Generation of 100 m, Hourly Land Surface Temperature Based on Spatio-Temporal Fusion

Yijie Tang, Qunming Wang, Xiaohua Tong and Peter M. Atkinson

***Abstract*—Landsat surface temperature (LST) is an important physical quantity for global climate change monitoring. Over the past decades, several LST products have been produced by satellite thermal infrared (TIR) bands or land surface models (LSMs). Recent research has increased the spatio-temporal resolution of LST products to 2 km, hourly based on Geostationary Operational Environmental Satellites (GOES)-R Advanced Baseline Imager (ABI) LST data. The spatial resolution of 2 km, however, is insufficient for monitoring at the regional scale. This paper investigates the feasibility of applying spatio-temporal fusion to generate reliable 100 m, hourly LST data based on fusion of the newly released 2 km, hourly GOES-16 ABI LST and 100 m Landsat LST data. The most accurate fusion method was identified through a comparison between several popular methods. Furthermore, a comprehensive comparison was performed between fusion (with Landsat LST) involving satellite-derived LST (i.e., GOES) and model-derived LSMs (i.e., European Centre for Medium-range Weather Forecasts (ECMWF) Reanalysis *v*.5 (ERA5)-Land). The spatial and temporal adaptive reflectance fusion model (STARFM) method was demonstrated to be an appropriate method to generate 100 m, hourly data, which produced an average root mean square error (RMSE) of 2.640 K, mean absolute error (MAE) of 2.159 K and average coefficient of determination (R^2) of 0.982 referring to the *in situ* time-series. Furthermore, inheriting the advantages of direct observation, and the fusion of Landsat and GOES for the generation of 100 m, hourly LST produced greater accuracy compared to the fusion of Landsat and ERA5-Land LST in the experiments. The generated 100 m, hourly LST can provide important diurnal data with fine spatial resolution for various monitoring applications.**

***Index Terms*—Land surface temperature (LST); Landsat; GOES; ERA5; spatio-temporal fusion.**

I. INTRODUCTION

Land surface temperature (LST), as an important parameter in the energy exchange between the land surface and atmosphere, has been researched extensively in recent years [1]-[3]. LST is central to many applications including mapping the urban heat island effect [4], [5], forest fire monitoring [6], [7] and drought monitoring [8], [9]. Moreover, it is acknowledged that large-scale monitoring of LST, especially global scale monitoring, is extremely valuable for climate change research [10], [11]. Given the increasing demand for large-scale LST data, several spatially continuous LST products have been released in recent decades. Generally, these products are divided into two categories: one is derived from the thermal infrared (TIR) bands of satellite sensors (e.g., band 6 of the Landsat Enhanced Thematic Mapper (ETM+) sensor) or passive microwave (PMW) measurement, and the other is from land surface models (LSMs) (e.g., the Global Land Data Assimilation System (GLDAS) model) produced by combining observations from a variety of sources [12], [13].

For the generation of LST products, a basic requirement is the spatial and temporal continuity of the original LST data. For LST retrieved from PMW measurement, it is possible to preserve the spatial continuity as PMW radiation can penetrate clouds. The limitation of these LST products, however, is the coarse spatial resolution, coupled with the scanning gap between orbits [14]. For example, the data acquired by the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) of the Earth Observing System present noticeable gaps over middle and low latitudes and have a coarse spatial resolution of 25 km. Thus, considering the abovementioned limitations, the TIR bands of the images acquired from optical sensors can serve as a common data source for deriving LST products. For LST obtained from TIR sensors, the main problem lies in the large number of spatial gaps due to cloud contamination and sensor deficiencies [15].

Up to now, several studies focused on generating spatially complete LST products by developing gap filling algorithms and integrating complementary multi-source LST data [14]-[17]. Furthermore, research was undertaken to increase the spatial and temporal resolutions of LST products through fusion-based approaches [18]-[20]. This research holds open the possibility of spatially continuous daily LST data. LST, however, always presents strong spatial and temporal heterogeneity. In the temporal dimension, the LST can vary greatly each hour. Thus, daily LST products may fail to support applications where diurnal LST change information may be important. To meet the demand of more frequent (e.g., hourly) monitoring, it is of great necessity to develop methods to further increase the temporal resolution of LST products.

Recently, a 2 km, all-sky, hourly LST product was generated by Jia *et al.* [21]. The Geostationary Operational Environmental Satellites (GOES)-R, which is operated by the National Oceanic and Atmospheric Administration (NOAA), can provide global diurnal observations [22]. The LST product retrieved from GOES-16 Advanced Baseline Imager (ABI) can provide 10 km, hourly LST over North and South America, and 2 km, hourly

Manuscript received XXX; revised XXX; accepted XXX. This research was supported by the National Natural Science Foundation of China under Grants 42222108, 42221002 and 42171345. *(Corresponding author: Q. Wang.)*

Y. Tang and X. Tong are with the College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China.

Q. Wang is with the College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai 200092, China, and also with the Key Laboratory of Ethnic Language Intelligent Analysis and Security Governance of MOE, Minzu University of China, Beijing 100081, China (e-mail: wqm11111@126.com).

P.M. Atkinson is with the Faculty of Science and Technology, Lancaster University, Lancaster LA1 4YR, UK; Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK.

data covering the contiguous US (CONUS) and Mexico [23]. However, due to frequent cloud cover, there always exists a large amount of data loss in the ABI LST product. Thus, to obtain spatially complete, hourly LST other auxiliary data have been applied to fill the gaps in the ABI LST product. Specifically, the European Centre for Medium-range Weather Forecasts (ECMWF) Reanalysis *v*.5 (ERA5)-Land product can provide ~10 km×10 km (0.1 degree) all-sky, hourly observations [24]. Ultimately, Jia *et al.* [21] filled the spatial gaps and produced 2 km, hourly all-sky LST data based on a spatio-temporal dynamic model constructed from ERA5-Land (simplified as ERA5 hereafter). Although the generated 2 km, hourly LST product has great advantages in characterizing high-frequency surface thermal dynamics at medium scale, it cannot meet the demands of fine spatial resolution monitoring. For example, in forest fires monitoring, the boundary of fire may not be captured clearly in the 2 km spatial resolution product.

To obtain LST with the potential for monitoring at both fine spatial and temporal resolutions, auxiliary LST products at fine spatial resolution are required. As a typical TIR sensor, Landsat 8 Thermal Infrared Sensor (TIRS) band 10 can detect LST at a spatial resolution of 100 m. The Landsat 8 LST product is generated from this band by a single channel algorithm. The temporal resolution of this product, however, is 16 days, restricting greatly applications requiring frequent monitoring. To combine the advantages of the 2 km, hourly LST and 100 m, 16-day Landsat LST products, spatio-temporal fusion methods can be adopted [25]-[27]. Spatio-temporal fusion aims at generating images with both fine spatial and temporal resolutions, by combining temporally sparse, fine spatial resolution images and temporally dense, coarse spatial resolution images [28]-[30]. So far, several spatial weighting-based [31]-[33] (e.g., the spatial and temporal adaptive reflectance fusion model (STARFM) [34], the enhanced STARFM (ESTARFM) method [31] and the spatial weighting-based virtual image pair fusion (VIPSTF-SW) approach [32]), spatial unmixing-based [35]-[37] (e.g., the unmixing-based data fusion (UBDF) model [35], the spatial temporal data fusion approach (STDFA) [38] and the blocks-removed spatial unmixing (SU-BR) method [36]), learning-based [39], [40] (e.g., the Sparse representation-based SpatioTemporal reflectance Fusion Model (SPSTFM) [39]) and hybrid methods [41-43] (e.g., Flexible Spatiotemporal DAta Fusion (FSDAF) approach [41]) were developed. Although spatio-temporal fusion methods were proposed initially for fusing the surface reflectance of images from different data sources, the feasibility of their application to the fusion of LST images was also demonstrated over recent years [44]-[46].

This research aims at generating 100 m, hourly LST images, by spatio-temporal fusion of 2 km, hourly GOES-16 ABI LST and 100 m Landsat LST images acquired at the same time. So far, several studies have focused on the generation of fine spatial resolution images with hourly or even finer temporal resolution. For example, Inamdar *et al.* [47] acquired 1 km, hourly LST by fusing GOES LST, MODIS LST and MODIS normalized difference vegetation index (NDVI). Zaksek and Ostir [48] further increased the resolution of LST to 1 km, 15 min, by merging spinning enhanced visible and infrared imager (SEVIRI) and MODIS LST. Wu *et al.* [49] fused Landsat, MODIS and GOES LST to obtain the 100 m, 30 min LST images. Due to gaps in the GOES data, however, the model failed to generate temporally continuous 24-hour LST data. Furthermore, Quan *et al.* [50] first generated the 100 m daily LST by fusing Landsat and MODIS LST, and then increased the spatial resolution to the diurnal scale based on a diurnal temperature cycle model. Recently, Ma *et al.* [51] obtained 60 m, 30 min LST data by fusing satellite (i.e., Landsat and MODIS) and LSM (community land model version 5.0 (CLM 5.0))-based data. However, almost all of these studies failed to acquire temporally continuous hourly data due to the lack of spatially and temporally seamless hourly basic data. That is, most of the research utilized hourly LST products retrieved from the original satellite sensor data containing different degrees of gaps. Furthermore, several studies utilized LSM-based data for generating spatially and temporally continuous LST, but then the fusion task may be challenging due to the difference between the satellite and LSM-based data.

In this paper, as an alternative solution, the newly released spatially and temporally seamless 2 km, hourly GOES-16 ABI product was applied, to generate continuous 100 m, hourly LST data. Specifically, as the original 2 km, hourly LST product employed in this research was measured directly using a satellite TIR sensor (i.e., GOES-16 ABI), it is more advantageous for revealing the continuous real LST compared to LSMs (e.g., GLDAS and ERA5), which depend greatly on the reliability of the model. Thus, the generated LST in this paper has the potential to produce greater accuracy than most of the related research based on LSMs. In this paper, to compare the accuracy of 100 m, hourly LST derived from satellite observations and LSMs, the Landsat LST images were fused with 2 km GOES-16 ABI LST images and 10 km ERA5 LST images, respectively. Generally, the generated 100 m, hourly LST has great potential for applications experiencing great spatial and temporal heterogeneity, such as urban heat flux monitoring [51], [52] and growth monitoring of smallholder crops. The contribution of this paper can be summarized as follows.

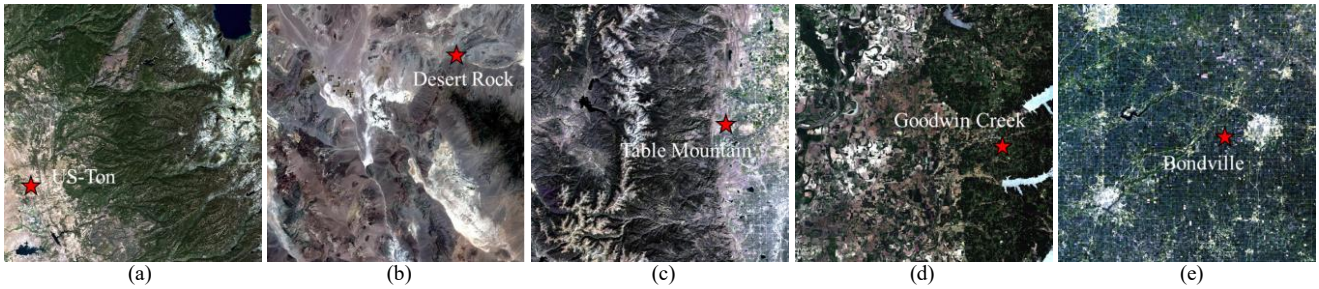
1) The feasibility of applying spatio-temporal fusion to produce 100 m, hourly LST data was evaluated, by fusion of the newly released 2 km, hourly GOES-16 ABI LST and 100 m Landsat LST images. Moreover, the most accurate method was identified through comparison across several popular spatio-temporal fusion methods.
2) The advantage of using the satellite-derived data (i.e., GOES-16 ABI LST) in generating 100 m, hourly LST was validated through a comprehensive comparison between fusing Landsat with GOES-16 ABI (satellite-based LST) LST data and fusing Landsat with ERA5 (LSM) LST data.

The remainder of this paper is organized as follows. In Section II, the study area and data are introduced. In Section III, the methods applied to generate 100 m, hourly LST are presented in detail. In Section IV, the performances of the spatio-temporal fusion methods were compared for generation of the 100 m, hourly data, and a comparison was made between the use of GOES and ERA5 LST data. Section V discusses the advantages and potential of this research. Finally, conclusions are summarized in Section VI.

II. Study Area and Data

A. Study area

To examine the effectiveness of the proposed scheme for producing 100 m, hourly LST data, five regions located in the United States were selected, and one validation site chosen in each region. The Landsat 8 data of the five study areas are shown in Fig. 1. The five study areas all cover a spatial extent of 100 km ×100 km. Specifically, Regions 1 and 2 are located in California and Regions 3-5 are located in Colorado, Mississippi and Illinois, respectively. For the five ground sites, Desert Rock (40°43′13.19″ N, 116°01′10.63″ W), Table Mountain (40°07′32.05″ N, 105°14′15.90″ W), Goodwin Creek (34°15′16.92″ N, 89°52′22.44″ W) and Bondville (40°03′05.58″ N, 88°22′23.70″ W) are from the set of Surface Radiation (SURFRAD) stations, while US-Ton (38°25′51.24″ N, 120°57′57.60″ W) is from the core AmeriFlux network. The main land covers of the five stations are woody savannas, arid shrubland, grass and shrub, pastureland, and cropland, respectively. It is noted that not all the stations for the SURFRAD and core AmeriFlux networks were included in the assessment. The reason is that for some stations parts of the Landsat LST data were contaminated by cloud and for some other stations the spatial heterogeneity of the LST data in the regions was not representative for validation.



(a) (b) (c) (d) (e)

Fig. 1. The five study areas (red, green and blue bands of Landsat 8 data as RGB). (a)-(e) represents Regions 1-5, respectively. (a) and (b) are located in California and (c), (d) and (e) are located in Denver, Mississippi and Illinois, respectively.

Table 1 Acquisition time of Landsat LST for the five regions

Region 1	Region 2	Region 3	Region 4	Region 5
2 August 2020	12 May 2021	29 April 2020	13 January 2021	9 February 2019
3 September 2020	28 May 2021	4 September 2020	19 April 2021	4 August 2019
5 October 2020	13 June 2021	6 October 2020	22 June 2021	7 October 2019

B. Data

(1) Satellite sensor data. The Landsat 8 surface temperature and GOES-16 ABI LST satellite sensor products are included in this research. The Landsat 8 surface temperature product was acquired from the United States Geological Survey (https://earthexplorer.usgs.gov/), which was generated by applying the single channel algorithm to Collection 2 Level 1 Thermal Infrared Sensor (TIRS) band 10 of Landsat 8. The specific acquisition time of Landsat LST used in this research is listed in Table 1. The GOES-16 ABI LST is a 2 km, hourly product providing all-sky observation for the CONUS and Mexico regions from July 2017 to June 2021, which can be obtained from http://glass.umd.edu/allsky_LST/. Specifically, this product was generated by a surface energy balance-based method. By constructing a spatio-temporal dynamic model from ERA5 data and assimilating the clear-sky GOES LST to the model, continuous GOES LST images series were reconstructed. Then, the diurnal cloud radiative effect (CRE) was calculated from satellite radiation products and the cloudy pixels were reconstructed [21]. The GOES-16 ABI LST utilized in this research is hourly data acquired on the same dates as the Landsat LST.

(2) LSM data. ERA5 is the fifth generation ECMWF atmospheric reanalysis of the global climate available from 1979 to the present [53], which is available at https://www.ecmwf.int/. By optimally combining observations and model data, the reanalysis of ERA5 produces continuous data across the world for several surface parameters. In addition, the skin temperature product of ERA5 reanalysis data was applied, which can provide hourly LST at the spatial resolution of ~10 km×10 km (0.1 degree).

(3) In situ data. To validate the accuracy of the generated 100 m, hourly LST images, ground-based *in situ* data were included in this research. Specifically, four sites from the SURFRAD stations (https://gml.noaa.gov/grad/surfrad/) and one site from the Core AmeriFlux stations (https://ameriflux.lbl.gov/) are involved. The SURFRAD networks can provide the upwelling and downscaling thermal infrared parameters at a temporal resolution of 1 min covering the CONUS. The ground measurement data provided by the Core AmeriFlux sites are presented at 30 min or 1 hour temporal resolution. *In situ* data from the SURFRAD and the Core AmeriFlux stations can provide hourly surface upwelling and atmospheric downwelling longwave radiation, which are core parameters in the derivation of ground-based LST. Generally, the ground-based LST can be calculated with the surface upwelling and atmospheric downwelling longwave radiation using the Stefan-Boltzmann law:

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

where T_s is the estimated LST, and $L^{\uparrow}$ and $L^{\downarrow}$ are surface upwelling and atmospheric downwelling longwave radiation,

respectively. σ is the Stefan-Boltzmann's constant ($5.67\times10^{-8}\mathrm{Wm^{-2}K^{-4}}$) and ε_b is the broadband emissivity calculated based on the MODIS/Aqua land surface temperature/3-band emissivity daily L3 global 1 km product (MYD21A1) ([54])

$$\varepsilon_b = 0.2122\cdot\varepsilon_{29} + 0.3859\cdot\varepsilon_{31} + 0.4029\cdot\varepsilon_{32} \quad (2)$$

where ε_{29}, ε_{31} and ε_{32} are the narrowband emissivities of bands 29, 31 and 32 of the MYD21A1 product.

III. METHODS

A. Spatio-temporal fusion for 100 m, hourly LST generation

In this research, 100 m, hourly LST images were generated by spatio-temporal fusion of GOES-16 ABI and Landsat LST images. Generally, the 100 m, hourly LST images for a certain time were estimated using the corresponding GOES-16 ABI LST image at the time, together with a Landsat-GOES LST image pair at a temporally close time. The spatio-temporal fusion process can be summarized as follows.

$$\hat{\mathbf{L}}(h_t) = \mathbf{L}(h_0) + f(\mathbf{G}(h_t) - \mathbf{G}(h_0)) \quad (3)$$

where h_0 is the acquisition time (known time) of the Landsat-GOES LST image pair, and h_t is the target time for prediction (prediction time). $\mathbf{L}$ and $\mathbf{G}$ represent the Landsat and GOES-16 ABI LST images, respectively. f is a downscaling function characterized by different spatio-temporal fusion methods.

In this paper, three spatio-temporal fusion methods (i.e., STARFM [34], VIPSTF-SW [32] and FSDAF [41] methods) were evaluated for the generation of 100 m, hourly LST data. For STARFM, f is a function to weight the neighboring similar pixels by considering comprehensively the spectral difference between the coarse and fine image pairs, the temporal difference between two coarse images and the spatial difference to the center pixel [34]. For the VIPSTF-SW approach, a virtual image pair is first generated to simulate data closer to the target prediction time, by applying a linear transformation model constructed by the two coarse spatial resolution images. Then, the fine and coarse virtual image pairs are fused by the same spatial weighting strategy in STARFM [32]. Furthermore, the f of the FSDAF method incorporates the strategy of spatial weighting to the spatial unmixing-based spatio-temporal fusion method. Specifically, FSDAF applies spatial unmixing to estimate the temporal change of each class, and then performs spatial weighting to distribute the residuals [41].

B. The framework for 100 m, hourly LST generation

The procedure for the generation of 100 m, hourly LST is shown in Fig. 2. The GOES-16 ABI LST and Landsat LST were selected as the main data sources for 100 m, hourly LST generation. Specifically, the hourly GOES-16 ABI LST images were first resampled and reprojected according to the Landsat LST acquired at the same time. Then, to further decrease the influence of data inconsistency on fusion, a data correction step was conducted according to the correlation between the upscaled Landsat LST and GOES-16 ABI LST acquired at the closest time. The specific process of data correction will further be introduced in Section III-*C*. By applying spatio-temporal fusion, the 100 m, hourly LST is produced. Finally, the accuracy of the generated 100 m, hourly LST is validated by *in situ* LST. Moreover, to compare the accuracy of LST generated from satellites sensors and LSMs, the ERA5 LST is also fused with Landsat LST to generate 100 m, hourly LST.

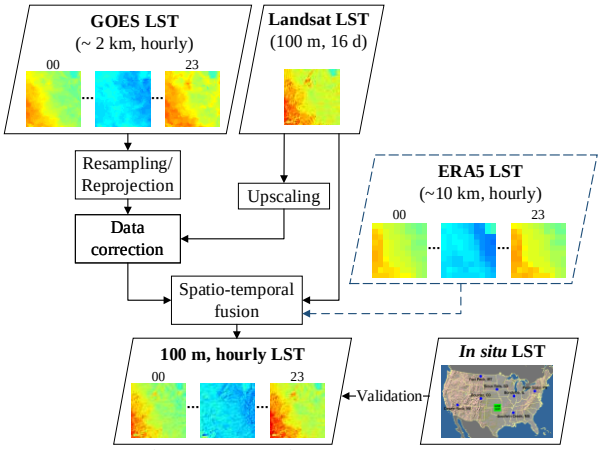


Fig. 2. The procedure for generation of 100 m, hourly LST.

C. Data correction

The difference between GOES and Landsat LST data has been one of the most important issues restricting the accuracy of spatio-temporal fusion. Thus, to decrease the influence of data inconsistency, a linear correction model was applied. The process of data correction between GOES-16 ABI and Landsat LST data is shown in Fig. 3. First, the Landsat LST acquired at t_k (temporally close to the GOES-16 ABI data at the prediction time t_p) is upscaled to the spatial resolution of GOES-16 ABI LST. Then, by applying the least squares regression between the upscaled Landsat LST (i.e., $\mathbf{L}_{\mathrm{up\_scale}}$) and GOES-16 ABI LST (i.e., $\mathbf{M}$) at t_k, the correction coefficients a and b can be obtained, as shown in the fitting model below

$$\mathbf{L}_{\mathrm{up\_scale}} = a\mathbf{M} + b \quad (4)$$

The linear transformation is then applied to the GOES-16 ABI LST at the prediction time t_p to preserve the consistency with Landsat LST. It is noted that for hourly GOES-16 ABI LST, the data acquired on the same date will be corrected collaboratively with the same correction coefficient according to the temporally close Landsat and GOES-16 ABI LST image pair, based on the assumption that the relation is consistent throughout the whole day. It would also be worthwhile to consider spatially adaptive regression models when the spatial extent of the study area is large, due to the difference in observation angles between Landsat and GOES.

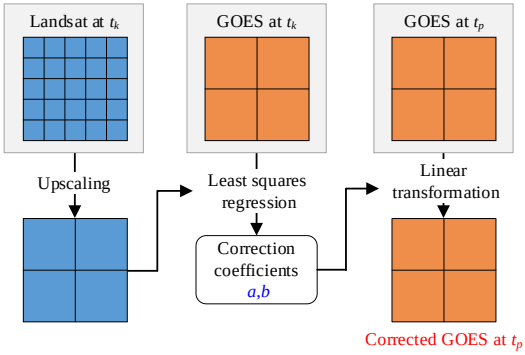


Fig. 3. Data correction of GOES-16 ABI LST.

D. Validation strategy

In this research, the effectiveness of the 100 m, hourly LST was validated in both the spatial and temporal domains (i.e., spatial validation and temporal validation). For spatial validation, the Landsat and GOES-16 ABI/ERA5 LST images at t_k, together with GOES-16 ABI/ERA5 LST images at t_p were treated as known images. By applying spatio-temporal fusion, the Landsat-like 100 m LST at t_p can be predicted. With the original 100 m Landsat LST images at t_p (known in advance), the accuracy of fusion can be evaluated in the spatial domain.

For temporal validation, the 100 m, hourly LST images will be generated for 24 hours within the day. Specifically, the Landsat LST and the GOES-16 ABI LST temporally closest to the acquisition time of Landsat LST were selected as known image pairs. By fusing the known image pairs with each GOES-16 ABI LST within the acquisition day, 100 m LST for 24 hours can be generated. The temporal validation is then performed by checking the *in situ* LST time-series and the value of the corresponding pixels in the generated 100 m, hourly LST time-series.

For spatial and temporal validation, three statistical metrics were applied for evaluating the accuracy of the generated 100 m LST. In this research, the root mean square error (RMSE), the mean absolute error (MAE) and the coefficient of determination (R^2) were adopted. The statistical metrics used in the experiment are defined as follows:

$$RMSE = \frac{1}{mn}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{P}_{(i,j)} - \mathbf{R}_{(i,j)})} \tag{5}$$

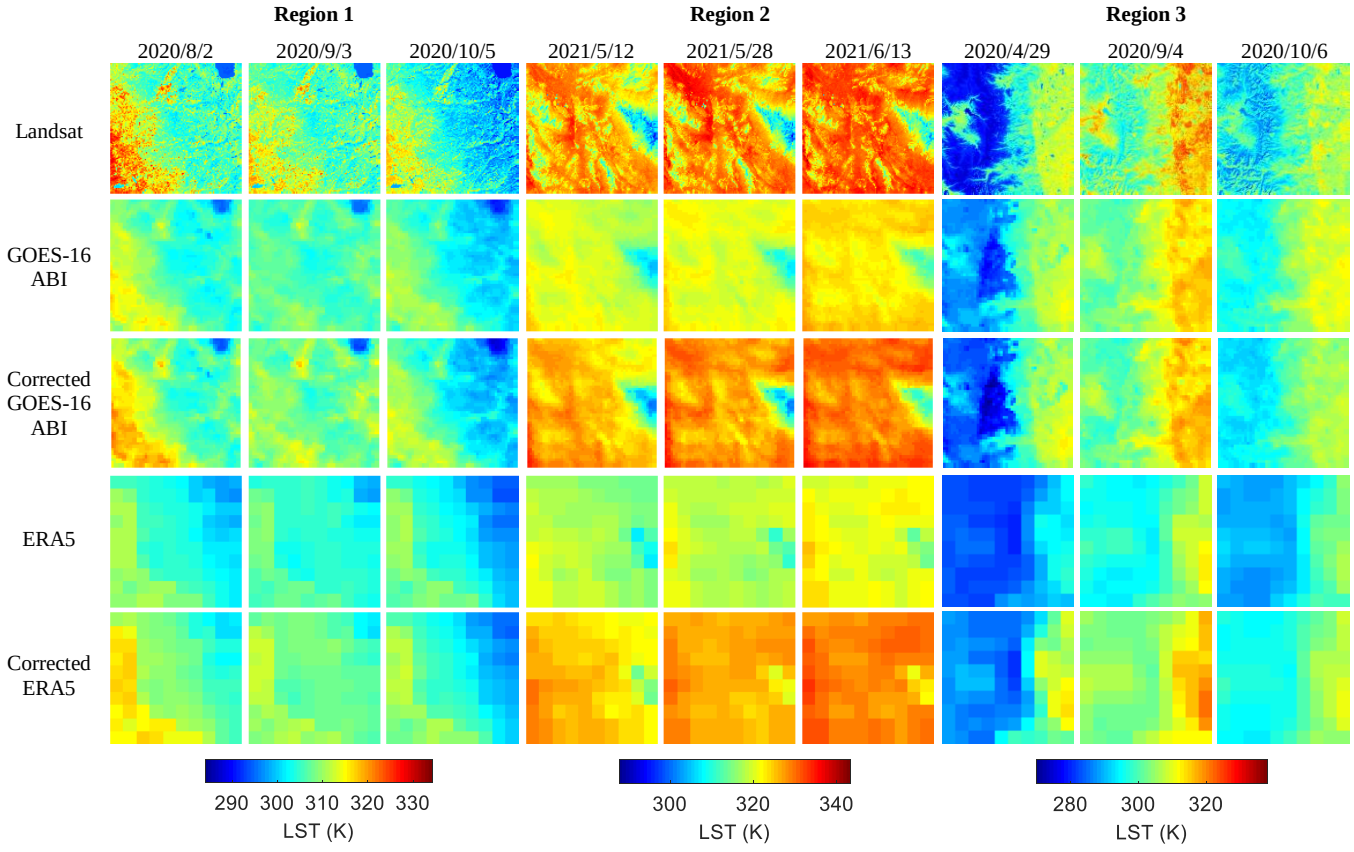
$$MAE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left|\mathbf{P}_{(i,j)} - \mathbf{R}_{(i,j)}\right| \tag{6}$$

$$R^2 = \frac{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{P}_{(i,j)} - \mu_P)(\mathbf{R}_{(i,j)} - \mu_R)\right)^2}{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{P}_{(i,j)} - \mu_P)^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{R}_{(i,j)} - \mu_R)^2\right)} \tag{7}$$

where $\mathbf{P}$ and $\mathbf{R}$ represent the prediction and reference data with the same size of $m \times n$, respectively. $\mathbf{P}_{(i,j)}$ is the pixel located at (i, j) in $\mathbf{P}$. μ_P and μ_R represent the mean values of $\mathbf{P}$ and $\mathbf{R}$, respectively.

IV. Experiments

The experiments in this research are divided into five parts. Section IV-*A* presented the LST data utilized in the experiments. In Section IV-*B*, the performance of different spatio-temporal fusion methods was compared both spatially and temporally, to find the most suitable method for generation of the 100 m, hourly LST images. In Section IV-*C*, a comparison was made between the generation of 100 m, hourly LST images by fusing Landsat LST with GOES LST and fusing Landsat LST with ERA5 LST. In Section IV-*D*, the advantage of data correction was analyzed. Section IV-*E* investigated the impact of similar pixel selection scheme on spatio-temporal fusion of Landsat and GOES LST images. To further investigate the feasibility of generating 100 m, hourly LST over a longer period, an experiment producing LST for one month was implemented in Section IV-*F*.



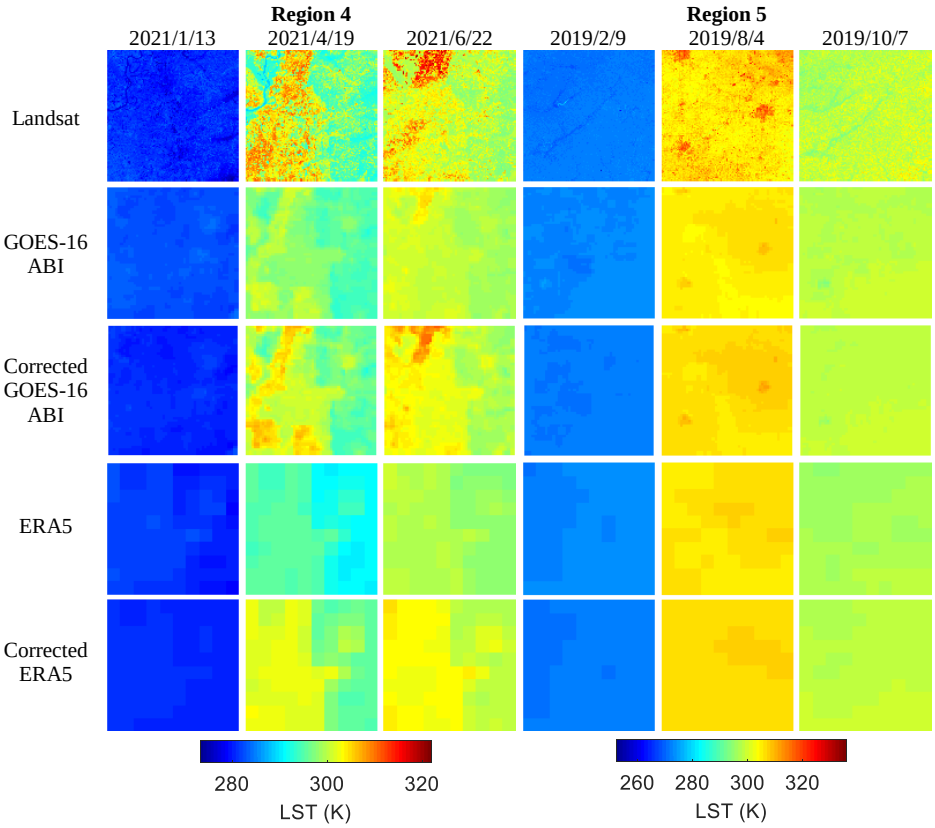


Fig. 4. Landsat, original and corrected GOES-16 ABI, and original and corrected ERA5 LST for the five regions.

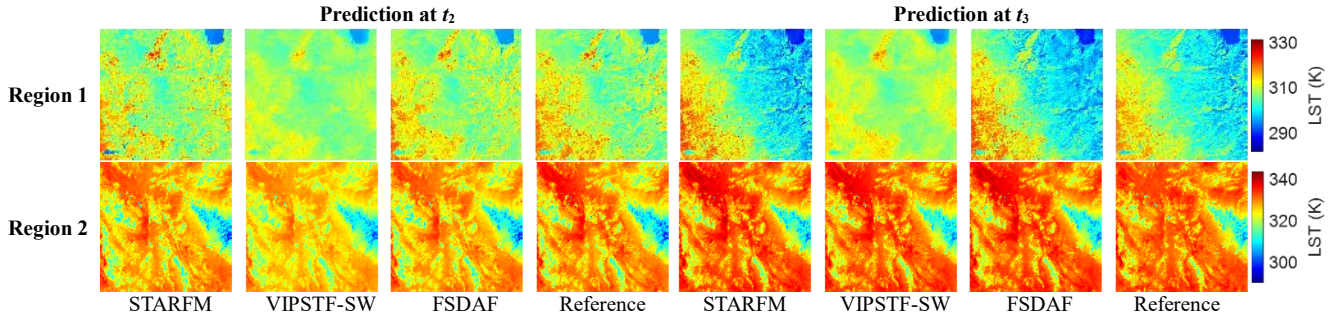


Fig. 5. Spatial validation by applying different spatio-temporal fusion methods.

A. LST data presentation

The Landsat, GOES-16 ABI and ERA5 LST data covering the five regions were used in the experiments in this section. For each region, LST images acquired at three separate times were used for validation. As GOES-16 ABI and ERA5 are both hourly LST products, the LST images acquired temporally closest to the Landsat LST (i.e., UTC 18:00, 19:00, 18:00, 17:00, 17:00 for Regions 1-5, respectively) are presented, as shown in Fig. 4. It is noted that for the five regions, the dates were selected according to data integrity and small areas of data loss were filled by interpolation methods. Considering that there exist some differences in LST between the three data sources, the data correction method presented in Section III-*C* was applied to both the GOES-16 ABI and ERA5 LST images, with the Landsat images as the basis for correction. As seen clearly in Fig. 4, LST from the three different data sources tends to be similar after correction. In the following parts, for convenience, the three dates for each region are referenced as t_1, t_2 and t_3.

B. Comparison between different spatio-temporal fusion methods in the generation of 100 m, hourly LST data

In this section, a comparison between the different methods was performed to identify the most accurate method for the generation of 100 m, hourly data. Specifically, the comparison was performed in both the spatial and temporal domains.

1) Spatial validation. The data for Regions 1 and 2 were selected for spatial validation. For each region, the LST image pair at t_1 and t_2 was selected as the known image pair, and the Landsat LST images at t_2 and t_3 were predicted, in turn. The results of different spatio-temporal fusion methods (i.e., STARFM, VIPSTF-SW and FSDAF) are shown in Fig. 5. From visual inspection, it is found that for both regions, the predictions of STARFM are the closest to the reference. Furthermore, a quantitative assessment was also performed, as shown in Table 2. It is noted that STARFM produced the smallest RMSE and the

largest R^2 in most cases.

Table 2 Accuracy of spatial validation results generated by different spatio-temporal fusion methods

		Method	Prediction at t_2	Prediction at t_3
RMSE (K)	Region 1	STARFM	**1.472**	**1.774**
		VIPSTF-SW	2.798	2.280
		FSDAF	1.628	2.048
	Region 2	STARFM	**1.743**	1.813
		VIPSTF-SW	2.931	**1.652**
		FSDAF	1.767	1.809
R^2	Region 1	STARFM	**0.919**	**0.963**
		VIPSTF-SW	0.853	0.962
		FSDAF	0.900	0.940
	Region 2	STARFM	0.954	0.922
		VIPSTF-SW	**0.957**	**0.924**
		FSDAF	0.954	0.919

2) Temporal validation. In the temporal validation, the GOES LST images and Landsat LST at the closest acquisition time were selected as the known image pair, together with the GOES LST at the other 23 hour times to predict the 100 m, hourly LST for 24 hours. The corresponding accuracy of different spatio-temporal fusion methods is listed in Table 3. Considering the MAE, the prediction of STARFM is the most accurate. For example, for Region 5, STARFM produces the smallest MAE of 2.179 K, which is 0.603 K and 0.132 K smaller than for VIPSTF-SW and FSDAF, respectively. For the R^2, STARFM also provides the largest values for most of the regions. Although VIPSTF-SW and FSDAF also produce satisfactory performance in several regions, STARFM is the most accurate for comprehensive inspection. Considering the performances of the three methods for both spatial and temporal validation, STARFM can serve as an appropriate solution for generating 100 m, hourly LST data.

Table 3 Accuracy of temporal validation results generated by different spatio-temporal fusion methods

		Region 1	Region 2	Region 3	Region 4	Region 5
MAE (K)	STARFM	2.180	1.714	2.976	**1.745**	**2.179**
	VIPSTF-SW	2.819	**1.501**	3.720	2.096	2.782
	FSDAF	**2.120**	1.723	**2.843**	5.023	2.311
R^2	STARFM	0.989	**0.988**	**0.985**	0.970	**0.980**
	VIPSTF-SW	**0.992**	**0.988**	0.949	**0.986**	0.976
	FSDAF	**0.992**	**0.988**	0.943	0.972	0.979

C. Comparison between the use of GOES and ERA5 data in generation of 100 m, hourly LST data

In this section, a comparison was made between fusing Landsat with GOES LST and fusing Landsat with ERA5 LST. The data utilized in this section are the same as those for Section IV-*B*. The most accurate method identified in Section IV-*B*, that is, the STARFM approach, was applied to the fusion task in this section. The comparison was also conducted in both the spatial and temporal domains.

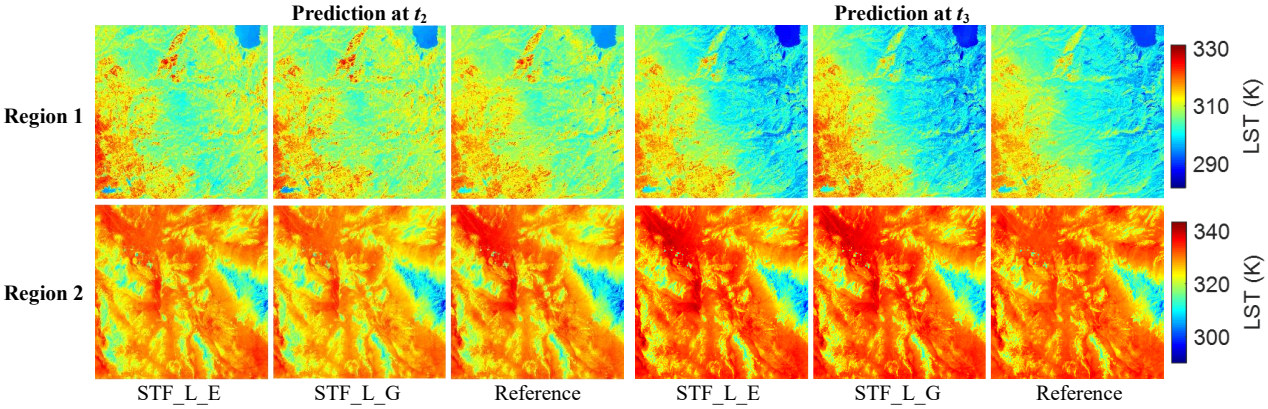


Fig. 6. Results of spatial validation for Regions 1 and 2.

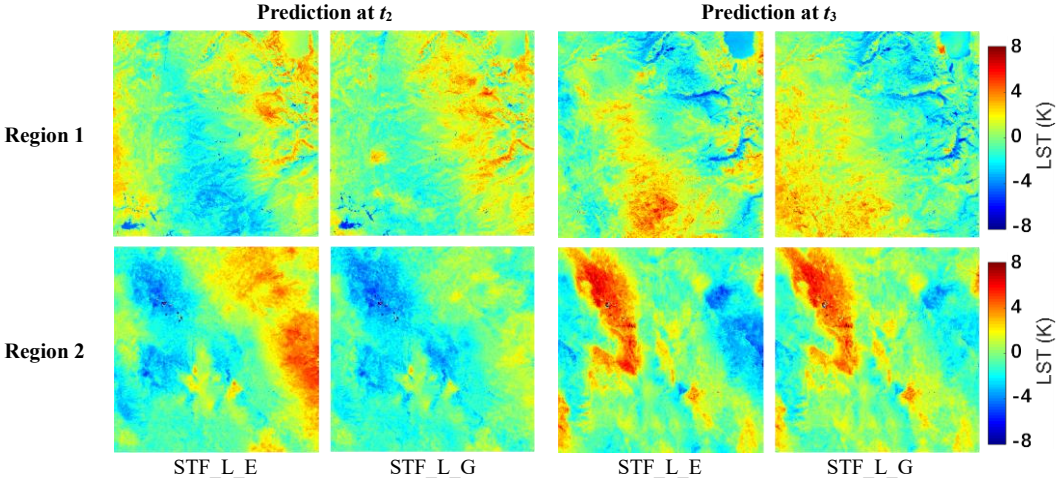


Fig. 7. Error images of spatial validation results for Regions 1 and 2.

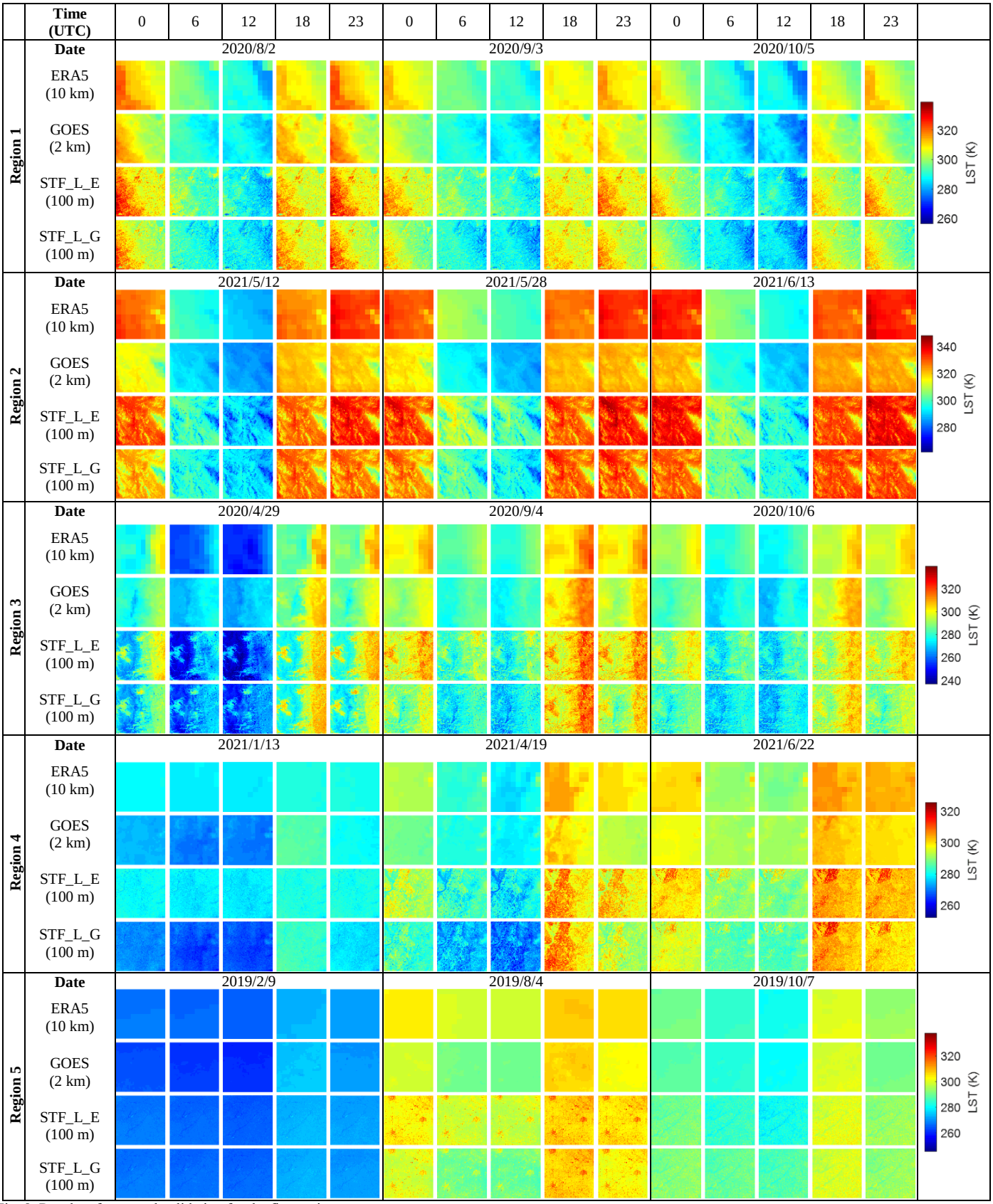


Fig. 8. Results of temporal validation for the five regions.

1) Spatial validation. For simplicity, the 100 m results predicted by fusion of Landsat and GOES-16 ABI LST is abbreviated as STF_L_G, while that predicted by fusing Landsat and ERA5 LST is named STF_L_E. The spatial validation results are shown in Fig. 6. From visual inspection, we can find that the predictions by STF_L_G and STF_L_E are both similar

to the reference LST. To depict the difference more clearly, the error images are presented in Fig. 7. For Region 1, the color is obviously lighter for STF_L_G prediction, indicating a smaller error. Specifically, for the STF_L_G prediction, the blue in the lower middle of t_2 and the right upper of t_3, and red in the lower middle of t_3 are lighter than for the STF_L_E prediction. Thus, the STF_L_G prediction is demonstrated to be more accurate in terms of spatial validation.

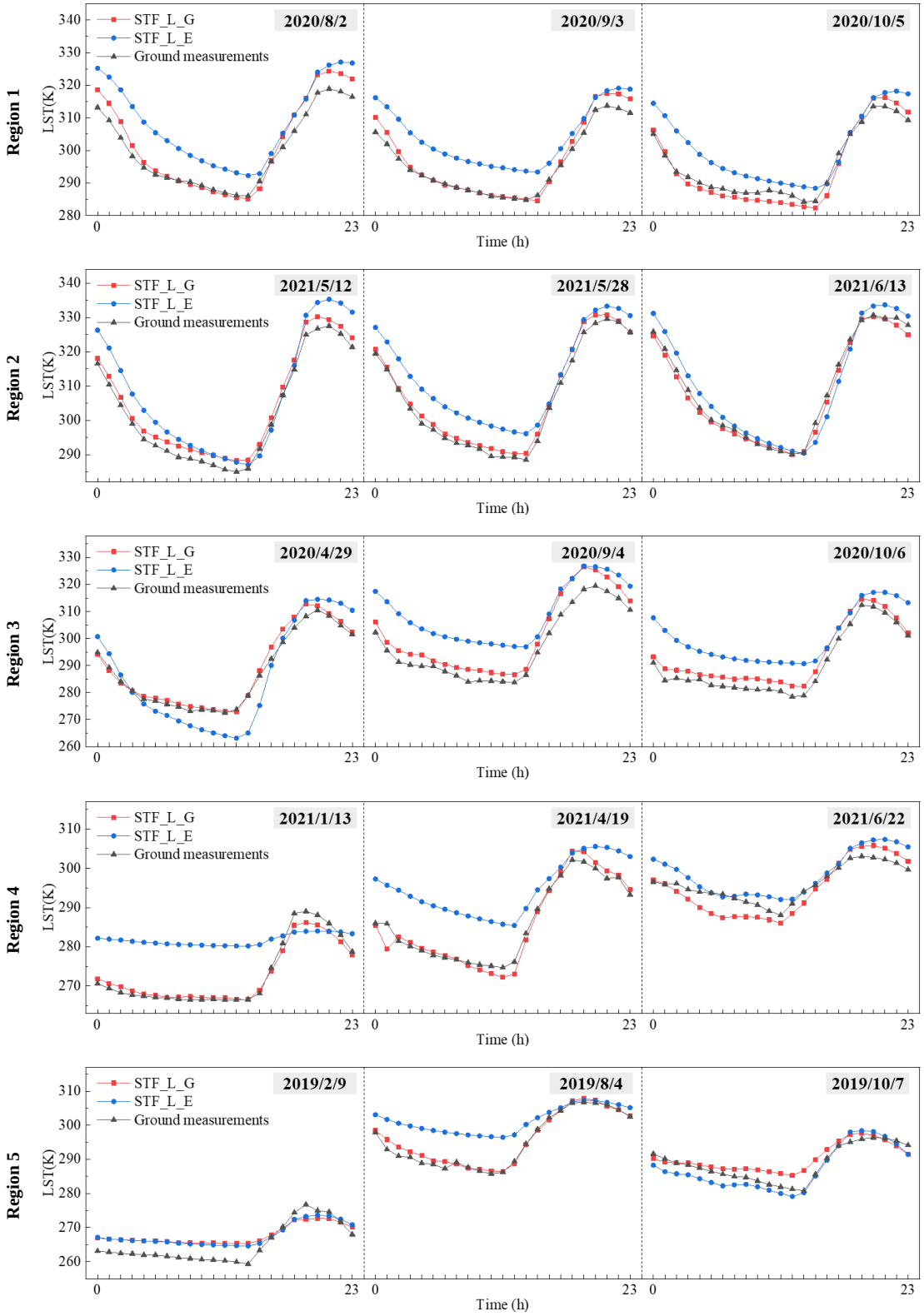


Fig. 9. Accuracy of temporal validation results for the five regions.

Quantitative assessment was performed for the STF_L_G and STF_L_E predictions, as listed in Table 4. Generally, STF_L_G presents greater accuracy for both regions. When predicting the 100 m LST at t_2, STF_L_G always produces greater accuracy

than STF_L_E. Specifically, STF_L_G produces RMSEs of 1.472 K and 1.743 K for Regions 1 and 2, which are 0.193 K and 0.294 K smaller than for STF_L_E. Also, the R^2 increases by 0.018 and 0.062 when applying STF_L_G compared to STF_L_E. For the prediction at t_3, the R^2 of STF_L_G are 0.028 and 0.013 larger than STF_L_E for Regions 1 and 2, respectively. From the spatial validation, it is found that the 100 m LST generated by fusing Landsat and GOES-16 ABI LST leads to more satisfactory performance than that produced by fusion of Landsat and ERA5 LST images.

Table 4 Accuracy of spatial validation results for Regions 1 and 2

			Prediction at t_2	Prediction at t_3
RMSE (K)	Region 1	STF_L_G	**1.472**	1.774
		STF_L_E	1.665	**1.760**
	Region 2	STF_L_G	**1.743**	**1.813**
		STF_L_E	2.037	2.031
R^2	Region 1	STF_L_G	**0.919**	**0.963**
		STF_L_E	0.901	0.935
	Region 2	STF_L_G	**0.954**	**0.922**
		STF_L_E	0.892	0.909

(2) Temporal validation. In temporal validation, the coarse spatial resolution LST image (i.e., ERA5 or GOES) and Landsat LST at the closest acquisition time were selected as the known image pair, together with the coarse spatial resolution LST at the other 23 hour times to predict the 100 m, hourly LST for 24 hours. For clearer presentation, the predictions at UTC 0:00, 6:00, 12:00, 18:00 and 23:00 are shown in Fig. 8. In Fig. 8, the four rows represent the 10 km ERA5, 2 km GOES, 100 m STF_L_E and STF_L_G predictions, from top to bottom. It is obvious that the 100 m results contain more details compared to the original 10 km and 2 km LST images, such as the clearer boundaries between high and low LST areas. Generally, the predictions preserve the color of the original coarse spatial resolution LST images. For all five regions, the STF_L_G and STF_L_E predictions both present diurnal variation in LST.

To further compare the performance of the STF_L_G and STF_L_E predictions, the quantitative assessment for temporal validation was performed, as shown in Fig. 9. Due to the absence of reference data in the spatial domain, the ground measurements (i.e., the *in situ* data) were applied for temporal validation. For all five regions, the STF_L_G prediction tends to be closer to the ground measurement compared to the STF_L_E prediction. For Regions 1 and 2, the STF_L_G predictions are almost coincident with the ground measurements. Although the STF_L_E predictions present a similar trend to the ground measurements, they produce larger LST values than the ground measurements, especially for the predictions far away from the known time. For Regions 3-4, although both STF_L_G and STF_L_E fail to produce accurate predictions at several time points, the predictions of STF_L_G are obviously closer to the ground measurements generally.

The scatterplots between the predicted and measured LST for the five regions are shown in Fig. 10. It is noted that the points in STF_L_G predictions are closer to the $y=x$ line compared to the STF_L_E predictions. More precisely, STF_L_G produces MAEs of 2.180 K, 1.714 K, 2.976 K, 1.745 K and 2.179 K for Regions 1-5, respectively, which are 5.203 K, 3.338 K, 6.204 K, 5.458 K and 1.605 K smaller than for STF_L_E correspondingly. Similarly, for RMSE, the predictions of STF_L_G are 5.403 K, 3.912 K, 6.661 K, 6.478 K and 2.094 K smaller than for STF_L_E for Regions 1-5, respectively. Furthermore, the R^2 for STF_L_G is 0.098, 0.051, 0.178, 0.143 and 0.064 larger than for STF_L_E for Regions 1-5, respectively. Generally, STF_L_G has a satisfactory performance in temporal validation, and produces greater accuracy than STF_L_E.

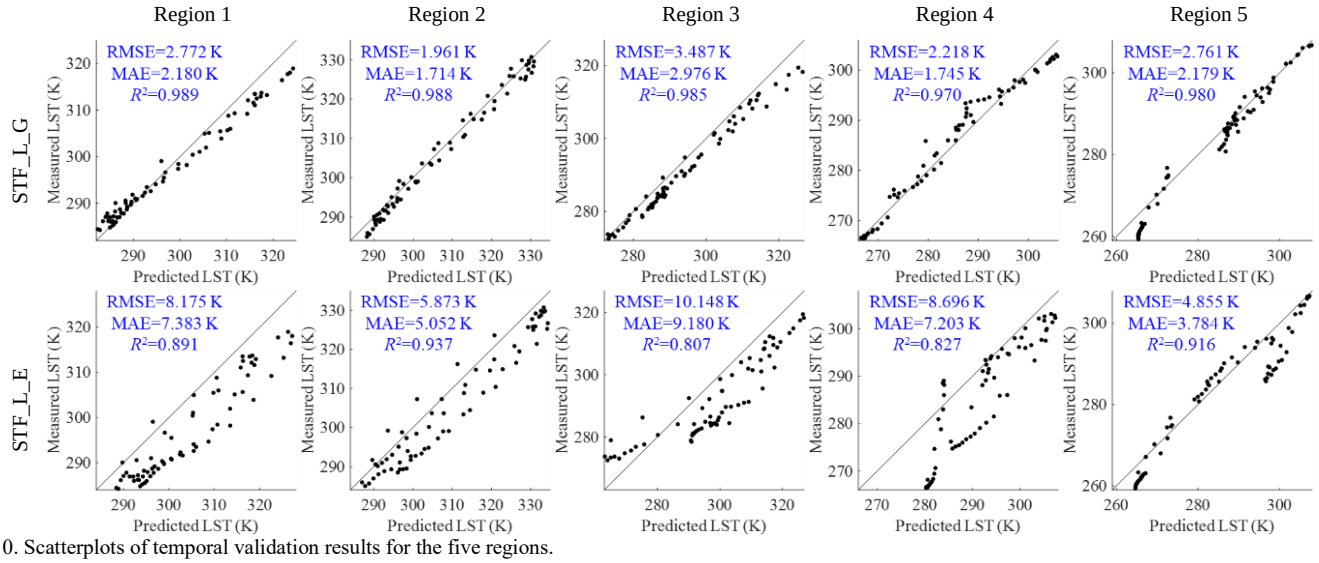


Fig. 10. Scatterplots of temporal validation results for the five regions.

D. The influence of data correction

In this paper, a data correction step was applied before spatio-temporal fusion. To validate the effectiveness of data correction, a comparison was performed between the original and corrected GOES/ERA5 LST-derived predictions. From the data presented in Fig. 4, it is noted that there exist obvious differences between the original GOES/ERA5 LST and Landsat LST. Generally, the original GOES/ERA5 LST is lower than the Landsat LST, especially in areas with high LST. On the contrary, the corrected GOES/ERA5 LST is visually much more

consistent with the Landsat LST regarding the spatial contribution of LST. To quantitatively evaluate the performance of data correction, the RMSEs between the original (or corrected) GOES/ERA5 LST and the upscaled Landsat LST were calculated, as shown in Fig. 11. It can be seen that the RMSE decreases obviously by applying data correction for both GOES and ERA5 LST in all five regions. For example, for Region 2, the RMSE for GOES LST decreases by 2.971 K, 3.888 K and 3.849 K for t_1, t_2 and t_3, respectively. For Region 3, the RMSE decreases by 3.703 K, 5.006 K and 3.610 K for ERA5 LST.

It is acknowledged that the difference between two data sources tends to be an important factor restricting the accuracy of spatio-temporal fusion. To further investigate the influence of data correction on fusion, spatio-temporal fusion was applied based on the original (i.e., STF_L_originalG) and the corrected GOES-16 ABI LST (i.e., STF_L_correctedG), respectively, with Region 4 as an example. Fig. 12 shows the absolute bias of prediction based on the original and corrected GOES-16 ABI LST. Overall, STF_L_correctedG leads to greater accuracy than STF_L_originalG in most cases. Specifically, STF_L_correctedG produces a RMSE of 2.218 K and MAE of 1.745 K, which are 2.254 K and 1.864 K smaller than for STF_L_originalG, respectively. Moreover, the R^2 increases by 0.0358 when applying STF_L_correctedG. Thus, the data correction process can decrease the difference between LST images from various data sources and, furthermore, increase the accuracy of the predicted 100 m, hourly LST.

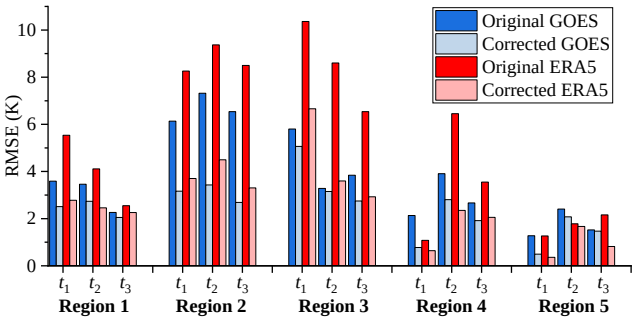


Fig. 11. The RMSEs between the original (or corrected) GOES/ERA5 LST and the upscaled Landsat LST.

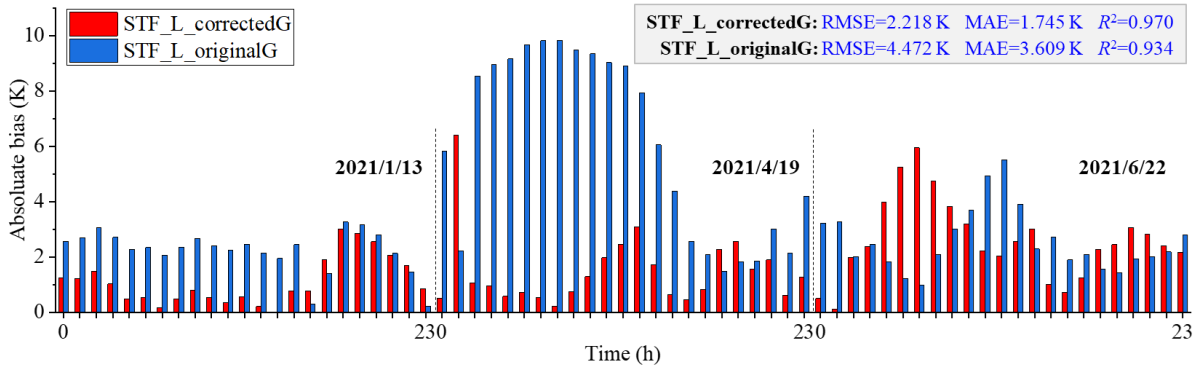


Fig. 12. Absolute bias for prediction based on original and corrected GOES-16 ABI LST (Region 4 as an example).

E. Comparison between different similar pixel selection schemes for STARFM

As spatio-temporal fusion methods such as STARFM were originally proposed for the fusion of surface reflectance data, when applied to the fusion of LST data, the effective scheme of similar pixel selection needs to be identified. In spatio-temporal fusion of reflectance data, similar pixels were selected by considering the properties of neighboring pixels from all available multispectral bands. When applied to fuse LST data, however, there exists only one band for determination of similar pixels. To the best of our knowledge, very few studies have discussed the selection scheme in LST fusion. Thus, in this section, the performances of two similar pixel selection schemes were compared, one using the TIR band, while the other using reflectance bands. The experiment was conducted using the LST data from the spatial validation, together with the corresponding reflectance data. The accuracies achieved by applying different versions of STARFM are listed in Table 5. As can be noted, there exists no obvious difference between the accuracies based on the use of TIR or reflectance bands for similar pixel selection. Thus, the similar pixels can be selected using either TIR or reflectance data, which does not have an obvious impact on the performance of STARFM.

Table 5 Accuracy of spatial validation results generated by different versions of STARFM (in terms of different similar pixel selection schemes; TIR means selecting similar pixels according to the TIR band, while REF means selecting similar pixels according to the reflectance bands)

		Band	Prediction at t_2	Prediction at t_3
RMSE (K)	Region 1	TIR	**1.472**	**1.774**
		REF	1.476	1.782
	Region 2	TIR	**1.743**	**1.813**
		REF	1.744	1.814
R^2	Region 1	TIR	**0.919**	**0.963**
		REF	**0.919**	**0.963**
	Region 2	TIR	**0.954**	**0.922**
		REF	**0.954**	**0.922**

F. Generation of 100 m, hourly LST for one month

In practice, there exists a great need for continuous monitoring over months or even years. The approach developed here has great potential for generating 100 m, hourly LST for long time periods. It is noted that the three dates selected for Region 2 last over a month. Thus, the predicted 100 m, hourly LST can be applied for reconstructing the missing LST images within one month, together with the 2 km GOES-16 ABI LST images over the month. Specifically, the LST at a certain time of the day was predicted based on the temporally closest two known 100 m LST image pairs (e.g., the closest two among 13 May, 28 May and 12 June) at the corresponding time point

within the day and related 2 km LST GOES-16 ABI images. For example, when predicting the 100 m LST at UTC 12:00 on 18 May, the 100 m LST at UTC 12:00 on 13 May and 28 May, and the GOES-16 ABI LST images at UTC 12:00 for the three days were included for fusion. An accuracy assessment based on the ground measurements (i.e., the *in situ* data) was performed for the predicted 100 m, hourly LST for one month, as shown in Fig. 13. Generally, the curve of the predicted LST can fit well with that for the measured LST, except for the several days experiencing very large LST fluctuations. The predicted LST produces a RMSE, MAE and R^2 of 4.333 K, 3.017 K and 0.914, respectively. For the scatterplot, the points are close to the $y=x$ line, indicating satisfactory performance. Checking the error distribution in Fig. 13(c), it is noted that the error presents a normal distribution, and about 64.9% of the errors fall in the range of -3 K to 3 K.

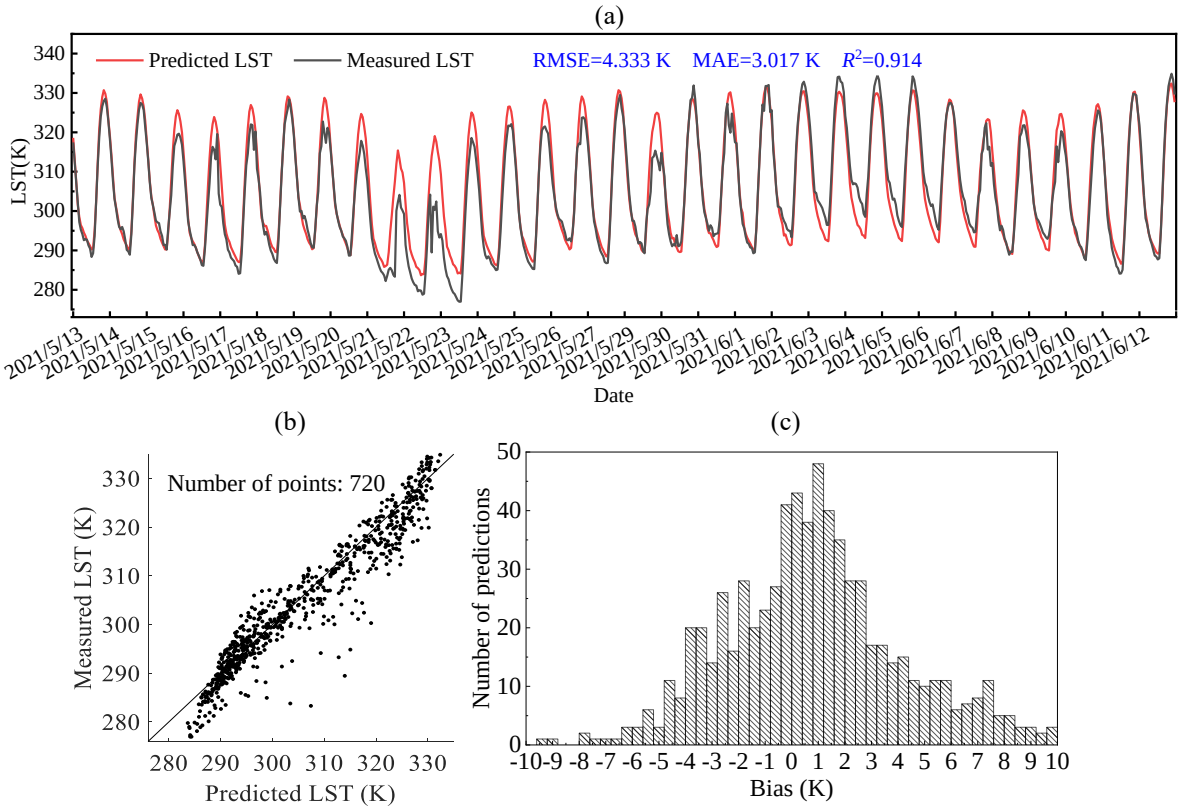


Fig. 13. Accuracy of generated 100 m, hourly LST for one month (Region 2 as an example). (a) Predicted and measured LST (i.e., the *in situ* data). (b) Scatterplot. (c) Diagram of error distribution.

V. Discussion

A. Comparison between GOES and ERA5 LST

In current research, satellite data-derived LST images are fused with LSMs to produce LST images with fine spatial and temporal resolutions. This practice, however, suffers from several limitations. First, as LSMs combine a series of land surface states and fluxes, the accuracy depends greatly on the algorithms and the modeling structure. It is possible that they can provide satisfactory simulation where stations exist, but may fail to simulate accurately the LST at locations spatially far away from the stations. On the contrary, satellite sensors have advantages as they can provide "real" and continuous observations of the Earth's surface. Thus, the LSMs always differ from satellite sensor data which results in great uncertainty in spatio-temporal fusion involving LSM and satellite data. Second, the LST products always have a coarse spatial resolution (e.g., ~14 km × 14 km (1/8 degree) for North American Land Data Assimilation System (NLDAS) and ~28 km × 28 km (1/4 degree) for GLDAS), making fusion with fine spatial resolution satellite LST images challenging. To cope with these problems, two satellite-derived data products, that is, Landsat and GOES LST images, were used in this paper.

Table 6 The R^2 between GOES/ERA5 and Landsat LST images

		t_1	t_2	t_3
Region 1	GOES	**0.940**	**0.661**	**0.956**
	ERA5	0.723	0.460	0.818
Region 2	GOES	**0.777**	**0.716**	**0.789**
	ERA5	0.367	0.119	0.245
Region 3	GOES	**0.896**	**0.913**	**0.924**
	ERA5	0.650	0.690	0.734
Region 4	GOES	**0.460**	**0.737**	**0.744**
	ERA5	0.099	0.599	0.347
Region 5	GOES	0.351	**0.164**	0.347
	ERA5	**0.386**	0.042	**0.502**

To further compare the differences between satellite sensor data and LSMs, the R^2 between Landsat and GOES-16 ABI (or ERA5) LST data were calculated, as listed in Table 6. To exclude the influence of differences in spatial resolution, both the Landsat and GOES-16 ABI LST images were upscaled to the spatial resolution of ERA5 images (i.e., 10 km) before calculating the R^2. In Table 6, it is noted that the R^2 between Landsat and GOES-16 ABI is obviously larger than that between

Landsat and ERA5 in most cases. For example, in Region 1, the R^2 for GOES is 0.940, 0.661 and 0.956 for the three times, which are 0.217, 0.201 and 0.138 larger than for ERA5. Thus, the GOES-16 ABI LST images are demonstrated to be more similar to the Landsat LST images, bringing great advantages for spatio-temporal fusion of LST data.

In this paper, the GOES-16 ABI LST and ERA5 LST images employed in fusion were at different spatial resolutions. To avoid the influence of spatial resolution, the input 2 km GOES-16 ABI LST data were upscaled to the same spatial resolution as that of the ERA5 LST data (10 km). The accuracy of the 100 m LST generated by fusing the Landsat LST with the upscaled GOES-16 ABI LST (STF_L_upscaledG) is shown in Table 7. It is noted that STF_L_upscaledG still produces much greater accuracy than STF_L_E. Thus, the advantage of using GOES-16 ABI LST for spatio-temporal fusion lies mainly in its origin, that is, it is satellite-based, rather than LSM-based.

Table 7 Accuracy of the 100 m LST generated by fusion of upscaled GOES-16 ABI (10 km)/ERA5 and Landsat LST images

		Region 1	Region 2	Region 3	Region 4	Region 5
R^2	STF_L_upscaledG	**0.988**	**0.987**	**0.971**	**0.981**	**0.980**
	STF_L_E	0.891	0.937	0.807	0.827	0.916
MAE	STF_L_upscaledG	**2.237**	**2.181**	**1.864**	**4.087**	**2.129**
	STF_L_E	7.383	5.052	9.180	7.203	3.784

B. Selection of spatio-temporal fusion methods for generating 100 m, hourly LST

In this paper, three popular methods were investigated to generate 100 m, hourly LST data, and the performances were compared. The reason for choosing these methods is that they are typical single image pair-based spatio-temporal fusion methods. That is, they require only one LST image pair at the known time, together with one coarse spatial resolution LST image at the prediction time. Although a number of spatio-temporal fusion methods (such as ESTARFM and learning-based spatio-temporal fusion methods) were proposed and demonstrated to have great accuracy, they usually require more than one image pair as input. For LST data, however, the temporal heterogeneity is always strong, and it is necessary to acquire the temporally closest image pair. The use of multiple image pairs will increase the uncertainty in the spatio-temporal fusion process, as the LST changes between the multiple image pairs can be very large.

C. Application of the generated 100 m, hourly LST

This paper increased the spatial resolution of the 2 km, hourly GOES-16 ABI products to 100 m, providing LST images at a finer spatial resolution, with the potential to support more applications compared to the original product. Fig. 14 presents a case of vegetation monitoring to demonstrate the advantages of the generated 100 m, hourly data, by comparing the diurnal variation of the single pixel in the original 2 km GOES LST image and the corresponding 400 pixels in the generated 100 m LST. Fig. 14(d) shows that the grey lines (LST of the 400 pixels in the generated 100 m LST) cover a larger range than the red line (the single pixel in the 2 km GOES-16 ABI LST), indicating diverse diurnal LST change of the pixels at the spatial resolution of 100 m. Thus, the generated 100 m, hourly LST data can undoubtedly provide more valuable information for applications requiring real-time LST data.

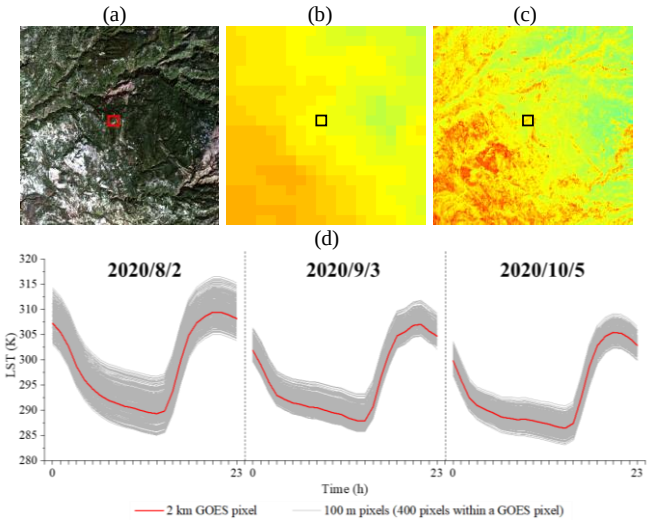


Fig. 14. The diurnal variation of the LST of pixels (with Region 1 as an example). (a) Sub area (red, green and blue bands of Landsat 8 data as RGB). (b) 2 km GOES observation. (c) Generated 100 m LST. (d) The diurnal LST variation of the pixels marked in black frame in (b) and (c). (The red line represents the single pixel in the 2 km GOES-16 ABI LST, while the grey lines are the LST of the 400 pixels in the generated 100 m LST).

For urban heat island (UHI) studies, the 100 m, hourly LST data can also serve as an important data source. For example, when LST images with short revisit times, but coarse spatial resolution (e.g., 2 km) are applied to UHI monitoring, considering the great heterogeneity of urban areas, they may fail to represent the complicated texture of urban LST. On the contrary, when LST images with fine spatial resolution, but long revisit times (e.g., 1 day or longer) are applied to UHI monitoring, the diurnal LST change cannot be represented. As urban LST normally varies greatly within a day, the diurnal LST variation tends to be an important part of UHI monitoring. Thus, considering the great spatial and temporal heterogeneity of urban areas, the 100 m, hourly LST product generated in this research can provide valuable data support for UHI monitoring [2], [55], [56]. Many other applications require 100 m, hourly LST (or LST with finer spatial and temporal resolutions), such as the monitoring of fire spread, crop growth and flood inundation, all of which may change greatly during a single day. The generated 100 m, hourly data make fine spatial resolution, diurnal variation monitoring possible, and increases the spatial and temporal accuracy for related monitoring.

D. The uncertainty in reconstruction of 100 m, hourly LST time-series

The results in Section IV-*F* validated the effectiveness of the spatio-temporal fusion scheme in producing 100 m, hourly LST for a long period of one month. The strategy, however, has several limitations in the case involving a long period. First, the error in the original GOES-16 ABI LST images can impact the accuracy of the 100 m, hourly LST time-series generation. It is noticed in Fig. 13(a) that there always exist larger errors near the extreme points. This phenomenon can be connected to the accuracy for the predicted hourly data of the separate three days for Region 2 in Fig. 9. To investigate this issue, the accuracy of

the GOES-16 ABI LST images was also evaluated based on the corresponding ground measurements, as shown in Fig. 15. Generally, when the GOES-16 ABI LST images have great consistency with the ground measurements, the performance of the STF_L_G prediction is also satisfying. Conversely, when there is a large error in the GOES-16 ABI LST images (e.g., near the extreme points), the accuracy of the STF_L_G prediction also decreases. Second, acquisition of timely Landsat LST data tends to be a challenging issue. It is acknowledged that the performance of spatio-temporal fusion tends to be compromised when the time interval between the known and prediction time is large. Due to the long revisit time (i.e., 16 days) and frequent cloud cover for Landsat LST images it can be difficult to acquire temporally close Landsat LST images (e.g., the time of known Landsat LST in Region 5 ranges for months), making spatio-temporal fusion challenging in this case.

For the uncertainty caused by data, it would be important to further enhance the data correction scheme for GOES-16 ABI LST data (e.g., construct a fitting model for better description of the relationship between Landsat and GOES LST, coupled with LST data from other satellites with greater temporal resolution). Furthermore, for the data acquisition issue, it will be of great value to take full advantage of the effective information in the observed Landsat LST time-series and include other data to decrease the uncertainty in prediction. For example, in current schemes, the Landsat data with spatial gaps (e.g., by cloud contamination) are usually abandoned, although they contain part of the valuable information temporally closer to the prediction time. It would be interesting to fill the gaps in the observed Landsat LST first by spatial reconstruction methods. Then, the 100 m, hourly LST for the given days can be obtained by fusing with the corresponding GOES-16 ABI LST images. It should be noted that the uncertainty caused by spatial reconstruction should also be considered in the generation of 100 m, hourly LST time-series.

Furthermore, the spatio-temporal fusion models also contain uncertainty. It is acknowledged that a number of spatio-temporal fusion models (e.g., STARFM) are performed based on the assumption that the temporal changes in the fine and coarse spatial resolution images are consistent. Due to the difference in the scale, however, they are also different. It is essentially an ill-posed problem to recover the fine spatial resolution temporal changes from the coarse spatial resolution temporal changes, especially in regions with great heterogeneity. This issue, however, is common in spatio-temporal fusion, which needs to be tackled in future.

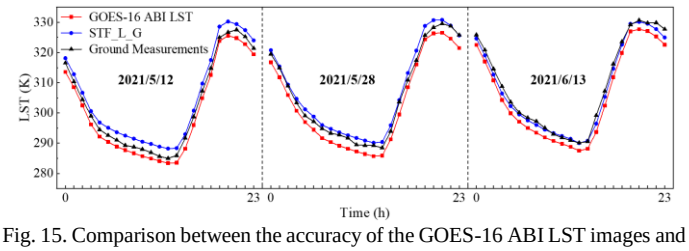


Fig. 15. Comparison between the accuracy of the GOES-16 ABI LST images and the STF_L_G prediction (Region 2 as an example).

VI. Conclusion

This paper evaluates the feasibility of applying spatio-temporal fusion for generating 100 m, hourly LST images, by fusion of the newly released 2 km, hourly GOES-16 ABI LST and 100 m, 16-day Landsat LST. The STARFM method was identified as the most accurate spatio-temporal fusion method in the fusion task in this paper. Furthermore, we compared the performance of the satellite sensor TIR bands-derived LST (i.e., GOES-16 ABI LST) and the LSM (i.e., ERA5 LST) in the fusion with the Landsat LST data. Compared to LSMs that have been applied widely for fine spatial and temporal LST generation, the LST data derived from satellite sensors (i.e., Landsat and GOES) have the unique advantage of observing LST variation directly and accurately. Generally, the 100 m, hourly LST images generated by spatio-temporal fusion can provide fine spatial resolution, diurnal LST monitoring, and have great potential to provide valuable data sources for applications requiring fine spatial and temporal resolutions to deal with fine spatial and temporal heterogeneity (e.g., UHI variation, crop growth, fire spread and flood inundation). The conclusions of this paper are summarized as follows.

1) 100 m, hourly LST can be generated by applying spatio-temporal fusion to blend 100 m Landsat and 2 km, hourly GOES-16 ABI LST images.
2) STARFM was demonstrated to be a more suitable method to generate 100 m, hourly LST compared to VIPSTF-SW and FSDAF methods.
3) Compared to the fusion of Landsat and ERA5 LST data, the fusion of Landsat and GOES-16 ABI LST data can produce 100 m, hourly LST with greater accuracy.
4) Data correction can decrease the difference between images from two data sources, and increase the accuracy of 100 m, hourly LST generation effectively.

References

[1] A. Jia, H. Ma, S. Liang, and D. Wang, "Cloudy-sky land surface temperature from VIIRS and MODIS satellite data using a surface energy balance-based method," *Remote Sensing of Environment*, vol. 263, pp. 112566, 2021.

[2] A. Siddiqui, G. Kushwaha, B. Nikam, S. K. Srivastav, A. Shelar, and P. Kumar, "Analysing the day/night seasonal and annual changes and trends in land surface temperature and surface urban heat island intensity (SUHII) for Indian cities," *Sustainable Cities and Society*, vol. 75, pp. 103374, 2021.

[3] X. Zheng, Z.-L. Li, T. Wang, H. Huang, and F. 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*, vol. 268, pp. 112774, 2022.

[4] Q. Meng, L. Zhang, Z. Sun, F. Meng, L. Wang, and Y. Sun, "Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: A 12-year case study in Beijing, China," *Remote Sensing of Environment*, vol. 204, pp. 826-837, 2018.

[5] H. Shen, L. Huang, L. Zhang, W. Penghai, and C. Zeng, "Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: a 26-year case study of the city of Wuhan in China," *Remote Sensing of Environment*, vol. 172, pp. 109-125, 2016.

[6] M. J. Wooster, G. J. Roberts, L. Giglio, D. P. Roy, P. H. Freeborn, L. Boschetti, C. Justice, C. Ichoku, W. Schroeder, D. Davies, A. M. S. Smith, A. Setzer, I. Csiszar, T. Strydom, P. Frost, T. Zhang, W. Xu, M. C. de Jong, J. M. Johnston, L. Ellison, K. Vadrevu, A. M. Sparks, H. Nguyen, J. McCarty, V. Tanpipat, C. Schmidt, and J. San-Miguel-Ayanz, "Satellite remote sensing of active fires: History and current status, applications and future requirements," *Remote Sensing of Environment*, vol. 267, pp. 112694, 2021.

[7] D. Pflugmacher, W. B. Cohen, and R. E. Kennedy, "Using Landsat-derived disturbance history (1972–2010) to predict current forest structure," *Remote Sensing of Environment,* vol. 122, pp. 146-165, 2012.

[8] T. Hu, L. J. Renzullo, A. I. J. M. van Dijk, J. He, S. Tian, Z. Xu, J. Zhou, T. Liu, and Q. Liu, "Monitoring agricultural drought in Australia using MTSAT-2 land surface temperature retrievals," *Remote Sensing of Environment,* vol. 236, 2020.

[9] F. Xie, and H. Fan, "Deriving drought indices from MODIS vegetation indices (NDVI/EVI) and Land Surface Temperature (LST): Is data reconstruction necessary?," *International Journal of Applied Earth Observation and Geoinformation,* vol. 101, pp. 102352, 2021.

[10] J. Ma, J. Zhou, F.-M. Göttsche, S. Liang, S. Wang, and M. Li, "A global long-term (1981–2000) land surface temperature product for NOAA AVHRR," *Earth System Science Data,* vol. 12, no. 4, pp. 3247-3268, 2020.

[11] M. Kilibarda, T. Hengl, G. B. M. Heuvelink, B. Gräler, E. Pebesma, M. Perčec Tadić, and B. Bajat, "Spatio‐temporal interpolation of daily temperatures for global land areas at 1 km resolution," *Journal of Geophysical Research: Atmospheres,* vol. 119, no. 5, pp. 2294-2313, 2014.

[12] D. Long, L. Yan, L. Bai, C. Zhang, X. Li, H. Lei, H. Yang, F. Tian, C. Zeng, X. Meng, and C. Shi, "Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach," *Remote Sensing of Environment,* vol. 246, pp. 111863, 2020.

[13] M. Rodell, P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll, "The global land data assimilation system," *Bulletin of the American Meteorological Society,* vol. 85, no. 3, pp. 381-394, 2004.

[14] P. Wu, Z. Yin, C. Zeng, S. B. Duan, F. M. Gottsche, X. Li, X. Ma, H. Yang, and H. Shen, "Spatially continuous and high-resolution land surface temperature product generation: a review of reconstruction and spatiotemporal fusion techniques," *IEEE Geoscience and Remote Sensing Magazine,* vol. 9, no. 3, pp. 112-137, 2021.

[15] F. Xu, J. Fan, C. Yang, J. Liu, and X. Zhang, "Reconstructing all-weather daytime land surface temperature based on energy balance considering the cloud radiative effect," *Atmospheric Research,* vol. 279, pp. 106397, 2022.

[16] H. Shen, X. Li, Q. Cheng, C. Zeng, G. Yang, H. Li, and L. Zhang, "Missing information reconstruction of remote sensing data: a technical review," *IEEE Geoscience and Remote Sensing Magazine,* vol. 3, no. 3, pp. 61-85, 2015.

[17] W. Zhao, and S.-B. Duan, "Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data," *Remote Sensing of Environment,* vol. 247, pp. 111931, 2020.

[18] W. Zhan, Y. Chen, J. Zhou, J. Wang, W. Liu, J. Voogt, X. Zhu, J. Quan, and J. Li, "Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats," *Remote Sensing of Environment,* vol. 131, pp. 119-139, 2013.

[19] M. Stathopoulou, and C. Cartalis, "Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation," *Remote Sensing of Environment,* vol. 113, no. 12, pp. 2592-2605, 2009.

[20] C. Yoo, J. Im, D. Cho, Y. Lee, D. Bae, and P. Sismanidis, "Downscaling MODIS nighttime land surface temperatures in urban areas using ASTER thermal data through local linear forest," *International Journal of Applied Earth Observation and Geoinformation,* vol. 110, pp. 102827, 2022.

[21] A. Jia, S. Liang, and D. Wang, "Generating a 2-km, all-sky, hourly land surface temperature product from Advanced Baseline Imager data," *Remote Sensing of Environment,* vol. 278, pp. 113105, 2022.

[22] S. Mallick, and T. A. Jones, "Impact of adaptively thinned GOES-16 all-sky radiances in an ensemble Kalman filter based WoFS," *Atmospheric Research,* vol. 277, pp. 106304, 2022.

[23] S. Pestana, and J. D. Lundquist, "Evaluating GOES-16 ABI surface brightness temperature observation biases over the central Sierra Nevada of California," *Remote Sensing of Environment,* vol. 281, pp. 113221, 2022.

[24] S. Qin, K. Wang, G. Wu, and Z. Ma, "Variability of hourly precipitation during the warm season over eastern China using gauge observations and ERA5," *Atmospheric Research,* vol. 264, pp. 105872, 2021.

[25] B. Chen, and B. Huang, "Comparison of spatiotemporal fusion models: a review," *Remote Sensing,* vol. 7, no. 2, pp. 1798-1835, 2015.

[26] H. Zhang, J. M. Chen, B. Huang, H. Song, and Y. Li, "Reconstructing seasonal variation of Landsat vegetation index related to leaf area index by fusing with MODIS data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 7, no. 3, pp. 950-960, 2014.

[27] H. Zhang, B. Huang, M. Zhang, K. Cao, and L. Yu, "A generalization of spatial and temporal fusion methods for remotely sensed surface parameters," *International Journal of Remote Sensing,* vol. 36, pp. 4411-4445, 2015.

[28] M. Belgiu, and A. Stein, "Spatiotemporal image fusion in remote sensing," *Remote Sensing,* vol. 11, no. 7, pp. 818, 2019.

[29] X. Zhu, F. Cai, J. Tian, and T. K.-A. Williams, "Spatiotemporal fusion of multisource remote sensing data: literature survey, taxonomy, principles, applications, and future directions," *Remote Sensing,* vol. 10, no. 4, pp. 527, 2018.

[30] J. Zhou, J. Chen, X. Chen, X. Zhu, Y. Qiu, H. Song, Y. Rao, C. Zhang, X. Cao, and X. Cui, "Sensitivity of six typical spatiotemporal fusion methods to different influential factors: a comparative study for a normalized difference vegetation index time series reconstruction," *Remote Sensing of Environment,* vol. 252, pp. 112130, 2021.

[31] X. Zhu, J. Chen, F. Gao, X. Chen, and J. G. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sensing of Environment,* vol. 114, no. 11, pp. 2610-2623, 2010.

[32] Q. Wang, Y. Tang, X. Tong, and P. M. Atkinson, "Virtual image pair-based spatio-temporal fusion," *Remote Sensing of Environment,* vol. 249, pp. 112009, 2020.

[33] T. Hilker, M. A. Wulder, N. C. Coops, J. Linke, G. McDermid, J. G. Masek, F. Gao, and J. C. White, "A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS," *Remote Sensing of Environment,* vol. 113, no. 8, pp. 1613-1627, 2009.

[34] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, no. 8, pp. 2207-2218, 2006.

[35] R. Zurita-Milla, J. Clevers, and M. E. Schaepman, "Unmixing-based Landsat TM and MERIS FR data fusion," *IEEE Geoscience and Remote Sensing Letters,* vol. 5, no. 3, pp. 453-457, 2008.

[36] Q. Wang, K. Peng, Y. Tang, X. Tong, and P. M. Atkinson, "Blocks-removed spatial unmixing for downscaling MODIS images," *Remote Sensing of Environment,* vol. 256, pp. 112325, 2021.

[37] C. Gevaert, and F. García-Haro, "A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion," *Remote Sensing of Environment,* vol. 156, pp. 34-44, 2015.

[38] M. Wu, Z. Niu, C. Wang, C. Wu, and L. Wang, "Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model," *Journal of Applied Remote Sensing,* vol. 6, no. 1, pp. 063507, 2012.

[39] H. Song, and B. Huang, "Spatiotemporal reflectance fusion via sparse representation," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 10, pp. 3707-3716, 2012.

[40] M. Das, and S. K. Ghosh, "Deep-STEP: a deep Learning approach for spatiotemporal prediction of remote sensing data," *IEEE Geoscience and Remote Sensing Letters,* vol. 13, no. 12, pp. 1984-1988, 2016.

[41] X. Zhu, E. H. Helmer, F. Gao, D. Liu, J. Chen, and M. A. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," *Remote Sensing of Environment,* vol. 172, pp. 165-177, 2016.

[42] X. Li, G. M. Foody, D. S. Boyd, Y. Ge, Y. Zhang, Y. Du, and F. Ling, "SFSDAF: An enhanced FSDAF that incorporates sub-pixel class fraction change information for spatio-temporal image fusion," *Remote Sensing of Environment,* vol. 237, pp. 111537, 2020.

[43] M. Liu, W. Yang, X. Zhu, J. Chen, X. Chen, L. Yang, and E. H. Helmer, "An Improved Flexible Spatiotemporal DAta Fusion (IFSDAF) method for producing high spatiotemporal resolution normalized difference vegetation index time series," *Remote Sensing of Environment,* vol. 227, pp. 74-89, 2019.

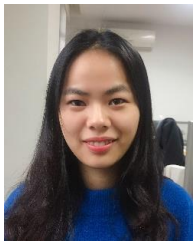
[44] Q. Weng, P. Fu, and F. Gao, "Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data," *Remote Sensing of Environment,* vol. 145, pp. 55-67, 2014.

[45] Z. Yin, P. Wu, G. M. Foody, Y. Wu, Z. Liu, Y. Du, and F. Ling, "Spatio-temporal fusion of land surface temperature based on a convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 59, no. 2, pp. 1808-1822, 2021.

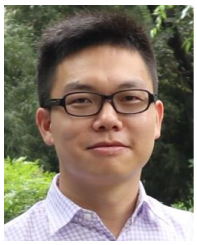
[46] B. Huang, J. Wang, H. Song, D. Fu, and K. Wong, "Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring," *IEEE Geoscience and Remote Sensing Letters,* vol. 10, no. 5, pp. 1011-1015, 2013.

[47] A. K. Inamdar, A. French, S. Hook, G. Vaughan, and W. Luckett, "Land surface temperature retrieval at high spatial and temporal resolutions over the southwestern United States," *Journal of Geophysical Research: Atmospheres,* vol. 113, no. D7, 2008.

[48] K. Zaksek, and K. Ostir, "Downscaling land surface temperature for urban heat island diurnal cycle analysis," *Remote Sensing of Environment*, vol. 117, pp. 114-124, 2012.
[49] P. Wu, H. Shen, L. Zhang, and F.-M. Göttsche, "Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature," *Remote Sensing of Environment*, vol. 156, pp. 169-181, 2015.
[50] J. Quan, W. Zhan, T. Ma, Y. Du, Z. Guo, and B. Qin, "An integrated model for generating hourly Landsat-like land surface temperatures over heterogeneous landscapes," *Remote Sensing of Environment*, vol. 206, pp. 403-423, 2018.
[51] J. Ma, H. Shen, P. Wu, J. Wu, M. Gao, and C. Meng, "Generating gapless land surface temperature with a high spatio-temporal resolution by fusing multi-source satellite-observed and model-simulated data," *Remote Sensing of Environment*, vol. 278, pp. 113083, 2022.
[52] G. Rios, and P. Ramamurthy, "A novel model to estimate sensible heat fluxes in urban areas using satellite-derived data," *Remote Sensing of Environment*, vol. 270, pp. 112880, 2022.
[53] B. Bell, H. Hersbach, A. Simmons, P. Berrisford, P. Dahlgren, A. Horányi, J. Muñoz‐Sabater, J. Nicolas, R. Radu, D. Schepers, C. Soci, S. Villaume, J. R. Bidlot, L. Haimberger, J. Woollen, C. Buontempo, and J. N. Thépaut, "The ERA5 global reanalysis: Preliminary extension to 1950," *Quarterly Journal of the Royal Meteorological Society*, vol. 147, no. 741, pp. 4186-4227, 2021.
[54] S.-B. Duan, Z.-L. Li, H. Li, F.-M. Göttsche, H. Wu, W. Zhao, P. Leng, X. Zhang, and C. Coll, "Validation of Collection 6 MODIS land surface temperature product using in situ measurements," *Remote Sensing of Environment*, vol. 225, pp. 16-29, 2019.
[55] J. Ma, S. Chen, X. Hu, P. Liu, and L. Liu, "Spatial-temporal variation of the land surface temperature field and present-day tectonic activity," *Geoscience Frontiers*, vol. 1, no. 1, pp. 57-67, 2010.
[56] D. Streutker, "Satellite-measured growth of the urban heat island of Houston, Texas," *Remote Sensing of Environment*, vol. 85, no. 3, pp. 282-289, 2003.



YijieTang received the B.S. degree from Nanjing Normal University, Nanjing, China, in 2019. She is currently pursuing the Ph.D. degree in Tongji University, Shanghai, China. Her research interests focus on remote sensing image fusion.



Qunming Wang received the Ph.D. degree from the Hong Kong Polytechnic University, Hong Kong, in 2015.

He is currently a Professor with the College of Surveying and Geo-Informatics, Tongji University, Shanghai, China. He was a Lecturer (Assistant Professor) with Lancaster Environment Centre, Lancaster University, Lancaster, U.K., from 2017 to 2018, where he is currently Visiting Professor. His 3-year Ph.D. study was supported by the hypercompetitive Hong Kong Ph.D. Fellowship and his Ph.D. thesis was awarded as the Outstanding Thesis in the Faculty. He has authored or coauthored 90 peer-reviewed articles in international journals such as *Remote Sensing of Environment*, *IEEE Transactions on Geoscience and Remote Sensing*, and *ISPRS Journal of Photogrammetry and Remote Sensing*. His research interests include remote sensing, image processing, and geostatistics.

Professor Wang is an Editorial Board member of *Remote Sensing of Environment*, and serves as Associate Editor for *Science of Remote Sensing* (sister journal of *Remote Sensing of Environment*) and *Photogrammetric Engineering & Remote Sensing*. He was Associate Editor for *Computers and Geosciences* (2017−2020).

Xiaohua Tong received the Ph.D. degree in traffic engineering from Tongji University, Shanghai, China, in 1999.

He is currently a Professor with the College of Surveying and Geoinformatics, Tongji University. He is an Academician of Chinese Academy of Engineering. He was a Research Fellow with Hong Kong Polytechnic University, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, between 2008 and 2009. His research interests include remote sensing, geographic information system, uncertainty and spatial data quality, and image processing for high-resolution and hyperspectral images.

Peter M. Atkinson received the Ph.D. degree from the University of Sheffield (NERC CASE award with Rothamsted Experimental Station) in 1990. More recently, he received the MBA degree from the University of Southampton in 2012.

He is currently Distinguished Professor of Spatial Data Science and Dean of the Faculty of Science and Technology at Lancaster University, UK. He was previously Professor of Geography at the University Southampton, where he is currently Visiting Professor. He is also Visiting Professor at the Chinese Academy of Sciences, Beijing. He previously held the Belle van Zuylen Chair at Utrecht University, the Netherlands, is a recipient of the Peter Burrough Award of the International Spatial Accuracy Research Association and is a Fellow of the Learned Society of Wales. The main focus of his research is in remote sensing, geographical information science and spatial (and space-time) statistics applied to a range of environmental science and socio-economic problems. He has published over 300 peer-reviewed articles in international scientific journals and around 50 refereed book chapters. He has also edited nine journal special issues and eight books.

Professor Atkinson is Editor-in-Chief of *Science of Remote Sensing*, a sister journal of *Remote Sensing of Environment*. He also sits on the editorial boards of several further journals including *Geographical Analysis*, *Spatial Statistics*, and *Environmental Informatics*.