu et al., 2020) and the FY-4 AGRI image of 4 km resolution (Li et al., 2023a).

The empirical parameters to make the algorithm operational were found to influence the accuracy of UnmixGO. We observed that the downscaling error of UnmixGO increases with the growth of the spectral matching threshold and the extent of the solution space (Fig. 11a and c), aligning with our expectations. For instance, a relaxed spectral matching threshold could oversimplify the surface heterogeneity of the experimental area, thereby weakening the performance of the TUM model in fitting the coarse pixel's LST. In a widening solution space, the TUM model may struggle to determine an optimal endmember combination. Interestingly, we found that changes in the size of the image window for determining similar coarse pixels (SCPs) have minimal impact on the downscaling accuracy of UnmixGO, especially when the window size is small (Fig. 11b). This may be explained by the algorithm's adaptive expansion of the window size when it failed to search for sufficient SCPs within the current window. A similar adaptive setting was applied to the spectral matching threshold.

Furthermore, the use of HLS images has been identified as a key factor influencing the performance of UnmixGO. Our observations from Fig. 11d indicate that the downscaling accuracy of UnmixGO improved when the acquisition time of the HLS reflectance image closely aligned with the GOES-R LST image. This phenomenon can be attributed to the high temporal resolution (2.5 days on average) of HLS images, which enables timely tracking of land cover changes (Claverie et al., 2018). This, in turn, allows the endmembers used by UnmixGO to accurately represent the up-to-date land surface types in the study area. In contrast, spatiotemporal LST fusion methods such as STARFM and its derivative methods (i.e., STITFM and BLEST) derive detailed surface change information solely from Landsat LST images with a lower temporal resolution of 8 to 16 days (Weng and Fu, 2014). The results of these methods may not accurately represent the true spatial distribution of LSTs in the study area due to limited availability of Landsat images caused by cloudy and rainy weather conditions. This issue becomes more pronounced in areas experiencing rapid surface changes. As demonstrated

in Fig. 12, the downscaling errors of BLEST, STITFM, and STARFM increased by 4.7 K, 6.3 K, and 6.53 K, respectively, as the heterogeneity index changed from 0 to 0.6. In comparison, the error increment for UnmixGO was 2.4 K.

#### 5.4. Unique contribution of each UnmixGO component to the LST downscaling accuracy

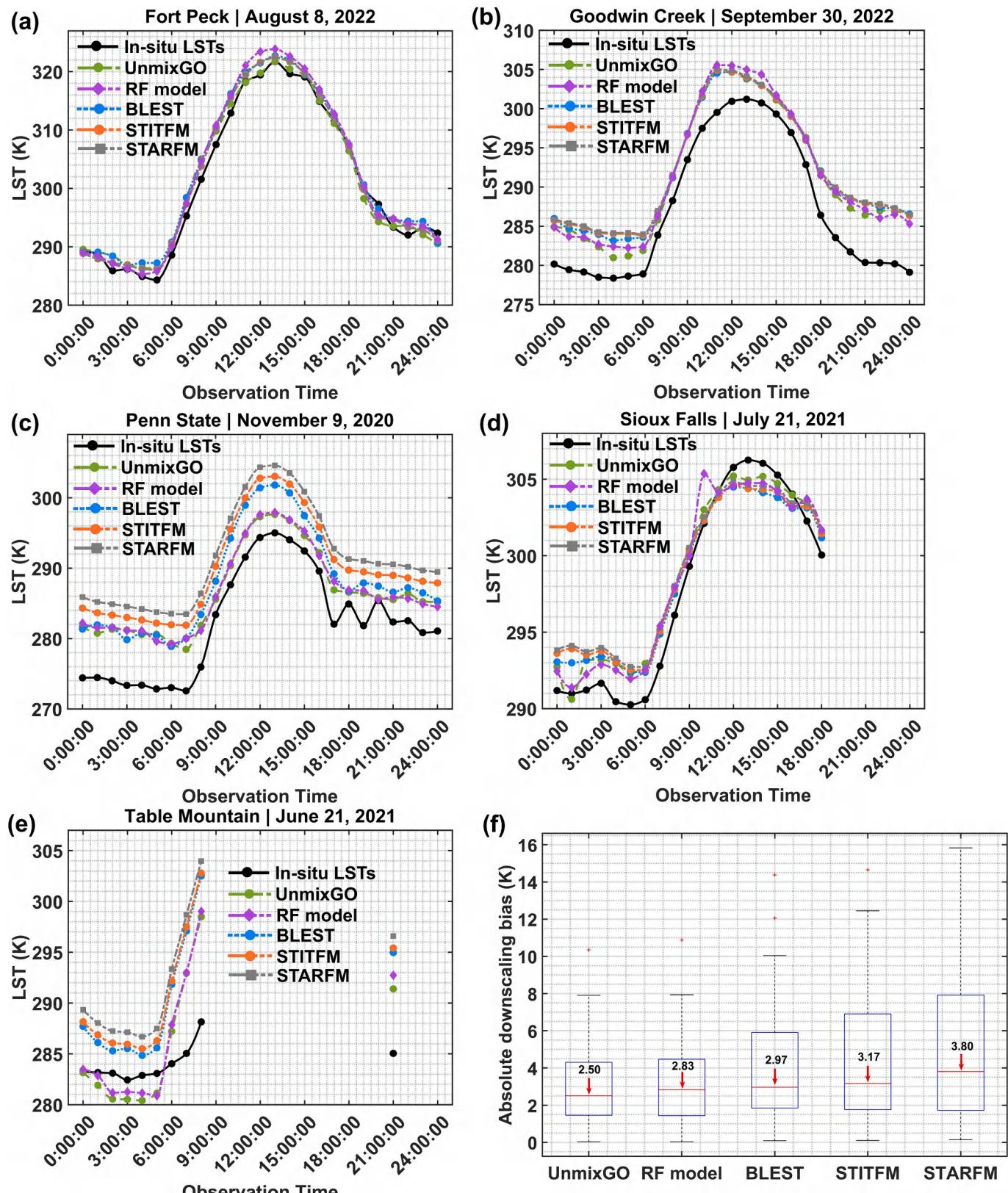
The UnmixGO model proposed in this study introduces three modifications to conventional temperature unmixing techniques, including (i) the adoption of a stepwise downscaling strategy, (ii) the determination of spatially adaptive endmember types, and (iii) the utilization of a kernel regression model to constrain the solution space of the system of TUM equations. These modifications have been demonstrated to effectively address the scale gap issue and optimize LST downscaling accuracy (Fig. 13). Specifically, the stepwise downscaling strategy resulted in a downscaling error 0.38 K lower than that obtained from the direct downscaling strategy. The second and third innovative modules featured by UnmixGO improve the LST downscaling accuracy by an average of 0.28 K and 1.11 K, respectively. Additionally, the integration of the kernel regression model and temperature unmixing model was found to yield an average improvement in accuracy of 13.17% compared to the kernel regression model (Fig. 6). Furthermore, the flexibility of the kernel regression model with respect to explanatory variables (Hutengs and Vohland, 2016) highlights the added value of UnmixGO in a future where multi-source satellite images such as hyperspectral images are increasingly accessible (Burke et al., 2021; Storch et al., 2023).

#### 5.5. Limitations and future work

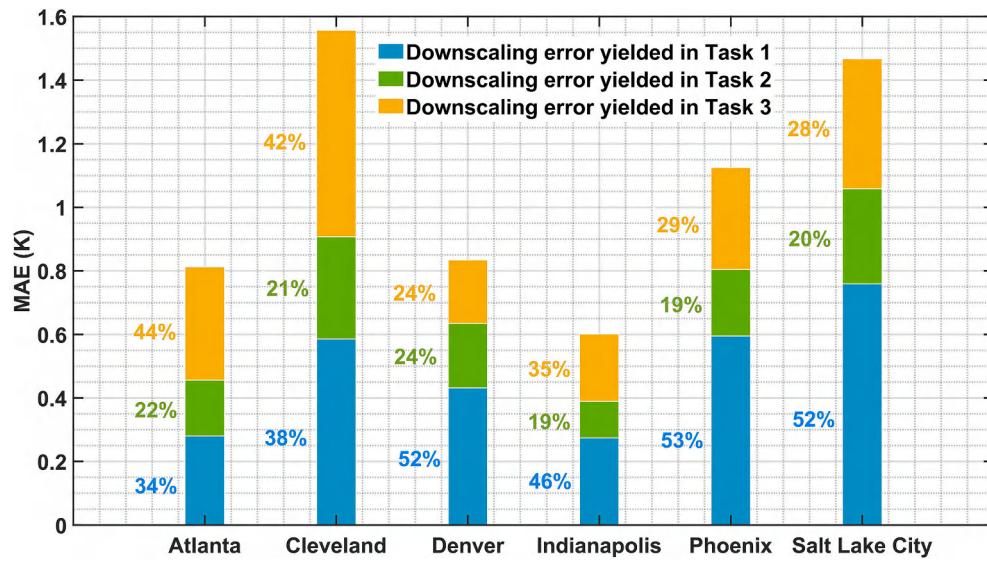
UnmixGO exhibits two limitations that may hinder its practical application. Firstly, the pixel-by-pixel downscaling approach of UnmixGO is computationally inefficient, involving repetitive operations such as determining endmember types, searching for similar coarse pixels, and solving for the system of TUM equations. This issue could be addressed by implementing parallel computations on supercomputers (Gao et al., 2021). Secondly, the linear TUM model used by UnmixGO does not accommodate thermal emissions from the vertical surfaces of ground objects within the satellite's observational footprint (Hu et al., 2016; McCabe et al., 2008; Zhan et al., 2013). Consequently, UnmixGO's downscaled LSTs may have additional uncertainties over areas with complex vertical structures (e.g., high-rise buildings), despite the proven robustness of the linear TUM model in bridging LSTs of different spatial scales at the cost of a slight drop in modeling accuracy (Danilina et al., 2012; Deng and Wu, 2013; Pu and Bonafoni, 2021).

Nevertheless, we can anticipate the added value of UnmixGO to studies on ecological thermal environments. For example, in Fig. A3 (see Appendix), UnmixGO is utilized to reveal seasonal surface temperature characteristics in the Dallas-Fort Worth metropolitan area located in Texas. The downscaled LST images uncover diurnal temperature signatures of different land surface types at high spatial resolution, which are challenging to interpret in the original GOES-R images. These exemplary results suggest that producing high spatiotemporal LST products at larger spatial scales (e.g., the continental United States) will offer new opportunities to understand the mechanisms by which ecosystems respond to global climate change.

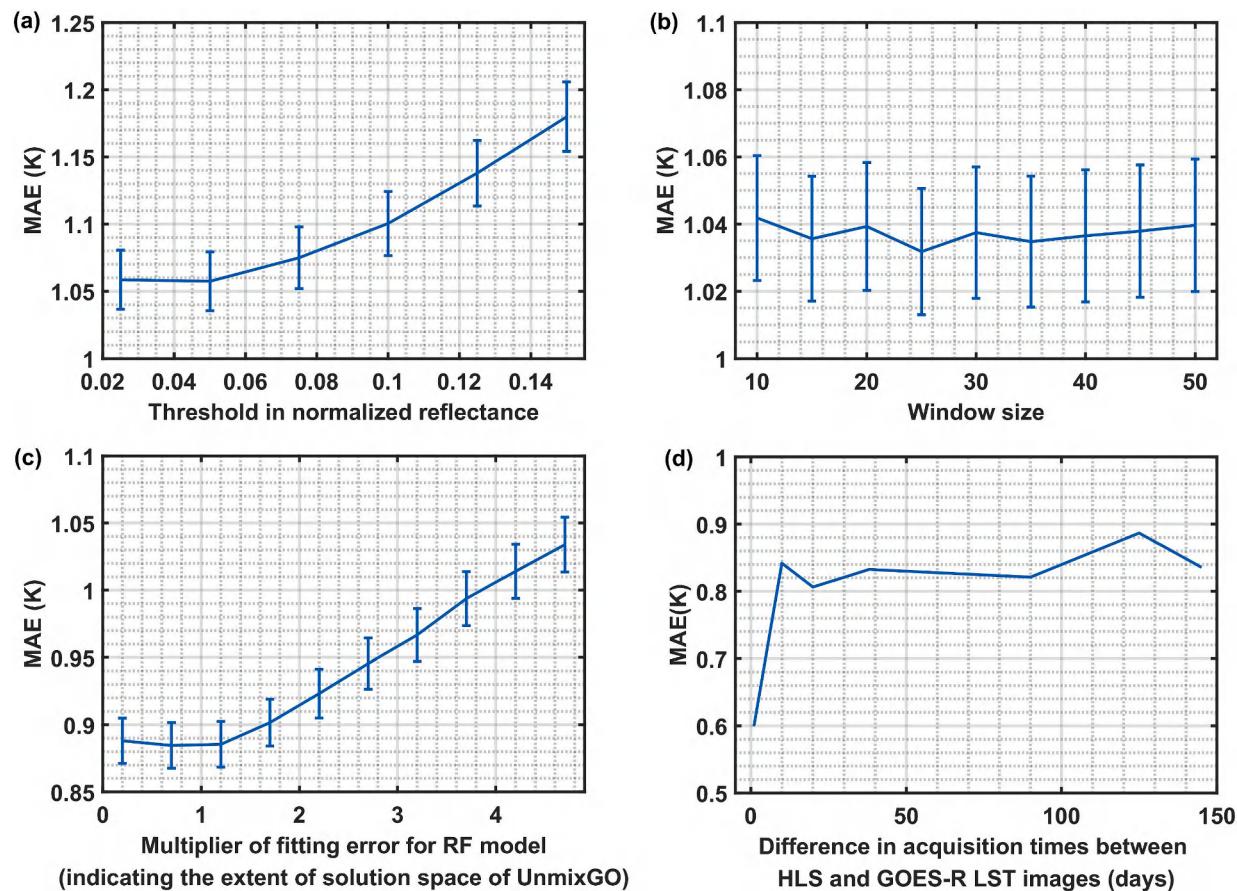
Furthermore, future research could be focused on examining the impact of varying satellite observation angles on LST downscaling accuracy. This issue should exist in our study, as none of the experimental areas fall within the nadir view of the GOES-R satellite, resulting in a spatial resolution of the used GOES-R images lower than their nominal resolution. However, our analysis did not reveal a significant effect of the viewing zenith angle (ranging from 43° to 63°) on the downscaled LSTs due to the limited number of samples available from the five SURFRAD sites. The scarcity of a validation dataset has hindered previous studies from addressing this issue (Quan et al., 2018; Weng and Fu,



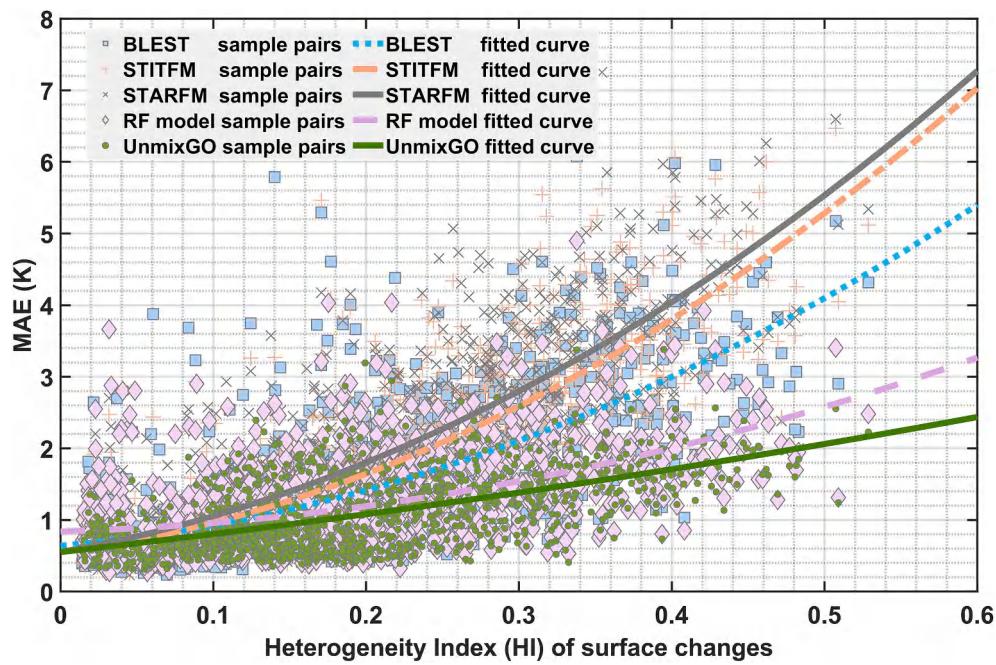
**Fig. 9.** (a) to (e): line graphs showing the field measured LSTs and the downscaled ones at each SURFRAD site. The missing data was due to the cloud contamination on GOES-R images and the curves were obtained through interpolating the discrete LST records; (f) boxplots demonstrating the MAEs for the downscaled results by all alternative methods with in-situ LSTs as reference data.



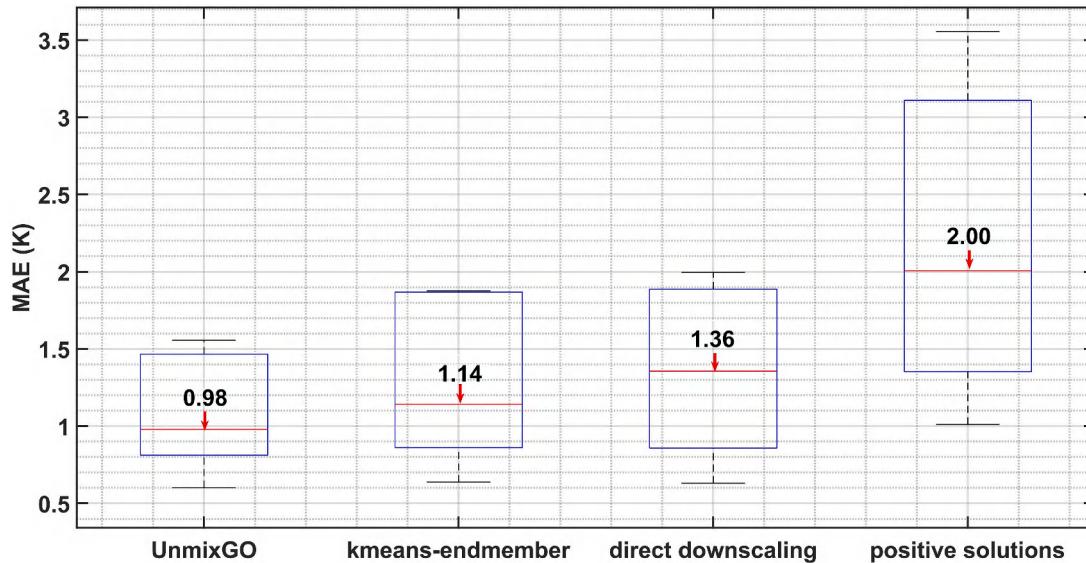
**Fig. 10.** The stacked bar chart demonstrating the MAE of UnmixGO downscaled LSTs yielded in the three consecutive downscaling tasks. The numbers annotated in the figure represent the error produced by UnmixGO in each downscaling task in the percentage of the total error. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference.



**Fig. 11.** Robustness test of UnmixGO's performance, measured in MAE, against varied parameter settings: (a) the accuracy of UnmixGO in response to the varying spectral matching thresholds; (b) the accuracy of UnmixGO in response to the varying size of the image window used to search for similar coarse pixels; (c) the accuracy of UnmixGO in response to the varying size of the solution space (i.e., the multiplier in front of  $\delta$  in Eq. (8)); (d) the accuracy of UnmixGO in response to the gap in acquisition times between the HLS and the GOES-R LST images. The error bars in figures (a) to (c) represent 10% of the standard deviation of the MAEs for all experimental regions. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference. Data in figure (d) were derived from the experimental results for the Indianapolis area.



**Fig. 12.** The accuracy of LST downscaling algorithms (in MAE) in response to the heterogeneity index (Eq. (14)) of surface changes in the experimental areas. The statistics are based on the accuracy assessment of the downscaled results from simulated GOES-R LST images, using the 100 m Landsat LST images as reference.



**Fig. 13.** The downscaling accuracy (measured in MAE) resulting from each component of UnmixGO being replaced by a conventional approach. ‘kmeans-endmember’ represents the use of Kmeans to determine the endmember types for each target coarse pixel; ‘direct downscaling’ represents a strategy for directly downscaling GOES-R LST images from 2 km to 100 m resolution; ‘positive solutions’ represents that the solved endmembers from the system of TUM equations are constrained to be positive.

2014; Wu et al., 2015). Encouragingly, recent work on correcting for angular effects in thermal infrared images shows promise for addressing this challenge (Teng et al., 2023).

## 6. Conclusion

In this paper, we introduced ‘UnmixGO’, which downscale surface temperature observations from geostationary satellites to obtain high spatiotemporal LST image data. The successful implementations of UnmixGO on both simulated and real GOES-R LST images across the contiguous United States demonstrates its robust performance. Moreover, this study makes the unique contributions to the temperature

unmixing methodology by: (i) validating the feasibility of a stepwise downscaling strategy to address the scale gap issue in the downscaling of geostationary satellite LST images; (ii) determining spatially adaptive endmember combinations for each target coarse pixel; and (iii) constraining the solution space for the system of temperature unmixing equations using predictions from a kernel regression model. Each of these innovative operations has been shown to help obtain optimized LST downscaling accuracy. Furthermore, the framework of UnmixGO allows for the utilization of multi-source image data to downscale the geostationary satellite LST images, thus meeting the opportunities brought by the continued growing accessibility of satellite images in the coming future.

## CRediT authorship contribution statement

**Fei Xu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xiaolin Zhu:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization. **Jin Chen:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis, Conceptualization. **Wenfeng Zhan:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

## Data availability

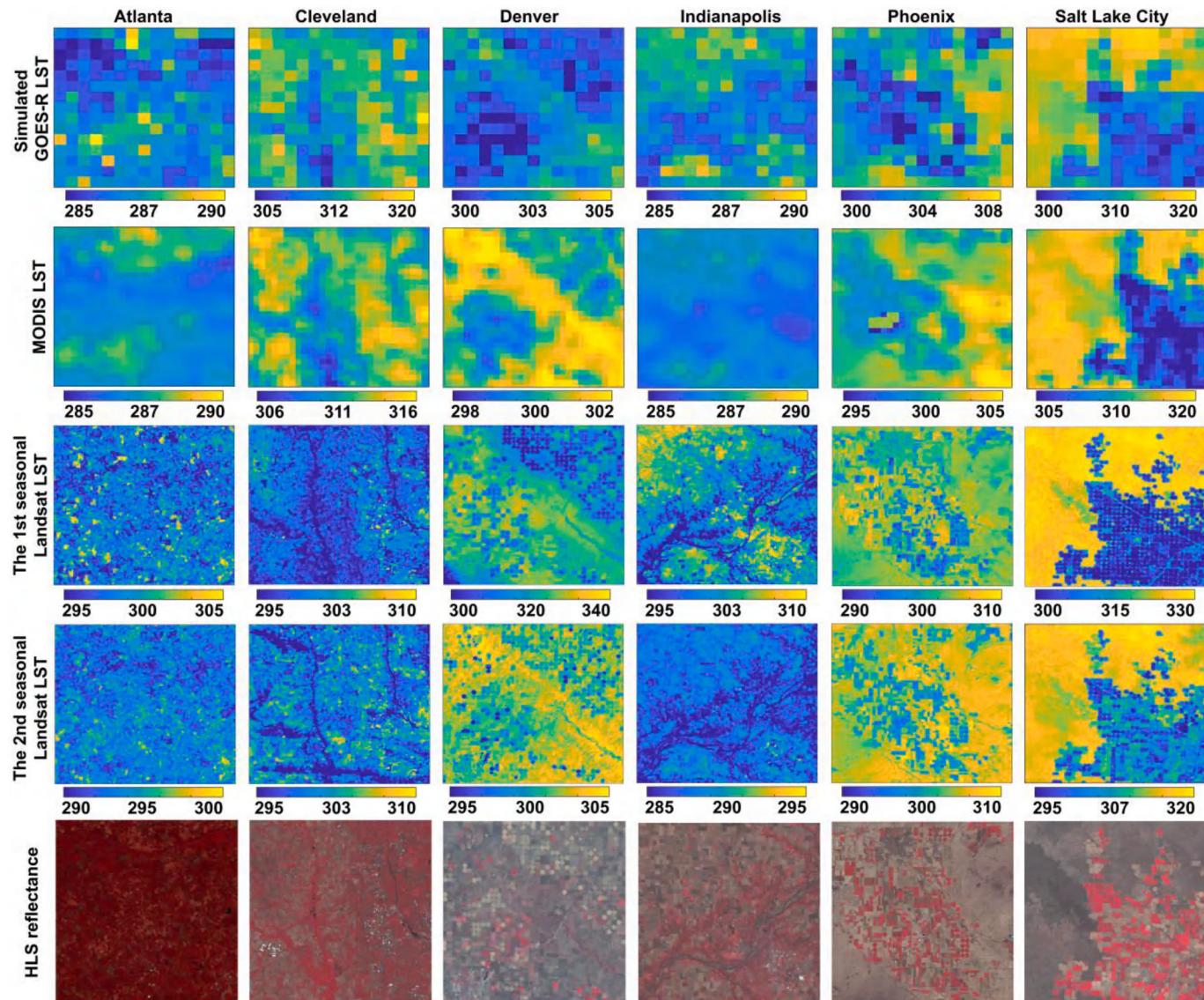
Data will be made available on request.

## Acknowledgments

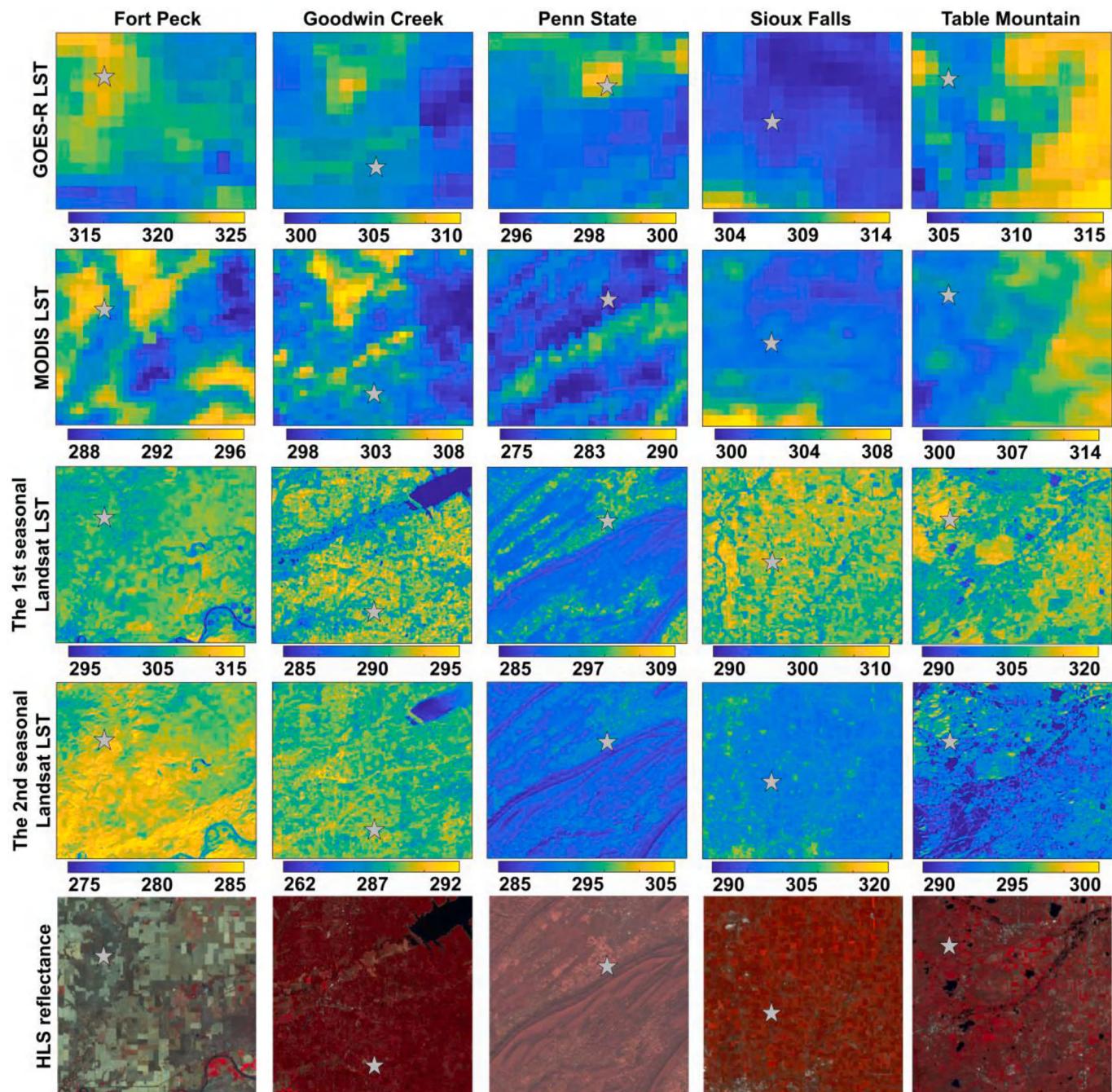
We sincerely thank the editors and anonymous reviewers for their suggestions that helped us improve the quality of the paper. This study was supported by the Guangdong Basic and Applied Basic Research Foundation (No. 2022B1515130001), National Natural Science Foundation of China (project No. 42271331), the Research Grants Council of Hong Kong (project No.15232923), and the Hong Kong Polytechnic University (project Nos. Q-CDBP and 4-ZZND).

## Appendix

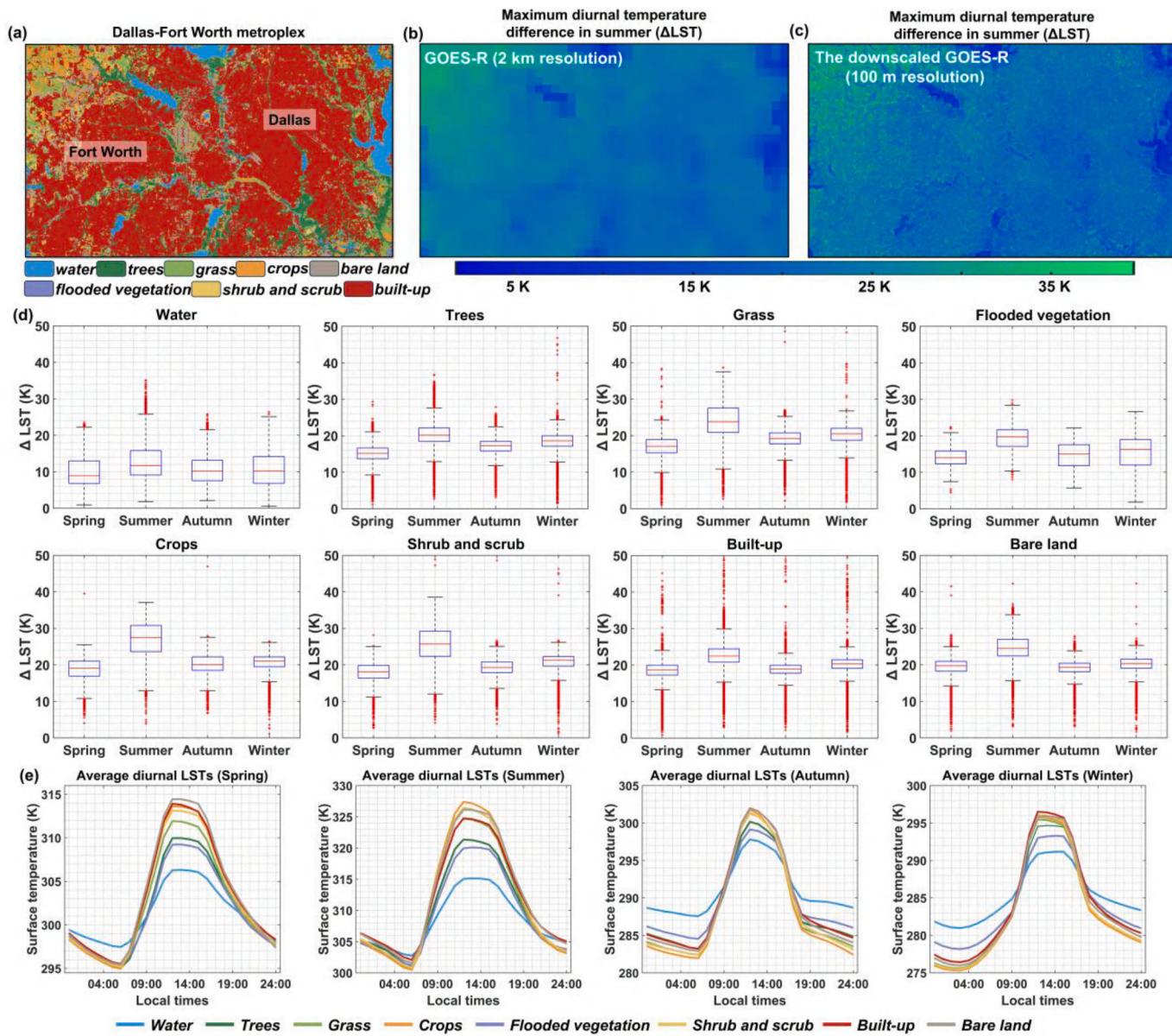
The demo codes and data for UnmixGO can be reached via both Google drive ([https://drive.google.com/drive/folders/17aG7RGMThqlrRR1v\\_G\\_Edtzs2WMyHGGI?usp=drive\\_link](https://drive.google.com/drive/folders/17aG7RGMThqlrRR1v_G_Edtzs2WMyHGGI?usp=drive_link)).



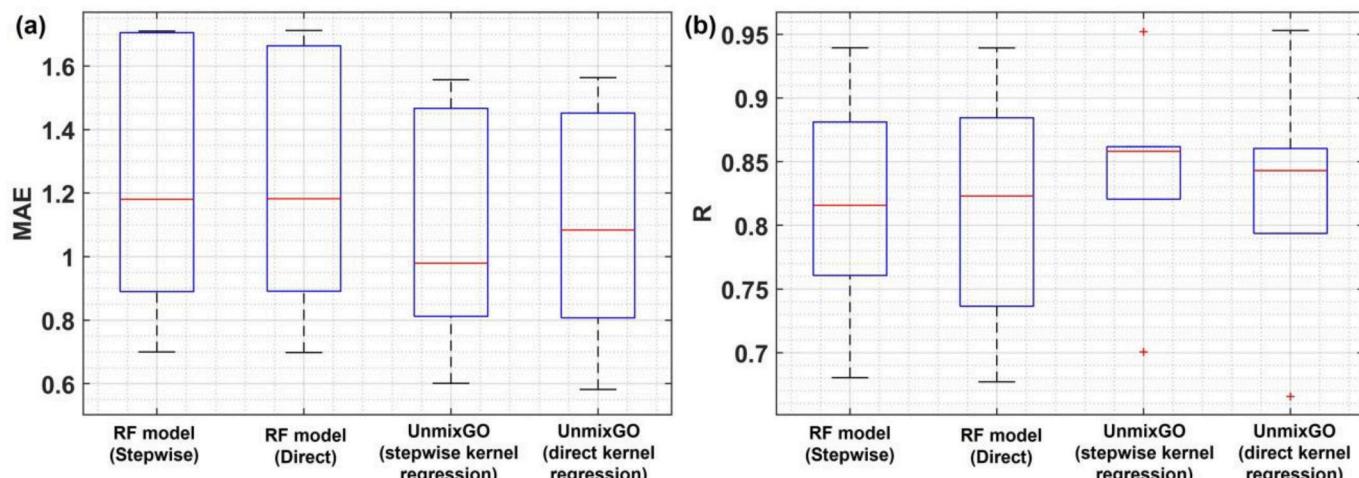
**Fig. A1.** The image data with 300 × 300 pixels used in the experimental areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City. Note that the color bars indicate the surface temperatures in Kelvin and the HLS images were rendered in the standard false color composite (R: 842 nm, G: 665 nm, B: 560 nm).



**Fig. A2.** The image data with  $300 \times 300$  pixels used in experiments conducted at SURFRAD sites of Fort Peck, Goodwin Creek, Penn State, Sioux Falls, and Table Mountain. Note that the color bars indicate the surface temperatures in Kelvin. The HLS images are rendered in standard false color composite. The stars in grey indicate the locations of SURFRAD sites.



**Fig. A3.** Results obtained by applying UnmixGO to GOES-R LST images covering the Dallas-Fort Worth metroplex area: (a) Spatial distributions of land cover types in the Dallas-Fort Worth metropolitan area at 100 m resolution obtained from the Dynamic World Atlas; (b) and (c) Two maps of the maximum diurnal temperature difference ( $\Delta LST$ ) for a day in summer obtained from the original GOES-R LST images and the downscaled GOES-R LST images by UnmixGO, respectively; (d) Box plots showing the  $\Delta LST$  of different land cover types in spring, summer, autumn, and winter; (e) Line graphs showing the average diurnal LSTs (from 0:00 to 24:00) of each land cover type for a day in spring, summer, fall, and winter, respectively.



**Fig. A4.** Boxplots showing the accuracy metrics (MAE and R) of the RF models operated with the 'stepwise' and 'direct' downscaling strategies (the first two boxplots in each subplot), as well as the metrics of UnmixGO integrated with the 'stepwise' and 'direct' kernel regression models (the final two boxplots in each subplot). The statistics were derived from the accuracy assessment of the downscaled results from simulated GOES-R images covering the areas of Atlanta, Cleveland, Denver, Indianapolis, Phoenix, and Salt Lake City.

## References

- Adams, J.B., Smith, M.O., Johnson, P.E., 1986. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site. *J. Geophys. Res. Solid Earth* 91 (B8), 8098–8112.
- Agam, N., Kustas, W.P., Anderson, M.C., Li, F., Neale, C.M., 2007. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sens. Environ.* 107 (4), 545–558.
- Augustine, J.A., Hodges, G.B., Cornwall, C.R., Michalsky, J.J., Medina, C.I., 2005. An update on SURFRAD—the GCOS surface radiation budget network for the continental United States. *J. Atmos. Ocean. Technol.* 22 (10), 1460–1472.
- Bechtel, B., Zaksék, K., Hoshyarpour, G., 2012. Downscaling land surface temperature in an urban area: a case study for Hamburg, Germany. *Remote Sens.* 4 (10), 3184–3200.
- Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci. Data* 5 (1), 1–12.
- Bindhu, V.M., Narasimhan, B., Sudheer, K.P., 2013. Development and verification of a non-linear disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale of Landsat thermal data to estimate evapotranspiration. *Remote Sens. Environ.* 135, 118–129.
- Bonafoni, S., 2016. Downscaling of Landsat and MODIS land surface temperature over the heterogeneous urban area of Milan. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9 (5), 2019–2027.
- Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Tait, A.M., 2022. Dynamic world, near real-time global 10 m land use land cover mapping. *Sci. Data* 9 (1), 251.
- Burke, M., Driscoll, A., Lobell, D.B., Ermon, S., 2021. Using satellite imagery to understand and promote sustainable development. *Science* 371 (6535), eabe8628.
- Carlson, A.R., Sibold, J.S., Negrón, J.F., 2021. Wildfire and spruce beetle outbreak have mixed effects on below-canopy temperatures in a Rocky Mountain subalpine forest. *J. Biogeogr.* 48 (1), 216–230.
- Chen, X., Chen, J., Jia, X., Somers, B., Wu, J., Coppin, P., 2011. A quantitative analysis of virtual endmembers' increased impact on the collinearity effect in spectral unmixing. *IEEE Trans. Geosci. Remote Sens.* 49 (8), 2945–2956.
- Chen, Y., Yang, Y., Pan, X., Meng, X., Hu, J., 2022. Spatiotemporal fusion network for land surface temperature based on a conditional variational autoencoder. *IEEE Trans. Geosci. Remote Sens.* 60, 1–13.
- Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Justice, C., 2018. The harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* 219, 145–161.
- Collins, E.F., Roberts, D.A., Borel, C.C., 2001. Spectral mixture analysis of simulated thermal infrared spectrometry data: an initial temperature estimate bounded TESSMA search approach. *IEEE Trans. Geosci. Remote Sens.* 39 (7), 1435–1446.
- Danilina, I., Gillespie, A.R., Balick, L.K., Mushkin, A., O'Neal, M.A., 2012. Performance of a thermal-infrared radiosity and heat-diffusion model for estimating sub-pixel radiant temperatures over the course of a day. *Remote Sens. Environ.* 124, 492–501.
- Deng, C., Wu, C., 2013. Estimating very high resolution urban surface temperature using a spectral unmixing and thermal mixing approach. *Int. J. Appl. Earth Obs. Geoinf.* 23, 155–164.
- Dennison, P.E., Charoensiri, K., Roberts, D.A., Peterson, S.H., Green, R.O., 2006. Wildfire temperature and land cover modeling using hyperspectral data. *Remote Sens. Environ.* 100 (2), 212–222.
- Dong, P., Gao, L., Zhan, W., Liu, Z., Li, J., Lai, J., Zhao, L., 2020. Global comparison of diverse scaling factors and regression models for downscaling Landsat-8 thermal data. *ISPRS J. Photogramm. Remote Sens.* 169, 44–56.
- Duan, S.B., Li, Z.L., 2016. Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: case study in northern China. *IEEE Trans. Geosci. Remote Sens.* 54 (11), 6458–6469.
- Duan, S.B., Li, Z.L., Li, H., Götsche, F.M., Wu, H., Zhao, W., Coll, C., 2019. Validation of collection 6 MODIS land surface temperature product using in situ measurements. *Remote Sens. Environ.* 225, 16–29.
- Dumitrescu, A., Brabec, M., Cheval, S., 2020. Statistical gap-filling of SEVIRI land surface temperature. *Remote Sens.* 12 (9), 1423.
- Enku, T., Melesse, A.M., 2014. A simple temperature method for the estimation of evapotranspiration. *Hydrol. Process.* 28 (6), 2945–2960.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 44 (8), 2207–2218.
- Gao, H., Zhu, X., Guan, Q., Yang, X., Yao, Y., Zeng, W., Peng, X., 2021. cuFSDAF: an enhanced flexible spatiotemporal data fusion algorithm parallelized using graphics processing units. *IEEE Trans. Geosci. Remote Sens.* 60, 1–16.
- Gevaert, C.M., García-Haro, F.J., 2015. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* 156, 34–44.
- Hu, L., Monaghan, A., Voogt, J.A., Barlage, M., 2016. A first satellite-based observational assessment of urban thermal anisotropy. *Remote Sens. Environ.* 181, 111–121.
- Huang, B., Wang, J., Song, H., Fu, D., Wong, K., 2013. Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring. *IEEE Geosci. Remote Sens. Lett.* 10 (5), 1011–1015.
- Hutengs, C., Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sens. Environ.* 178, 127–141.
- Keshava, N., Mustard, J.F., 2002. Spectral unmixing. *IEEE Signal Process. Mag.* 19 (1), 44–57.
- Kumar, D.A., 2022. Knowledge-based morphological deep transparent neural networks for remote sensing image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 15, 2209–2222.
- Lezama Valdes, L.M., Katurji, M., Meyer, H., 2021. A machine learning based downscaling approach to produce high spatio-temporal resolution land surface temperature of the Antarctic dry valleys from MODIS data. *Remote Sens.* 13 (22), 4673.
- Li, Z.L., Tang, B.H., Wu, H., Ren, H., Yan, G., Wan, Z., Sobrino, J.A., 2013. Satellite-derived land surface temperature: current status and perspectives. *Remote Sens. Environ.* 131, 14–37.
- Li, W., Ni, L., Li, Z.L., Duan, S.B., Wu, H., 2019. Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 12 (7), 2299–2307.
- Li, M., Guo, S., Chen, J., Chang, Y., Sun, L., Zhao, L., Yao, H., 2023a. Stability analysis of unmixing-based spatiotemporal fusion model: a case of land surface temperature product downscaling. *Remote Sens.* 15 (4), 901.
- Li, Y., Wu, H., Chen, H., Zhu, X., 2023b. A robust framework for resolution enhancement of land surface temperature by combining spatial downscaling and spatiotemporal fusion methods. *IEEE Trans. Geosci. Remote Sens.* 61, 1–14.
- Li, Z.L., Wu, H., Duan, S.B., Zhao, W., Ren, H., Liu, X., Zhou, C., 2023c. Satellite remote sensing of global land surface temperature: definition, methods, products, and applications. *Rev. Geophys.* 61 (1) e2022RG000777.

- Ma, J., Shen, H., Wu, P., Wu, J., Gao, M., Meng, C., 2022. Generating gapless land surface temperature with a high spatio-temporal resolution by fusing multi-source satellite-observed and model-simulated data. *Remote Sens. Environ.* 278, 113083.
- McCabe, M.F., Balick, L.K., Theiler, J., Gillespie, A.R., Mushkin, A., 2008. Linear mixing in thermal infrared temperature retrieval. *Int. J. Remote Sens.* 29 (17–18), 5047–5061.
- Mitraka, Z., Chrysoulakis, N., Doxani, G., Del Frate, F., Berger, M., 2015. Urban surface temperature time series estimation at the local scale by spatial-spectral unmixing of satellite observations. *Remote Sens.* 7 (4), 4139–4156.
- Moosavi, V., Talebi, A., Mokhtari, M.H., Shamsi, S.R.F., Niazi, Y., 2015. A wavelet-artificial intelligence fusion approach (WAIFA) for blending Landsat and MODIS surface temperature. *Remote Sens. Environ.* 169, 243–254.
- Pu, R., Bonafoni, S., 2021. Reducing scaling effect on downscaled land surface temperature maps in heterogenous urban environments. *Remote Sens.* 13 (24), 5044.
- Pu, R., Bonafoni, S., 2023. Thermal infrared remote sensing data downscaling investigations: an overview on current status and perspectives. *Remote Sens. Appl. Soc. Environ.* 100921.
- Quan, J., Zhan, W., Ma, T., Du, Y., Guo, Z., Qin, B., 2018. An integrated model for generating hourly Landsat-like land surface temperatures over heterogeneous landscapes. *Remote Sens. Environ.* 206, 403–423.
- Singh, R.K., Svystun, T., AlDahmash, B., Jönsson, A.M., Bhalerao, R.P., 2017. Photoperiod-and temperature-mediated control of phenology in trees—a molecular perspective. *New Phytol.* 213 (2), 511–524.
- Sismanidis, P., Keramitsoglou, I., Bechtel, B., Kiranoudis, C.T., 2016. Improving the downscaling of diurnal land surface temperatures using the annual cycle parameters as disaggregation kernels. *Remote Sens.* 9 (1), 23.
- Somers, B., Cools, K., Delalieux, S., Stuckens, J., Van der Zande, D., Verstraeten, W.W., Coppin, P., 2009. Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sens. Environ.* 113 (6), 1183–1193.
- Storch, T., Honold, H.P., Chabriat, S., Habermeyer, M., Tucker, P., Brell, M., Fischer, S., 2023. The EnMAP imaging spectroscopy mission towards operations. *Remote Sens. Environ.* 294, 113632.
- Teng, Y., Ren, H., Zhu, J., Jiang, C., Ye, X., Zeng, H., 2023. A practical method for angular normalization on land surface temperature using space between thermal radiance and fraction of vegetation cover. *Remote Sens. Environ.* 291, 113558.
- Valenzuela, A.Q., Reinke, K., Jones, S.D., 2023. A new methodology to assess spatial response models for satellite imagers using the optical design parameters of a generic sensor as independent variables. *IEEE Trans. Geosci. Remote* 61, 1–10.
- Wang, K., Liang, S., 2009. Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sens. Environ.* 113 (7), 1556–1565.
- Wang, Q., Shi, W., Atkinson, P.M., 2016. Area-to-point regression kriging for pan-sharpening. *ISPRS J. Photogramm. Remote Sens.* 114, 151–165.
- Wang, J.W., Chow, W.T., Wang, Y.C., 2020a. A global regression method for thermal sharpening of urban land surface temperatures from MODIS and Landsat. *Int. J. Remote Sens.* 41 (8), 2986–3009.
- Wang, J., Schmitz, O., Lu, M., Karsenberg, D., 2020b. Thermal unmixing based downscaling for fine resolution diurnal land surface temperature analysis. *ISPRS J. Photogramm. Remote Sens.* 161, 76–89.
- Wang, S., Luo, Y., Li, X., Yang, K., Liu, Q., Luo, X., Li, X., 2021a. Downscaling land surface temperature based on non-linear geographically weighted regressive model over urban areas. *Remote Sens.* 13 (8), 1580.
- Wang, H., Mao, K., Yuan, Z., Shi, J., Cao, M., Qin, Z., Tang, B., 2021b. A method for land surface temperature retrieval based on model-data-knowledge-driven and deep learning. *Remote Sens. Environ.* 265, 112665.
- Weng, Q., Fu, P., 2014. Modeling diurnal land temperature cycles over Los Angeles using downscaled GOES imagery. *ISPRS J. Photogramm. Remote Sens.* 97, 78–88.
- Wu, P., Shen, H., Zhang, L., Götsche, F.M., 2015. Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature. *Remote Sens. Environ.* 156, 169–181.
- Wu, P., Su, Y., Duan, S.B., Li, X., Yang, H., Zeng, C., Shen, H., 2022a. A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data. *Remote Sens. Environ.* 277, 113070.
- Wu, J., Xia, L., Chan, T.O., Awange, J., Zhong, B., 2022b. Downscaling land surface temperature: a framework based on geographically and temporally neural network weighted autoregressive model with spatio-temporal fused scaling factors. *ISPRS J. Photogramm. Remote Sens.* 187, 259–272.
- Xu, F., Somers, B., 2021. Unmixing-based Sentinel-2 downscaling for urban land cover mapping. *ISPRS J. Photogramm. Remote Sens.* 171, 133–154.
- Yang, G., Pu, R., Huang, W., Wang, J., Zhao, C., 2009. A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network. *IEEE Trans. Geosci. Remote Sens.* 48 (4), 2170–2178.
- Yin, Z., Wu, P., Foody, G.M., Wu, Y., Liu, Z., Du, Y., Ling, F., 2020. Spatiotemporal fusion of land surface temperature based on a convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* 59 (2), 1808–1822.
- Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Zhang, L., 2020. Deep learning in environmental remote sensing: achievements and challenges. *Remote Sens. Environ.* 241, 111716.
- Zakšek, K., Oštir, K., 2012. Downscaling land surface temperature for urban heat island diurnal cycle analysis. *Remote Sens. Environ.* 117, 114–124.
- Zhan, W., Chen, Y., Zhou, J., Li, J., 2010. An algorithm for separating soil and vegetation temperatures with sensors featuring a single thermal channel. *IEEE Trans. Geosci. Remote Sens.* 49 (5), 1796–1809.
- Zhan, W., Chen, Y., Zhou, J., Wang, J., Liu, W., Voogt, J., Li, J., 2013. Disaggregation of remotely sensed land surface temperature: literature survey, taxonomy, issues, and caveats. *Remote Sens. Environ.* 131, 119–139.
- Zhang, Q., Wang, N., Cheng, J., Xu, S., 2020. A stepwise downscaling method for generating high-resolution land surface temperature from AMSR-E data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 13, 5669–5681.
- Zhao, Y., Huang, B., Song, H., 2018. A robust adaptive spatial and temporal image fusion model for complex land surface changes. *Remote Sens. Environ.* 208, 42–62.
- Zhou, J., Chen, J., Chen, X., Zhu, X., Qiu, Y., Song, H., Cui, X., 2021. Sensitivity of six typical spatiotemporal fusion methods to different influential factors: a comparative study for a normalized difference vegetation index time series reconstruction. *Remote Sens. Environ.* 252, 112130.
- Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* 114 (11), 2610–2623.
- Zhu, X., Zhan, W., Zhou, J., Chen, X., Liang, Z., Xu, S., Chen, J., 2022. A novel framework to assess all-round performances of spatiotemporal fusion models. *Remote Sens. Environ.* 274, 113002.
- Zhukov, B., Oertel, D., Lanzl, F., Reinhackel, G., 1999. Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geosci. Remote Sens.* 37 (3), 1212–1226.