Spatially Continuous and High-Resolution Land Surface Temperature Product Generation

©SHUTTERSTOCK.COM/TRUEVISIONER

A review of reconstruction and spatiotemporal fusion techniques

PENGHAI WU, ZHIXIANG YIN, CHAO ZENG, SI-BO DUAN, FRANK-MICHAEL GÖTTSCHE, XIAOSHUANG MA, XINGHUA LI, HUI YANG, AND HUANFENG SHEN

Remotely sensed land surface temperature (LST) with spatial continuity and high spatiotemporal resolution (hereafter referred to as *high resolution*) is a crucial parameter for studying the thermal environment and has important applications in many fields. However, adverse atmospheric conditions, sensor malfunctioning, and scanning gaps between orbits frequently introduce spatial discontinuities into satellite-retrieved LST products. For a single sensor, a tradeoff occurs between temporal and spatial resolutions; therefore, it is almost impossible to obtain images in high resolution.

Recent advances in the reconstruction and spatiotemporal fusion of LST products have promoted active research to address this limitation. These methods have been applied most often in thermal remote sensing studies and have led to the development of numerous algorithms. However, to the best of our knowledge, no review exists in the literature that expatiates and summarizes the available LST reconstruction and spatiotemporal fusion algorithms. Therefore, this article introduces the principles and theories behind LST reconstruction and spatiotemporal fusion and provides an overview of the published research and algorithms.

METHODS FOR RETRIEVING LST

LST is a key parameter in environmental monitoring and ecological processes [1]–[3] and is also valuable in the study of heat-related health issues and urban heat islands [4]. From a climate perspective, LST is a significant index for evaluating land–atmosphere exchanges of energy and reveals dynamic changes in surface temperature both globally and in key areas [5]. Satellite remote sensing has provided unique and efficient approaches for obtaining LST over a vast region with acceptable revisit cycles (i.e., temporal resolution) and complete spatial coverage [6], [7].

The methods for retrieving satellite-based LST can be categorized into two types: thermal infrared (TIR)- and passive microwave (PMW)-based methods. The retrieval of LST from remotely sensed TIR data has attracted abundant attention

owing to the data's relatively fine spatial resolution and proper accuracy [6]. Various operational LST products can be obtained using effective TIR-based retrieval algorithms [6]. For example, LST products are available free of charge from the *FengYun-2* (*FY-2*) Visible Infrared Spin Scan Radiometer (VISSR) [8], *FY-3* Visible Infrared Scanning Radiometer (VIRR), Moderate-Resolution Imaging Spectroradiometer (MODIS) [7], and Meteosat Second-Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) [9].

However, the spatial continuity of LST products is strongly affected by pixels with invalid values caused by clouds or cloud shadows, hereafter referred to as *cloudy pixels*. On average, clear-sky conditions account for only about one-third of the daily weather worldwide, with lower percentages occurring in humid regions [10]. In particular, about two thirds of MODIS LST data are contaminated by clouds [11]. Furthermore, some cloud-free but naturally bright pixels are frequently misclassified as cloud coverage, and their corresponding LSTs are set to the missing pixel value [12]. In addition, missing information caused by defective TIR sensors, such as the *Landsat* Enhanced Thematic Mapper Plus (ETM+) scan line corrector (SLC)-off data [13], also introduces spatial discontinuity into LST products. Thus, spatial discontinuity is a common limitation in LST obtained from TIR data owing to cloud coverage, incorrect algorithm data, defective sensors, and other problems.

PMW radiation, which can penetrate nonprecipitating clouds, is minimally affected by atmospheric absorption and water vapor [14], [15]; therefore, the LST retrieved using this method has high spatial completeness. However, the limitations of PMW measurement lie in the existence of a wider scanning gap between orbits and lower spatial resolution compared with those of TIR measurement. The spatial resolution of PMW measurement, at 25 km for the Advanced Microwave Scanning Radiometer (AMSR) for the Earth Observing System (AMSR-E) and about 10 km for AMSR2, is much lower than that of TIR measurement. Wide gaps often exist between orbits over middle and low latitudes [16], such as the 25% gap in the AMSR-E image over the Chinese mainland. Therefore, LST with low spatial resolution retrieved from PMW measurements often includes spatial discontinuity, particularly on a global scale.

Another limitation is the tradeoff between the temporal and spatial resolutions of satellite-derived LST for a single sensor, whether TIR or PMW, which often makes it impossible to obtain LST with the high spatiotemporal resolution required by some applications [17]. Generally, LST retrieved from sensors with fine spatial resolution has poor temporal resolution, resulting in temporal discontinuity and vice versa. These two core limitations of satellite-derived LST products are summarized in Table 1, and the corresponding sample charts are shown in Figure 1(a).

The aforementioned issues severely hinder applications of LST products in many fields [18]. For example, surface urban heat islands (SUHIs) can be continuously observed with cloud-free *FY-2/3* or MODIS data; however, their associated spatial resolutions are too coarse for revealing detailed SUHI spatial patterns. Images with greater spatial detail are provided by *Landsat* and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). However, the long revisit cycles of these two products, at more than 15 days, and frequent cloud contamination cause significant differences in the LST scenes these sensors acquire in space–time observation [13]. Therefore, practical applications urgently require LST products with spatially continuous high-spatiotemporal-resolution (SCHR)-LST, as shown in Figure 1(b).

Owing to the growing number of available satellite-based LST products, many different approaches for generating SCHR-LST have been proposed, resulting in abundant research on these methods. Therefore, it is of great importance to present an overview of the state-of-the art technology in applied SCHR-LST techniques.

Although previous research of Zhan et al. includes a review on the disaggregation of LST to finer resolutions [19], it is presented from the perspective of thermal sharpening and temperature unmixing. However, several recently proposed methods (e.g., spatiotemporal fusion) for obtaining LST with finer temporal and spatial resolutions should also be surveyed. Similar to that in disaggregation, the purpose of spatiotemporal fusion is also to improve the spatiotemporal resolution of LST data using multiple sensors and multitemporal LST data without the need for auxiliary data. Spatiotemporal LST fusion has developed rapidly to become a research hotspot.

Therefore, a review of recent research conducted on this topic is needed. Furthermore, to the best of our knowledge, a thorough review of methods for deriving spatially continuous LST has not been performed thus far. The objectives of this article are to review methods for generating SCHR-LST products, describe the state-of-the-art technology, and identify the most promising research fields to obtain results that will ultimately benefit LST producers and developers of SCHR-LST algorithms.

BACKGROUND

THEORETICAL CONCEPTS

LST is affected by incoming solar radiation, cooling by wind, coastal effects, land cover, temperature inversion, and other effects [20], making the data highly variable in space and

TABLE 1. THE CORE LIMITATIONS OF SATELLITE-DERIVED LST PRODUCTS.

ISSUE	Spatial discontinuity type	Tradeoff between temporal and spatial resolution
TIR	Cloudy pixels, pixels misclassified as clouds, and defective sensors	Technical and budgetary limitations
PWM	Scanning gaps between orbits	Technical and budgetary limitations

time [21]. Satellite-derived LST is based on Planck's law [22], which states that total radiative energy emitted by the ground surface or subsurface increases with an increase in temperature.

Using the radiative transfer equation and assuming a cloudless atmosphere in local thermodynamic equilibrium, the channel infrared radiance I_{obs} received by a sensor at the top of the atmosphere can be written as [6]

$$I_{\text{obs}}(\lambda,\theta) = \tau_\lambda(\theta)(\varepsilon_\lambda B(\lambda, T_s) + [1 - \varepsilon_\lambda](R_{at\downarrow}(\lambda,\theta) + R_{sl\downarrow}(\lambda,\theta))) \\ + R_{at\uparrow}(\lambda,\theta) + R_{sl\uparrow}(\lambda,\theta) \tag{1'}$$

where λ is the wavelength, θ is the satellite viewing angle, ε_λ is the surface emissivity, $B(\lambda, T_s)$ is the Planck function describing radiance emitted at surface temperature (T_s), $R_{sl\downarrow}$ is the downward solar diffusion radiance, $R_{at\downarrow}$ is the downwelling thermal irradiance, $\tau_\lambda(\theta)$ is the atmospheric transmittance, $R_{at\uparrow}$ is the upward atmospheric thermal radiance, and $R_{sl\uparrow}$ is the upward path solar radiance.

Equation (1) shows that retrieval of LST from satellite TIR data depends on surface parameters such as emissivity, surface conditions such as heterogeneous and nonisothermal factors, and atmospheric effects such as water vapor and clouds [23]. These are responsible for the lack of clarity in the physical meaning of satellite-derived LST, whereas the atmospheric effects make TIR-based retrieval algorithms favorable only for images observed under clear-sky conditions [24]. The definition of satellite-derived LST for heterogeneous and nonisothermal surfaces is still controversial [23]. Although several classical retrieval methods, such as split-window, mono-window, and single-channel algorithms, have been widely used to generate LST data with acceptable accuracy [6], invalid values of cloud pixels have been observed in LST images.

Conversely, PMW data penetrate clouds, making all-sky subsurface temperature retrieval possible in cloudless and cloudy conditions based on various approaches, such as empirical and physics-based methods and neural networks. However, such techniques have ambiguous physical mechanisms and assumptions regarding the land surface emissivity and atmospheric effects, which reduces the feasibility of measuring and the accuracy of subsurface temperatures [6]. Moreover, subsurface temperatures retrieved from PMW data show an average value of the soil temperature from the land surface to a particular depth beneath the surface and differ from TIR LST obtained at several microns of depth. A conversion model developed to consider the thermal conductivity equation of soil would offer strong potential for the PMW LST data to match the LST value of TIR images.

As an important variable, LST can also be derived from the surface energy balance (SEB) equation by identifying the surface-related energy components [10]. Over the land surface, the SEB formula is applied as follows:

$$G = S_n - F_n - H - LE, \tag{2}$$

where G is the ground heat flux, S_n is the net solar (or shortwave) radiation derived from incoming shortwave radiation minus outgoing shortwave radiation, F_n reflects upwelling longwave radiation minus downwelling longwave radiation, and H and LE denote the sensible and latent heat flux, respectively.
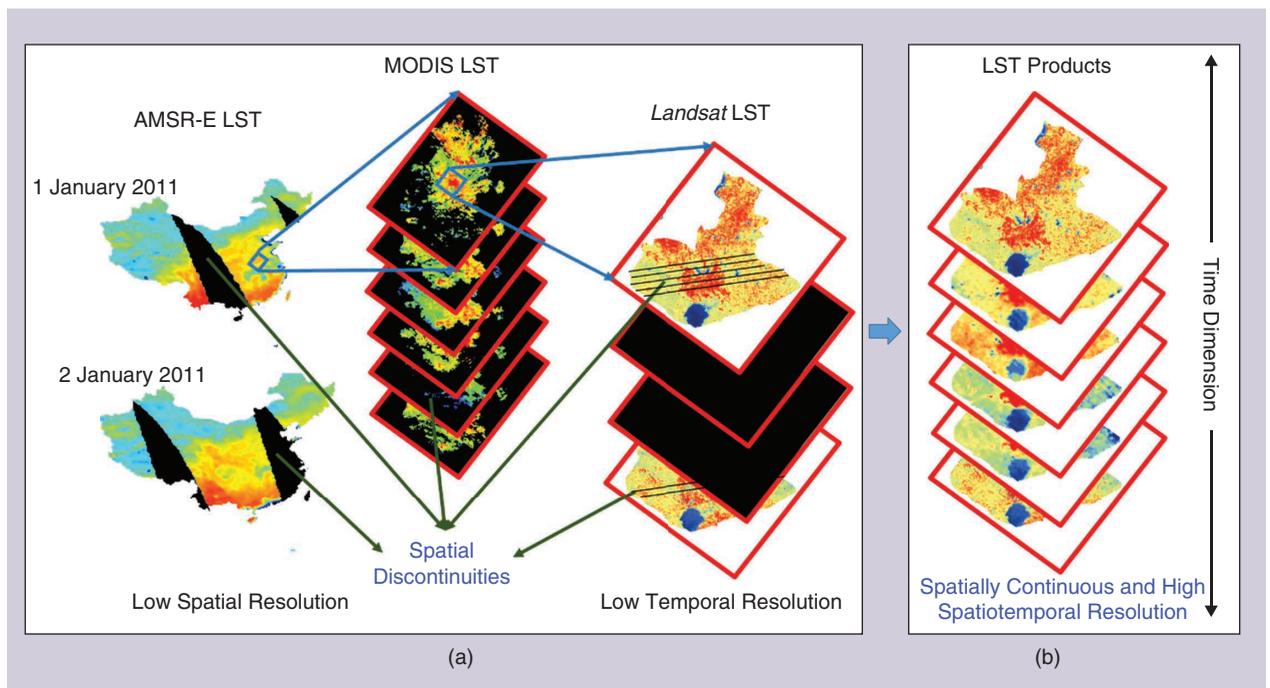

FIGURE 1. (a) A comparison of spatially discontinuous LST data and that of low spatial and temporal resolution. (b) An example of spatially continuous LST at high spatiotemporal resolution.

F_n, H, and LE are functions of LST and other variables, such as surface emissivity, air density, wind speed, humidity, and air temperature. LST can be derived from SEB [10] based on the knowledge of surface and atmospheric conditions. Because surface variables cannot be measured by satellite, the satellite data require substantial work before they can be used alone to derive LST [10].

CURRENT SATELLITE-BASED LST PRODUCTS

Satellite-derived LST estimation has improved significantly during the last few decades [6], resulting in LST products with acceptable accuracy, and the demand for remotely sensed LST data has increased in numerous applications. Operational LST products retrieved from different satellite sensors are freely available from various organizations.

Notably, significant projects have been funded to provide an accurate global view of LST in recent decades. These projects have developed techniques for merging archived data from various satellites into a combined long-term satellite record for climates, such as the European Space Agency LST Climate Change Initiative project for GlobTemperature data (http://data.globtemperature.info/); NASA Making Earth System Data Records for Use in Research Environments project;

and Research and Development of the System for Global Land Surface Parameter Product Set and Product Generation project (phase II), managed by the Ministry of Science and Technology of the People's Republic of China.

Newly released LST products can also be obtained from various satellites, such as ASTER onboard the *Terra* satellite, the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the *Suomi National Polar-Orbiting Partnership* (*S-NPP*) satellite, the Advanced Baseline Imager (ABI) onboard the *Geostationary Operational Environmental Satellite* (*GOES*)-R satellite, *Global Land Surface Satellite* (*GLASS*) products [25], and the Sea and Land Surface Temperature Radiometer (SLSTR) onboard the *Sentinel-3 A/B* satellites. Furthermore, some scholars provide open source software or executable code for generating *Landsat* LST products [26], [27], although consistent scene-based global *Landsat* LST values are currently provided free of charge by the U.S. Geological Survey (https://www.usgs.gov/land-resources/nli/landsat/landsat-collection-2-surface-temperature).

Other LST products, such as PMW LST, used in various publications are also available upon request from the corresponding authors [15]. Some of the commonly used operational satellite-derived LST products and their specifications are listed in Table 2.

TABLE 2. THE COMMON SATELLITE-DERIVED LST PRODUCTS AND THEIR SPECIFICATIONS.

LST DATA	SPATIAL RESOLUTION	TEMPORAL RESOLUTION	TEMPORAL EXTENT	REPORTING RESULTS OF LST PRODUCT ACCURACIES	DOWNLOAD LINK OR SOURCE
Aqua AMSR-E	25 km	One day	June 2002–Sept. 2011	~4–6 K [28]	From the corresponding author of [28]
Global Change Observation Mission AMSR2	10 km	One day	July 2012–present	~4–6 K [29]	From the corresponding author of [29]
FY-2F/2G VISSR	5 km	1 h	Nov. 2012–present	Have not been validated extensively [30]	http://satellite.nsmc.org.cn/portalsite/default.aspx
GOES Imager	4 km	hourly and 3 h	Sept. 1994–present	About 2 K [31]	https://www.ospo.noaa.gov/Products/land/glst/
MSG SEVIRI	3 km	15 min	Feb. 2005–present	Fewer than 2 K [32]	https://landsaf.ipma.pt/en/
GOES ABI	2 km	15 min	2018–present	2.5 K in accuracy in most cases of the validation [33]	https://www.star.nesdis.noaa.gov/goesr/product_land_lst.php
Terra/Aqua MODIS	1 and 5 km	One day	June 2000–present	Fewer than 1 K over homogeneous surfaces [34]	https://ladsweb.modaps.eosdis.nasa.gov/search/
GLASS	1 km and 0.05°	One day	1983, 1993, 2003, and 2013	2 K for *GLASS*-MODIS LST and 3 K for *GLASS*-Advanced Very High Resolution Radiometer LST [25]	http://glass.umd.edu/LST/
FY-3A/3B/3C VIRR	1 km	One day	Aug. 2009–present	Fewer than 2 K [35]	http://satellite.nsmc.org.cn/portalsite/default.aspx
Sentinel-3 SLSTR	1 km	One day	Feb. 2016–present	Fewer than 2 K [36]	https://eoportal.eumetsat.int/
S-NPP VIIRS	750 m	One day	Nov. 2011–present	~1.2 K at zenith angles of ≤40°	https://viirsland.gsfc.nasa.gov/Products/NASA/LSTESDR.html
Terra/ASTER	90 m	16 days	Mar. 2000–present	Fewer than 1.5 K over a selected number scenes [37], [38]	https://lpdaac.usgs.gov/products/ast_08v003/
Landsat TM/ETM+/TIRS	30 m	16 days	1984–present (North America)	Fewer than 2 K from 83% global validation data set [39], [40]	https://earthexplorer.usgs.gov/

DESCRIPTION OF THE PROBLEMS

RECONSTRUCTION

The main reason for spatially discontinuous LST is missing or cloudy pixels, which severely limits many LST applications [41]. Therefore, suitable methods for filling or amending these pixels are required. In particular, an efficient and cost-effective approach is reconstruction of these pixels by utilizing complementary information from different sources [42].

Reconstructing missing or cloudy information in satellite-derived LST data is a typical ill-posed inverse problem. To unify the terminology and explain the issue, we provide a brief introduction to the problem, describe the general concept of the reconstruction of LST data, and explain the notations and symbols.

The reconstruction of absent LST data requires the missing values to be estimated from the existing or remaining valid LST values and complementary spatial, multitemporal, spatiotemporal, or multisensor information. As shown in Figure 2, LST data $\mathbf{L}$ are given at observation time t_0, and $\mathbf{L} \in R^{m*n}$ $(\mathbf{L}; \Omega \in \mathbf{R}^2 \to \mathbf{R}^2)$, where Ω represents the spatial domain and comprises m × n points. We assume that domain Ω contains a missing or cloudy pixel region $\mathbf{MC}$ and a region $\mathbf{EC}$ with an existing and valid LST, i.e., $\Omega = \mathbf{MC} \cup \mathbf{EC}$ and $\mathbf{MC} \cap \mathbf{EC} = \mathbf{\Phi}$.

The purpose of the reconstruction is to estimate (i.e., reconstruct) an LST value at a position (x_0, y_0) located in $\mathbf{MC}$ from pixels in $\mathbf{EC}$, which could contain data from other observation times and sensors. LST reconstructions are based on reasonable assumptions and required to be visually natural and in agreement with all known thermal properties.

Given a missing LST pixel (x_0, y_0) obtained from a sensor (s_0) acquired at observation time t_0, its LST value can be reconstructed as

$$LST(x_0, y_0, t_0, s_0) = \sum_{x=1}^{m} \sum_{y=1}^{n} \sum_{t=t_0}^{t_p} \sum_{s=s_0}^{s_q} f(LST(x, y, t, s)), \quad (3)$$

where $(x_0, y_0) \in \mathbf{MC}$; $(x, y) \in \mathbf{EC}$; $f(\cdot)$ is a linear or nonlinear function denoting the relationships of all of the existing or remaining valid pixels; t is the observation time; s is the sensor; m and n are the rows and columns of the pending reconstructed LST data, respectively; and p and q are the numbers of the observation time and sensor, respectively. Complementary information from the same sensor can be exploited only if $s = s_0$, whereas spatial information can be exploited only if $t = t_0$. Furthermore, complementary information from different sensors with the same or similar observation times (e.g., t_0) can be used to reconstruct cloudy pixels by blending TIR LST (e.g., from MODIS) with PMW LST (e.g., from AMSE-R) or by combining the SEB equation and in situ data to produce spatially complete LST data sets.

Figure 2 summarizes two main groups of LST reconstruction methods: those using spatial, temporal, and spatiotemporal information and those using multisensor data. Furthermore, the reconstruction methods discussed in this review are based exclusively on satellite LST products; techniques based on TIR radiance are not covered.

SPATIOTEMPORAL FUSION

For technological or financial constraints, a tradeoff between temporal and spatial resolution always exists [43]; that is, a sensor providing LST data at fine spatial resolution has
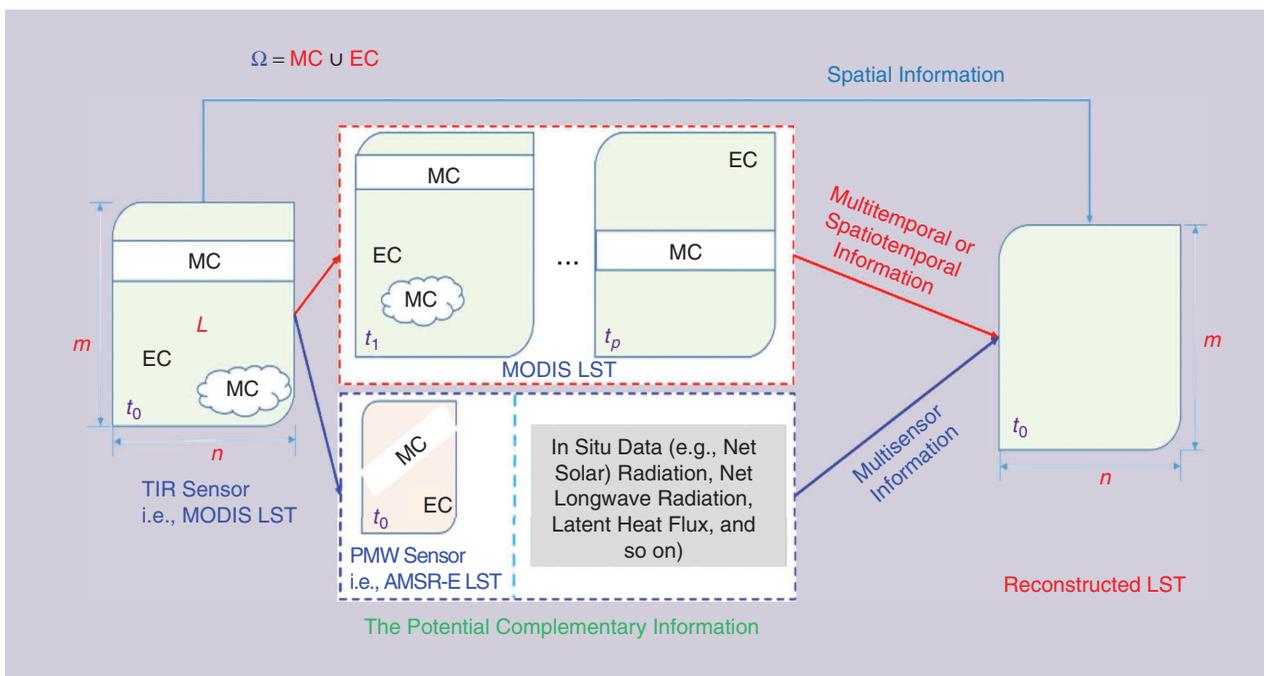

FIGURE 2. An LST reconstruction using complementary information based on spatial, multitemporal, spatiotemporal, and multisensor methods. EC: existing or cloudless region; MC: missing or cloudy region.

poor temporal resolution and vice versa. For example, geostationary satellites provide multispectral images of the observed Earth disk with high temporal resolution (e.g., 15 min) [44]. Nevertheless, the corresponding low spatial resolutions (3–5 km) limit the spatial details observed over heterogeneous landscapes [45]. Polar-orbiting sensors, such as *Landsat* Operational Land Imager and *Terra* MODIS, often provide LST at resolutions of about 100 and 1,000 m, respectively, which enables observation with greater detail in heterogeneous areas. However, their respective long revisit cycles of 16 days and a half day can shorten the duration for optimal observation, particularly over rapidly changing regions.

Several studies have reviewed methods for increasing the spatial or temporal resolution in remotely sensed LST data [19], [46]. Different terms were used to refer to the various methods for enhancing LST spatiotemporal resolution, e.g., *downscaling* [47], [48], *image fusion* [49], and *disaggregation* [19]. According to a recent study by Xia et al., these methods can be divided into two categories: kernel-driven techniques, which downscale LSTs with the reference of auxiliary data from multispectral sensors, and fusion-based approaches, which predict fine spatiotemporal resolution LSTs by blending information about temporal changes and spatial relationships obtained from other sensors [50]. Zhan et al. focused on methods dedicated to disaggregating LSTs and provided a comprehensive and systematic review of kernel-driven methods [19]. Various new satellite sensors and LST products have been presented recently (Table 1); accordingly, fusion-based approaches have advanced rapidly and attracted increasing attention. This review focuses on the development of fusion-based methods over recent decades.

It should be noted that the fusion-based approaches presented in this review are spatiotemporal fusion techniques, which differ from more traditional optical-image fusion methods, such as spatial–spectral fusion. Such traditional methods generally use only a panchromatic band to improve the detailed spatial information of their multispectral images [51].

In contrast, spatiotemporal fusion methods consider images with different spatial resolutions and acquisition times simultaneously. Furthermore, spatial–spectral fusion is often performed on raw digital numbers, whereas spatiotemporal fusion generally requires a physical parameter as input, e.g., reflectance or LST. The basic concept behind LST spatiotemporal fusion is to predict a fine-spatial-resolution LST at time t_0 using resampled LST of coarse spatial resolution at the same time as well as a fine-scale-conversion factor (SCF), as shown in Figure 3. The SCF can be obtained or learned from a fine-scale classification image or from multiple pairs of fine- and coarse-spatial-resolution LSTs of the same areas observed at various times (e.g., $t_1, t_2, \ldots t_p$). Fine-spatial-resolution LST at time t_0 thus can be expressed as

$$LST_H(i, t_0) = f(LST_L(i, t_0), \text{SCF}), \qquad (4)$$

where LST_H and LST_L represent fine- and resampled coarse-spatial-resolution LSTs, respectively, and i denotes the ith pixel.

CONNECTIONS BETWEEN RECONSTRUCTION AND SPATIOTEMPORAL FUSION

Many applications would benefit from spatially continuous LST products on a global scale, whereas SCHR-LST products would be ideal for many fine-scale applications. Figure 4 shows the relationship between reconstructed LST, spatiotemporal fusion, and the application of generated SCHR-LST. Thus, reconstruction and spatiotemporal fusion methods are often used together.

For example, Shen et al. employed reconstruction and spatiotemporal fusion methods to analyze long-term and fine-scale summer SUHIs in Wuhan City, China [13]. Lu et al. compared SUHIs during daytime (i.e., 10:30 a.m. local time) and nighttime (i.e., 12:30 a.m. local time) during different seasons in Hefei, China, based on reconstructed and fused *Landsat*-like LST data [52].

More recently, spatially complete and temporally continuous LST values were obtained for surface soil moisture mapping by combining reconstruction and spatiotemporal fusion methods [53]. In essence, we can extend the basic concept of spatiotemporal fusion by interpreting this method as a special case of reconstruction in which no auxiliary pixels are available in space at the predicted time (i.e., completely missing data). Nonetheless, this study reviews the methods separately without referring to them collectively as reconstruction methods because spatiotemporal fusion is better known and easily acceptable. Based on the low-resolution LST information at the predicted time, spatiotemporal fusion
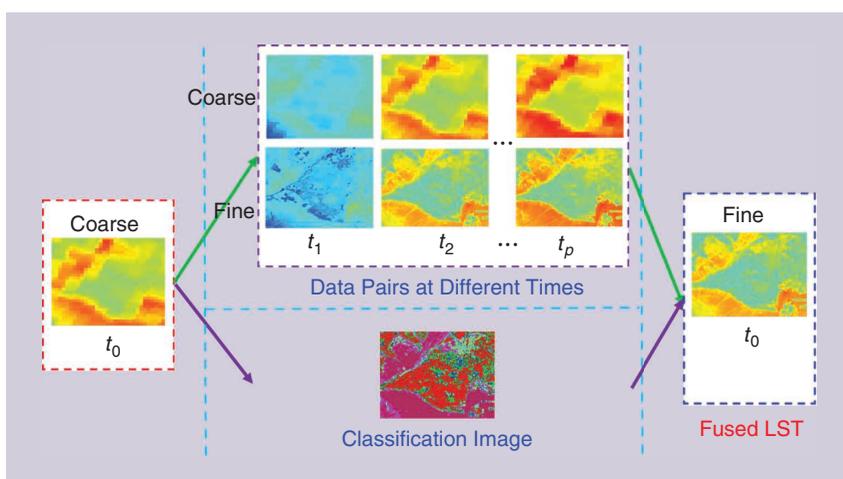

FIGURE 3. A spatiotemporal LST fusion with an SCF obtained from data pairs (fine and coarse resolution) at different times or a fine-scale classification image.

can be used to reconstruct remote sensing images affected by phenological and significant land cover changes [54]. Therefore, reconstruction and spatiotemporal fusion are closely interconnected, which is the main reason for reviewing these methods here.

RECONSTRUCTION FOR SPATIAL CONTINUITY

RECONSTRUCTION OF MISSING LST PIXELS

Although partially cloudy pixels are often treated as missing pixels, there are obvious differences when estimating LST for cloudy pixels and missing pixels from neighboring cloud-free pixels. In this review, missing pixels are regarded only as the case in which measurements are actually missing owing mainly to defective sensors (e.g., *Landsat* ETM+ SLC-off data) [13] and scanning gaps between orbits (e.g., *Aqua*/AMSR-E, Global Change Observation Mission/AMSR2 data), which can occur in TIR and PMW imagery.

Reconstruction techniques can be used to complete the missing data and enhance the usability of incomplete LST data. In the literature, several approaches have been proposed that can be roughly classified into three types [55]: 1) spatial methods, in which no additional reference information is required; 2) temporal techniques, in which the reference information of one area was obtained at different observational times; and 3) spatiotemporal approaches, in which the reference information is from additional spatial and temporal information.

An example of the spatiotemporal reconstruction of missing AMSR-E LST pixels is given in Figure 5. Proposed by Zeng et al. [56], the adopted spatiotemporal reconstruction method consists of three steps. First, a classification map should be obtained to differentiate the land surface types; the missing AMSR-E LST values of the target image (observed on 1 January 2011, in this case) can then be reconstructed using robust regression with the spatiotemporal information from other auxiliary LST images in each classification (such as those observed on 2 January 2011, in this case). Finally, a postprocessing procedure with spatial information is applied to eliminate outliers.

SPATIAL RECONSTRUCTION METHODS

Reconstruction techniques based exclusively on spatial information are the most traditional of the three approaches. These methods reconstruct missing information using the existing or remaining valid LST values and assume that they share the same statistical and geometrical structures with the missing values [57]. The most basic reconstruction methods are spatial interpolation approaches, e.g., spline function [58] and geostatistical methods. Motivated by the high spatiotemporal heterogeneity of LST, some researchers have considered more environmental variables and used multivariate interpolation methods, e.g., cokriging [59] and regression tree analysis [60].

TEMPORAL RECONSTRUCTION METHODS

Reconstruction methods based on multitemporal or time-series information use temporal images of the same region at different times to reconstruct missing pixels. The algorithms primarily employed include the linear temporal approach [61]–[63], harmonic analysis method [64], temporal Fourier analysis approach [65], wavelet transformation method [66], asymmetric Gaussian function fitting method [67], singular spectrum analysis algorithm [68], discrete cosine transform [69], diurnal temperature cycle (DTC)-based method [8], [70], [71], and multitemporal dictionary learning [72]. Multitemporal methods generally ignore data from geographically neighboring pixels and are appropriate when the differences are linked mainly to regular changes, such as those in observation conditions and
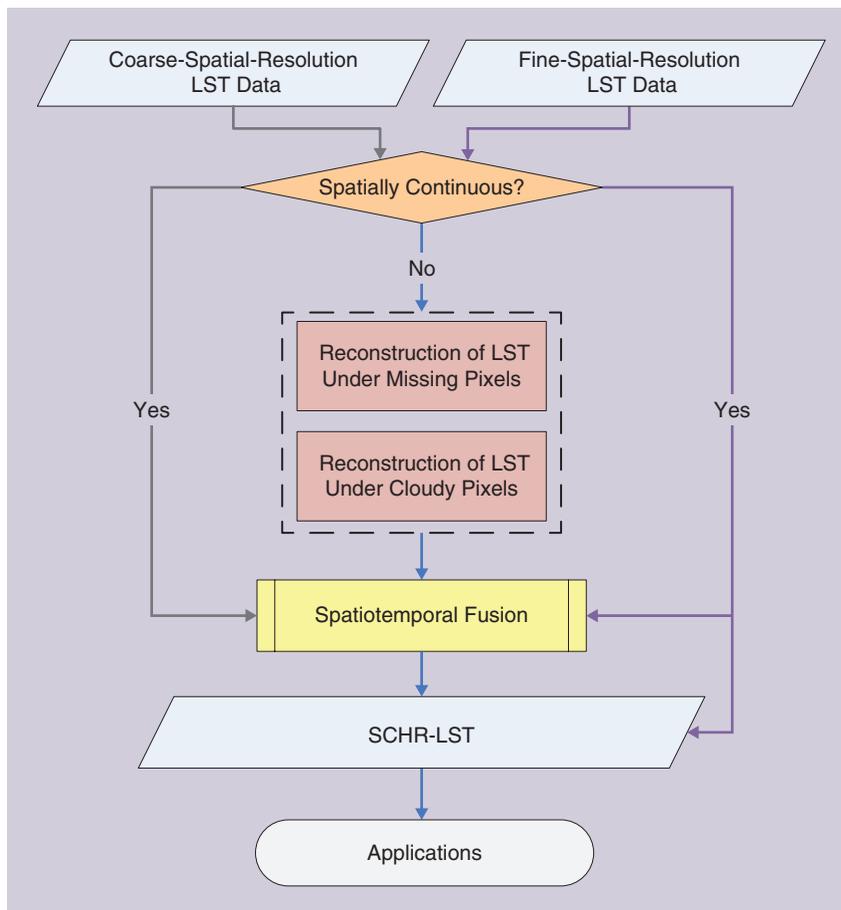

FIGURE 4. A schematic diagram showing the generation of SCHR-LST and the relationship between LST reconstruction and spatiotemporal fusion.

phenology. In contrast, abrupt changes, such as newly constructed buildings and sudden natural disasters, are difficult to reconstruct [42].

SPATIOTEMPORAL RECONSTRUCTION METHODS

Spatiotemporal approaches appear to be the most appropriate methods for reconstructing degraded remote sensing data with high variability in both space and time [55], [73]. The most basic strategy is to sequentially apply a spatial and temporal reconstruction method in which a multitemporal method is used first; a spatial reconstruction method is then applied if the first step fails to recover all missing pixels or identify abnormal pixels [56].

For example, Liu et al. introduced a spatiotemporal reconstruction framework for recovering missing pixels of *FY-2F* LST products that included multitemporal and spatial reconstruction methods [8]. Duan et al. proposed a spatiotemporal interpolation module for constructing missing AMSR-E temperature data caused by orbital gaps between satellite overpasses [24]. For LST image time series, Weiss et al. proposed a gap-filling method using neighboring valid data and data from other times, i.e., different calendar dates or multiannual data sets [74]. Pham et al. reported good results in applying a 3D gap-filling method to nine years of LST data observed over Australia between 2002 and 2011 [75].

Recently, Wang et al. proposed a new method for estimating MODIS or *Landsat* LST under cloudy skies that used solar–cloud–satellite geometry (SCSG) [76]. Their scheme required multitemporal or synthetic reference LST data (e.g., MOD11A2 averaged eight-day LST products) and a neighboring clear-sky region to calculate a ratio based on the SCSG.

Moreover, some hybrid methods that combine spatial or temporal approaches with spatiotemporal techniques have been proposed. These hybrid methods utilize the advantages of the individual approaches but avoid their disadvantages [77]. Many studies have demonstrated that hybrid and spatiotemporal methods usually yield lower errors [78].

A COMPARISON OF METHODS FOR RECONSTRUCTING MISSING PIXELS

Figure 6 shows reconstruction results for simulated missing pixels using different reconstruction methods. The original image in Figure 6(a) is a *Terra* MODIS LST image acquired on 10 October 2010, and Figure 6(b) is an LST image that simulates the missing pixels from Figure 6(a). Figure 6(c)–(e) shows the *Terra* MODIS LST images acquired on 4 September 2010, 20 September 2010, and 6 October 2010, respectively, which were used to reconstruct the missing pixels from Figure 6(b). The result of the spatial reconstruction method, shown in Figure 6(f), includes severe artifacts, particularly in
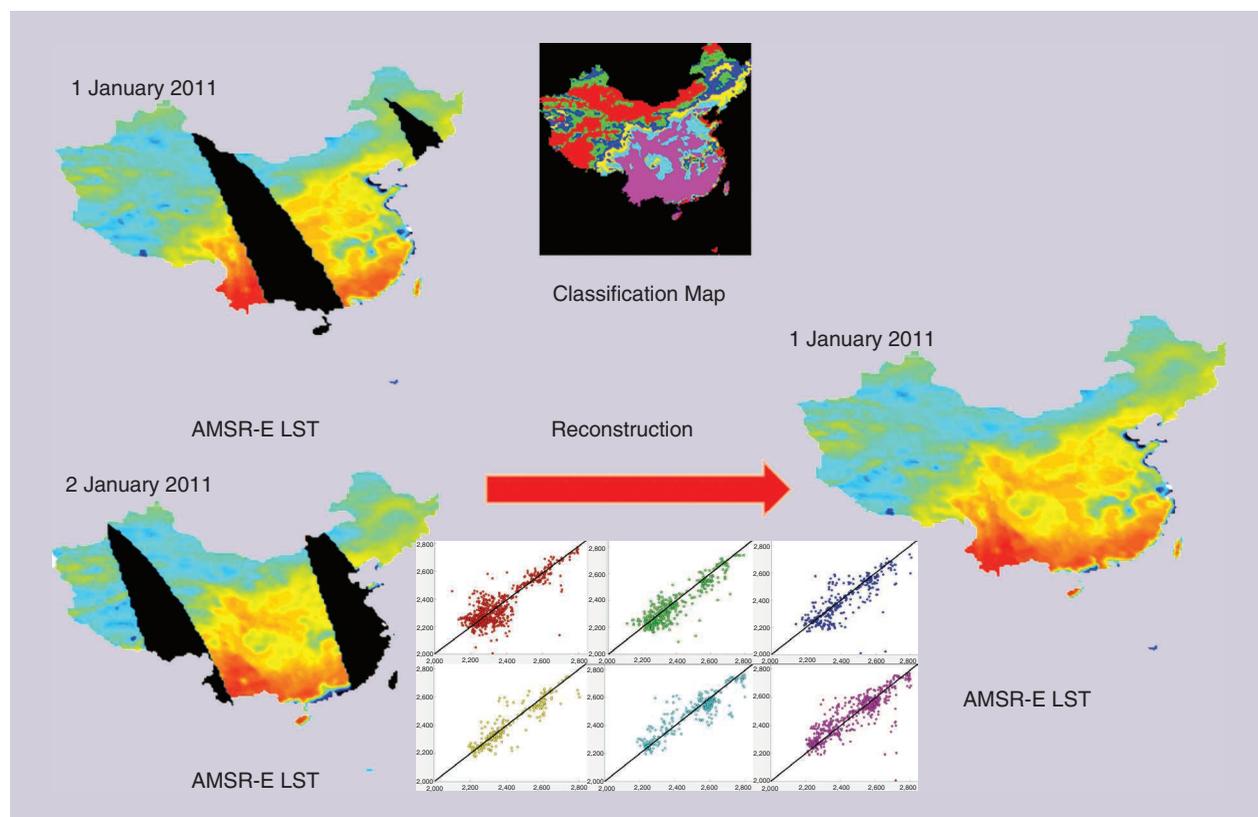

FIGURE 5. The spatiotemporal reconstruction of missing AMSR-E LST pixels. The target (observed on 1 January 2011) and multitemporal AMSR-E LST images (observed on 2 January 2011) are used to reconstruct the missing LST pixels for generating a spatially continuous AMSR-E LST (1 January 2011) [56]. The scatterplots between the target and multitemporal AMSR-E LST images correspond to those on the classification map.
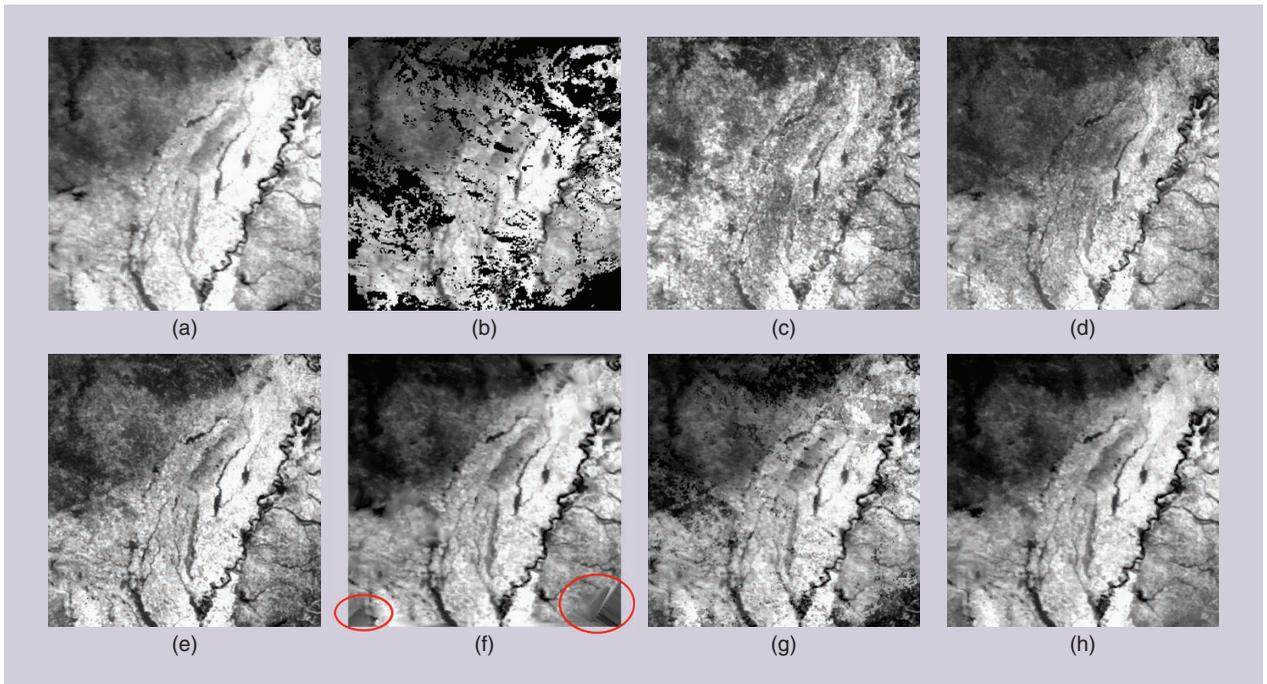
FIGURE 6. The simulations conducted for reconstructing missing pixels using different methods. (a) A *Terra* MODIS LST acquired on 10 October 2010. (b) A simulated LST image with missing pixels from (a). (c)–(e) *Terra* MODIS LSTs acquired on (c) 4 September 2010, (d) 20 September 2010, and (e) 6 October 2010. The reconstruction results followed the (f) spatial reconstruction method in [58], (g) temporal reconstruction approach in [63] using MODIS LSTs in (c)–(e), and (h) spatiotemporal reconstruction technique in [56] using the same auxiliary data as those employed in the temporal reconstruction method. The regions marked with red ovals show that the spatial reconstruction method brings severe artifacts.

the regions marked with a red ellipse. These images demonstrate that simple interpolation does not give a satisfactory result when the image data contain abundant missing values.

The results of the temporal reconstruction method are shown in Figure 6(g). Fewer artifacts are observed in this approach than in the spatial reconstruction. However, the temporal reconstruction method introduced apparent errors caused by temporal inconsistencies. When using the spatiotemporal reconstruction technique, shown in Figure 6(h), all missing pixels were reconstructed with no visible artifacts or errors. This experiment shows that spatiotemporal reconstruction methods are preferable for reconstructing missing pixels in LST images.

The quantitative assessment of the three methods given in Table 3 further demonstrates the superiority of the spatiotemporal reconstruction techniques over other reconstruction approaches. The correlation coefficient (CC) and root-mean-square error (RMSE) values of all filled pixels in

Figure 6 are provided for an overall quantitative evaluation. The temporal reconstruction method gave the worst result, whereas the best result was obtained using the spatiotemporal reconstruction approach, which is consistent with the findings in [78].

An experiment was conducted to illustrate the effects of the different methods using AMSR-E LST images that included with missing pixels caused by scanning gaps between orbits. The AMSR-E LST images with missing pixels acquired on 1–4 January 2020 are shown in Figure 7(a)–(d), and the results of reconstructing Figure 7(a) by spatial, temporal, and spatiotemporal methods are shown in Figure 7(e)–(g). The results obtained are similar to those of the simulated experiment. The spatial method presented inferior results, including abundant artifacts, and the temporally reconstructed pixels were not satisfactory in the boundary regions; the spatiotemporal approach gave the best visual effect.

APPLICATIONS AND LIMITATIONS OF THE RECONSTRUCTION METHODS

The primary limitations of the spatial, temporal, and spatiotemporal reconstruction approaches are discussed in this section. The spatial reconstruction technique is generally easy to implement at a minimum computational cost and is suitable for homogeneous images with a small number of

TABLE 3. THE CC AND RMSE VALUES OF THE RECONSTRUCTION IMAGES IN FIGURE 6.

INDICATOR	SPATIAL RECONSTRUCTION METHOD IN [58]	TEMPORAL RECONSTRUCTION METHOD IN [63]	SPATIOTEMPORAL RECONSTRUCTION METHOD IN [56]
CC	0.9828	0.976627	0.98686
RMSE (ºC)	1.34	2.95	0.9

missing pixels. However, it provides limited spatial information, blurred LST images with abundant missing pixels in heterogeneous landscapes, and unsatisfactory accuracy, as shown in Figure 6(f) and Figure 7(e).

The temporal reconstruction method is suitable for LST images with abundant pixels and adequate multitemporal information. However, this approach is highly sensitive to temporal inconsistency, such as that caused by land cover change or weather, as shown in Figure 7(f). Furthermore, selecting or designing a model to fit this multitemporal information can be challenging.

The spatiotemporal reconstruction method is appropriate for regions with abundant missing pixels. The resultant image has high accuracy and acceptable computational cost. However, this method underutilizes the available temporal and spatial information. Maximizing the use of all available temporal and spatial information is the main concept of the spatiotemporal reconstruction methods, which show promise.

It is noteworthy that all of the mentioned methods, in principle, are applied for recovering missing pixels in LST data. In practice, however, these methods are often used to reconstruct LST pixels under cloudy conditions [8], [73], [78] even though they can provide only hypothetical LST values under cloudless conditions. These hypothetical values are generally different from LSTs actually observed under cloudy conditions.

RECONSTRUCTING LSTs FOR CLOUDY PIXELS

Cloudy-sky conditions lead to abundant cloudy pixels in TIR LST data; thus, it is highly desirable to develop suitable methods for reconstructing cloud-contaminated LSTs. Numerous approaches have been presented specifically for reconstructing cloudy-pixel LSTs [10], [12], [14], [24], [79]–[81]. These reconstruction approaches can be categorized into two general groups. The first involves PMW data observed by satellite sensors (e.g., AMSR-E), which also retrieve the LST under clouds. The second contains SEB-based methods, in which additional assumptions or meteorological conditions are indispensable for estimating the errors between cloud-free and cloudy areas. An example of a PMW-based reconstruction of cloudy MODIS LST (MYD11A1) pixels is shown in Figure 8.

PMW-BASED METHODS

Given the different advantages, techniques for blending data from TIR and PMW sensors have strong potential for generating spatially continuous LST data sets with high accuracy and moderate spatial resolution [79]. Recently, several blending methods have been presented for TIR and PMW LST. Indian scholars utilized artificial neural networks to generate a spatially continuous MODIS LST data set under nonclear-sky conditions based on the microwave polarization difference index, MODIS LST, and other ancillary data [14]. Duan et al. proposed a promising framework for retrieving all-weather MODIS LST by integrating data from AMSR-E and MODIS sensors onboard the *Aqua* satellite [24].

Using a method based on the Bayesian maximum entropy theory, some scholars combined MODIS and AMSR-E LST data obtained during nighttime and daytime over different terrains (e.g., the Tibetan Plateau and Heihe River Basin) and heterogeneous landscape types [79], [82]. Their method is shown to be feasible for reconstructing all-sky LSTs with the combined TIR and microwave LST products.

However, the challenges of low retrieval accuracy and coarse spatial resolution of PMW data in this method have inspired new concepts. For example, Zhou et al. proposed a thermal sampling depth correction method to estimate the PMW LST over barren land, which had the potential for generating PMW LST with the same physical meaning and similar accuracy as those of TIR LST [83]. Holmes et al. modeled and quantified the systematic differences between PMW and TIR LST by considering both radiative transfer and sensing depth [84].

Given that presently used PMW LST downscaling methods do not fully consider the rapid time-varying characteristic of LST, Zhang et al. proposed a novel method for blending
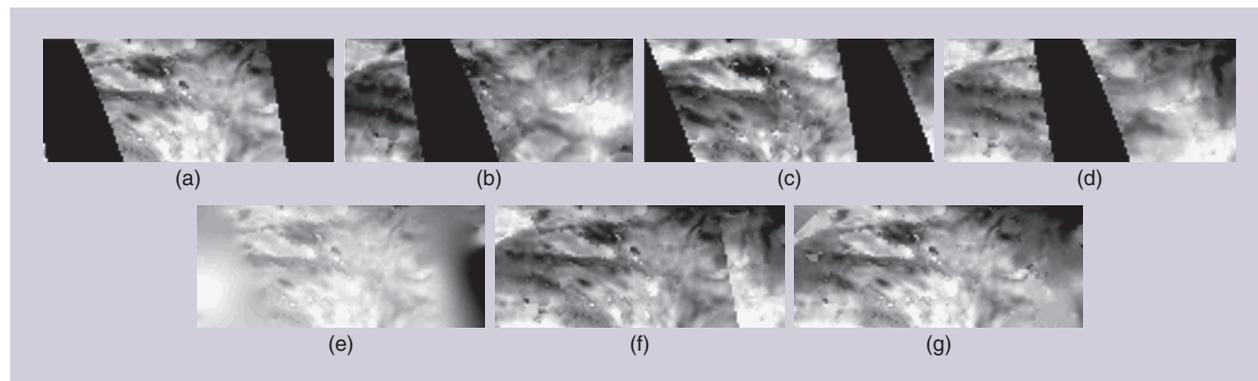

FIGURE 7. AMSR-E LST images with missing pixels (scanning gaps). (a)–(d) AMSR-E LST images with missing pixels acquired on 1–4 January 2020. The reconstruction results followed the (f) spatial reconstruction method in [58], (g) temporal reconstruction method in [63] using the AMSR-E LSTs in (b)–(d), and (e) spatiotemporal reconstruction method in [56] using the same auxiliary data as those used in the temporal reconstruction method.

TIR and PMW data based on the decomposition of LST in the temporal dimension, such as annual, diurnal, and real-time weather components [85]. These findings present a significant step forward in the global merger of PMW and TIR LST in generating all-weather LST products. In summary, coarse-spatial-resolution PMW data can be regarded as reference information for available TIR LST (e.g., MYD11A1 products) or other auxiliary data when retrieving spatially complete LSTs with fine spatial resolution [14], [24], [79], [83], [85].

SEB-BASED METHODS
Unlike the aforementioned techniques, which are based directly on remote sensing data, SEB-based approaches utilize models of the physical processes controlling the various surface parameters. In 2000, Jin creatively put forward an SEB-based neighboring-pixel (NP) method for obtaining cloudy-pixel LST values [10]. In this approach, the cloudy-pixel LSTs were interpolated from neighboring clear pixels and auxiliary in situ data (e.g., net solar radiation, net longwave radiation, and latent heat flux). Inspired by Jin's research, a temporal method [86], spatiotemporal method [87], and revised NP method [12] were proposed to recover cloud-covered pixels of LSTs from MSG/SEVIRI data or MOD11A1 products.

In addition, some two-step frameworks were developed to recover satellite-based LSTs contaminated by clouds [81], [88], which reduced the strong dependence on ground-based measurement. In that method, the cloudless theoretical LSTs were calculated in the first step, and an SEB-based procedure was then used to correct the filled values in the second step. A novel method was recently proposed for reconstructing LSTs in cloud-covered areas using integrated MODIS/*Terra* land products and MSG geostationary satellite data [89]. One main advantage of the method is that it considers the cumulative effect of solar radiation on LST rather than only instantaneous differences in solar radiation based on MSG time-series observations.

In addition, Martins et al. applied the SEB-based method to derive all-weather MSG/SEVIRI LST products, which provided the foundation for the first all-weather LST product generated by the Land Surface Analysis Satellite Application Facility (LSA-SAF) [90]. On the basis of numerical (i.e., the urban canopy model) and atmospheric (i.e., weather research and forecasting) method, Fu et al. used information related to the structure and form of urban buildings to calculate the SEB and proposed a physical model-based technique for estimating cloudy pixels of urban MODIS LSTs [80]. In summary, the actual LST values of cloudy pixels can be recovered with the help of auxiliary data by combining SEB- and NP-based or model-based methods.

A two-step LST data fusion framework was recently proposed for generating MODIS-like LSTs under all-weather conditions [91], given that spatially complete and temporally continuous LSTs of coarse spatial resolution can be obtained from the China Land Data Assimilation System (CLDAS). This innovative work maximizes the potential of MODIS LST retrievals and CLDAS LST and attempts to create synergy between reconstruction and spatiotemporal fusion.

A COMPARISON OF RECONSTRUCTION METHODS FOR CLOUDY PIXELS
Few studies provide a comparison of SEB- and PMW-based methods because they require different auxiliary data sets (e.g., from PMW sensors or in situ measurements) or models with abundant parameters, which are difficult to obtain simultaneously. Although Fu et al. concluded that the physical model-based approach was more effective than SEB- and PMW-based methods for reconstructing urban LSTs [80], their conclusion was not backed up by experimental results and quantitative evaluation.

To the best of our knowledge, the first quantitative comparison was given recently by Long et al. [91], in which an approximate SEB-based [88] and PMW-based method [24] was selected to evaluate the proposed technique.
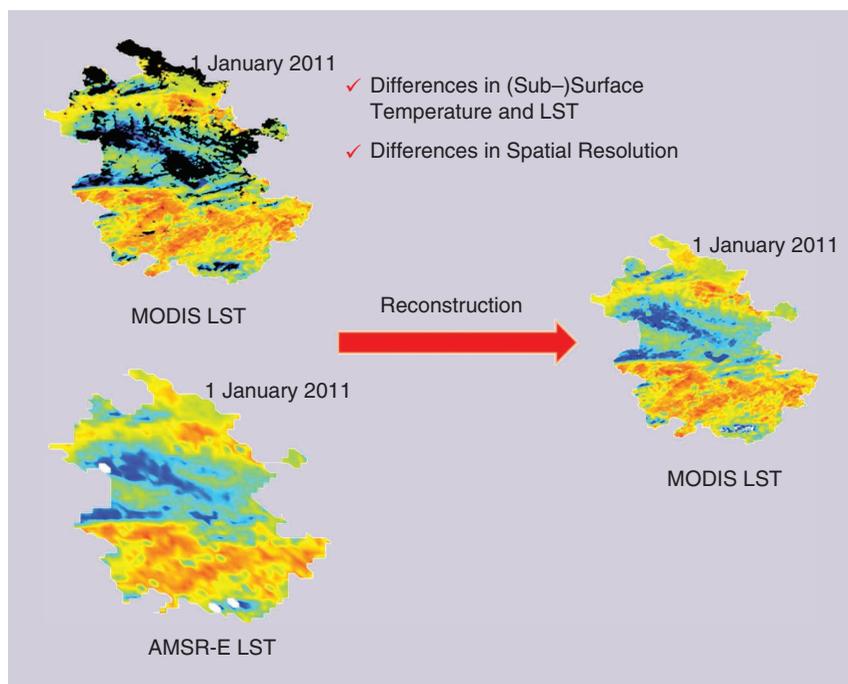

FIGURE 8. A PWM-based reconstruction of cloudy MODIS LST pixels. The target MODIS LST images (observed on 1 January 2011) and the AMSR-E LST images (observed on 1 January 2011) are used to reconstruct the cloudy MODIS LST pixels for generating spatially continuous MODIS LST (1 January 2011). PWM-based reconstruction needs to consider the differences in spatial resolution and the conversion from (sub-)surface temperatures to LST.

Hereafter, this approach is referred to as the *Long method*. Although the SEB-based technique includes multitemporal reconstruction and correction with SEB [88], Long et al. compared the LST estimated only from the multitemporal reconstruction with their results because the accuracy of the corrected LST was not improved [91].

The mean absolute error (MAE), RMSE, and coefficient of determination (CD) of the LSTs obtained from the PMW-based and Long methods were compared with in situ LSTs (comparison A), as shown in Table 4. All three indicators in the Long method were better than those in the PMW-based approach. Similarly, the three indicators of LSTs obtained from the Long method were better than those from the SEB-based technique when compared with in situ LSTs obtained at other times (comparison B).

APPLICATIONS AND LIMITATIONS OF PMW- AND SEB-BASED METHODS

The primary applications and limitations of PMW- and SEB-based approaches are discussed in this section. PMW-based techniques are suitable for LST retrievals in all-sky (i.e., cloudless and cloudy) conditions and can provide useful information for reconstructing the cloudy LSTs of TIR sensors. However, LSTs estimated from PMW have limitations in accounting for the spatiotemporal variability of microwave emissivity [92] owing to variation in land surface types. Moreover, subsurface temperatures retrieved from PMW differ from LSTs (also known as *skin temperature*) obtained from TIR sensors and require conversion to skin temperature [93].

Another limitation is that the resolution of the PMW data (e.g., 25 km) is significantly lower than that of TIR data (e.g., 100 m). These differences require consideration when fusing TIR and PMW LST products, such as recovering the LST at the spatial resolution of TIR data when a microwave pixel is partly cloudy. Therefore, high-precision, physics-based retrieval algorithms; effective conversion models; and an appropriate downscaling algorithm are urgently required in PMW-based reconstruction methods [24], [83]–[85].

SEB-based techniques provide a new approach for reconstructing the cloudy LSTs of TIR sensors without considering the differences in spatial resolution and conversion between subsurface temperature and LST. However, these methods require valid NPs from spatial, temporal, or spatio-temporal techniques [87], which are difficult to obtain in humid regions with consecutive rain events. Furthermore, these approaches make assumptions about environmental variables and require specific meteorological and hydrological observations to calculate LST errors between cloudless and cloudy areas, thereby introducing more error [80].

In summary, the PMW- and SEB-based methods can be used to effectively estimate LSTs under all-weather conditions compared with the spatial, temporal, and spatiotemporal reconstruction approaches. However, their limitations hinder their further application.

SPATIOTEMPORAL FUSION FOR HIGH RESOLUTION

By far, the most widely used spatiotemporal fusion methods employ the spatial and temporal adaptive reflectance fusion model (STARFM) and its enhanced version (ESTARFM) [43], [94]. Although these were originally proposed for surface reflectance, they can be applied to other environmental variables, such as the normalized difference vegetation index, leaf area index, and LSTs. Over the past decade, various spatiotemporal fusion methods have been proposed for obtaining LSTs with high spatiotemporal resolution, including 1) weighted function-based, 2) unmixing-based, 3) hybrid, and (4) learning-based methods. An example of weighted function-based spatiotemporal fusion methods is shown in Figure 9.

TABLE 4. A QUANTITATIVE EVALUATION OF THE RECONSTRUCTION METHODS FOR CLOUDY PIXELS.

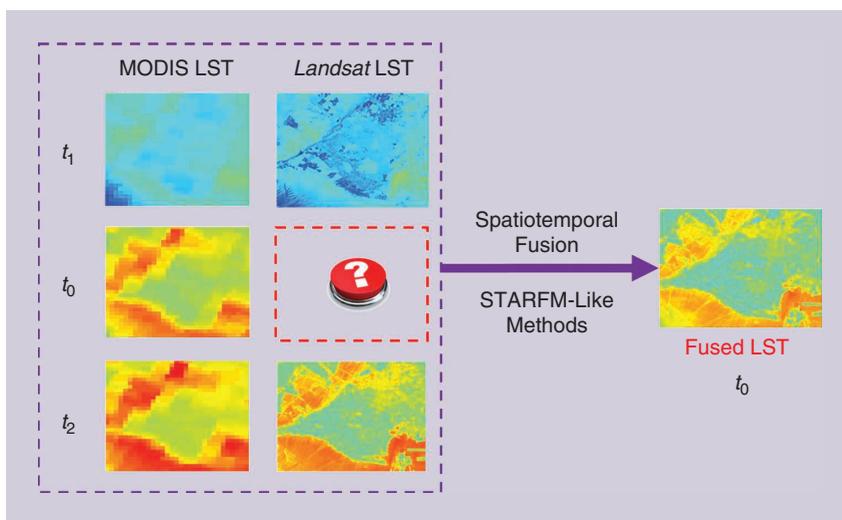| INDICATORS	COMPARISON A		COMPARISON B	
	PMW-BASED METHOD IN [24]	LONG METHOD [91]	LONG METHOD [91]	SEB-BASED METHOD IN [88]
MAE (°C)	5.96	2.31	2.49	3.85
RMSE (°C)	6.89	3.01	3.24	5.16
CD	0.8	0.91	0.93	0.81

FIGURE 9. The weighted function-based spatiotemporal fusion methods. The two pairs of MODIS and *Landsat* LST data (time t_1 and t_2) and one MODIS LST data set at the predicted time t_0 can be used to generate *Landsat*-like LST data of the predicted time t_0.

WEIGHTED FUNCTION-BASED METHODS

Since they were first proposed, weighted function-based methods, such as STARFM and ESTARFM, hereafter referred to collectively as *STARFM-like methods*, have gained considerable attention. Some studies directly adopted STARFM-like methods to generate daily fine-resolution thermal data and their derivative products, such as evapotranspiration and LST [95]–[99]. However, their performance, to a certain degree, degrades over heterogeneous areas because obvious differences exist between reflectance and LST.

STARFM-like methods with various improvements were designed for generating high-spatiotemporal-resolution LSTs. These enhancements were made in the model relationships among different-resolution pixels, weighted function design, number of sensors, and the temporal variation modeling or thermal landscape representation, which aimed to improve spatiotemporal LST patterns. With the aid of the weighted function, linear relationships were established between two MODIS LSTs of different times or a MODIS and *Landsat* LST of the same time to predict synthetic *Landsat*-scale LST data [100], [101].

Based on the STARFM framework, Huang et al. and Wu et al. designed bilateral filtering- and variation-based fusion methods, respectively, for producing high-spatiotemporal-resolution LSTs with higher accuracy [102], [103]. Weng et al. and Quan et al. modified ESTARFM to obtain LSTs with high spatiotemporal resolution by considering the annual temperature cycle (ATC) and urban thermal landscape heterogeneity [104], [105], respectively. Wu et al. proposed a spatiotemporal integrated temperature fusion model (STITFM) for generating fine-spatiotemporal-resolution LSTs with more than two sensors [17]. Weighted function-based methods generally have shown good performances in the literature and can be improved further by considering complex surface heterogeneity and rapid changes in the surface type.

UNMIXING-BASED METHODS

Unmixing-based methods regard temporal variation at a low-spatial-resolution pixel as a mixture of some component temporal variations at the corresponding high-spatial-resolution pixels. As a pioneering work, an unmixing-based fusion idea to combine multisensor satellite images observed at different times with different resolution was proposed [106].

Generally speaking, unmixing-based methods have two main challenges: serious errors associated with the unmixing process and a lack of within-class variability of the fine-scale information inside a coarser pixel [46]. Subsequently, unmixing-based approaches were designed to address these challenges. For example, specific constraints were incorporated into a unmixing process to ensure that the obtained reflectance values were positive and within a physically meaningful range [107]. To account for land cover change, a spatial and temporal data fusion approach (STDFA) and its modified version (ESTDFA) were put forward that use two or more image pairs with a constant and an adaptive-moving window, respectively [108], [109].

Huang and Zhang described an unmixing-based fusion technique capable of considering phenological differences and changes in the surface type [110]. However, owing to the complexity of temperature unmixing, these methods are often applied to reflectance data but not to LST data. Wu et al. estimated LSTs with high spatiotemporal resolution for environmental-process monitoring using STDFA directly and compared the results with those obtained by STARFM and ESTARFM [111]. Their results indicated that, although all three methods captured the spatiotemporal change information of the LSTs, they also showed performance differences, with ESTARFM, STDFA, and STARFM performing best to worst, respectively. In summary, unmixing-based methods are widely limited by inaccurate estimation numbers of endmembers, endmember spectral variability in multitemporal data, and spectral mixing nonlinearities. Furthermore, these techniques are challenged by land cover changes.

HYBRID METHODS

Hybrid methods usually integrate two or more techniques and contain weighted function- or unmixing-based approaches. Their purpose is to enhance the performance of data fusion by integrating the advantages of different methods. Representative approaches include the spatial and temporal reflectance unmixing model [112], flexible spatiotemporal data fusion (FSDAF) [113], the spatial–temporal remotely sensed images and land cover maps fusion model [114], bleed spatiotemporal temperature (BLEST) [105], and weighted combination of kernel-driven and fusion-based methods (CKFM) [50].

The first three methods combine ideas from unmixing theory, weighted function, Bayesian theory, and spatial interpolation. Hybrid methods can fuse reflectance images under complex conditions, e.g., heterogeneous landscapes or abrupt land cover changes [46]. Inspired by the spatiotemporal adaptive data fusion algorithm for temperature mapping [104] and FSDAF, the BLEST method combines multiple strategies for predicting hourly *Landsat*-like LSTs in heterogeneous areas. Specially, BLEST adopts weighted functions to preserve more details in LST images of fine spatial resolution. Moreover, it uses ATC and DTC models to describe nonlinear temporal patterns of LST variations between LST images of coarse resolution, and it combines the temperature-mixing model and downscaling technology to account for effects from complex and changeable landscapes.

However, assuming a linear or simple nonlinear relationship between the input and predicted LST data over complex landscapes might not be a sufficiently accurate approximation of reality. Furthermore, the literature shows that kernel-driven methods can capture abundant detail of the TIR band from visible bands of finer spatial resolution, whereas fusion-based techniques are

favored for their high prediction ability in both spatial and temporal terms. To take full advantage of the two method types, Xia et al. proposed the weighted CKFM for obtaining downscaled LST time series [50]. However, CKFM cannot be directly implemented when the time resolution is more than one year because the long time interval can render the regression function invalid and enlarge the spatial differences, and the land cover types can change [50].

LEARNING-BASED METHODS

Learning-based methods employ machine learning algorithms to model the relationship between observed coarse–fine image pairs and then predict the unobserved fine images. These techniques were proposed following the powerful nonlinear representation ability between input and output images, which is a major limitation of the first three approaches.

Although learning-based methods are popular in the remote sensing community, very few studies have been done to obtain high-spatiotemporal-resolution LST data. In previous research by Moosavi et al., a hybrid wavelet–artificial intelligence fusion approach was used to produce LST data with high spatiotemporal resolution, where the support vector regression and 2D wavelet transform were used to handle the nonlinear nature and nonstationary properties of the *Landsat* and MODIS LST data [115].

A deep learning (DL)-based spatiotemporal temperature fusion network (STTFN) method was designed recently to generate fine-spatiotemporal-resolution LST products with the *Landsat* and MODIS data [116]. The STTFN is more flexible and intelligent than other approaches because it forms potentially complicated relationships using training data without manually designed mathematical rules.

Moreover, high-spatiotemporal-resolution LST estimation using a DL approach was also proposed by the fusion of satellite and ground sensor data, such as from the wireless sensor network [117] and automatic weather system [118]. This fusion strategy, in contrast to multisource satellite data fusion techniques, can deliver LST estimates at a higher rate than the satellite revisit frequency but introduces errors in the spatial representativeness. Other spatiotemporal methods originally designed for reflectance images also have strong potential for LST images, e.g., Bayesian-based methods [119]–[121].

A SIMPLE COMPARISON OF SPATIOTEMPORAL FUSION METHODS

The recent work of Yin et al. [116] enables intercomparison among three typical methods: two weighted function-based (i.e., ESTARFM and STITFM) and a learning-based technique (i.e., STITN). These pioneering studies are chosen because of the popularity of the methods or accessibility of the codes. Figure 9 shows predictions from the three approaches by fusing *Landsat* and MODIS data and the actual *Landsat* LST images obtained on 24 June 2011 for area

1 (comprising 1,300 × 1.300 pixels) and on 26 February 2017 for area 2 (comprising 800 × 800 pixels).

The predicted LST images obtained from STITFM and ESTARFM on 24 June 2011 were higher than the actual values, particularly for STITFM, as shown in Figure 9. The predictions from ESTARFM were physically closer to the actual LST image for 26 February 2017, although some considerably higher values were observed in the red rectangular region shown in Figure 10. However, the predicted LST images generated by STTFN were the most consistent with the actual LST images.

The quantitative evaluations and computation efficiency for the fusion results of ESTARFM, STITFM, and STTFN are summarized in Table 5. The highest structural similarity (SSIM) and lowest RMSE were recorded by the proposed STTFN, whereas lower SSIM and higher RMSE values were recorded in ESTARFM and STITFM. ESTARFM was the most time-consuming method; the time expenditure of STTFN was close to that of STITFM. For STTFN, the most computational time was spent on training.

LIMITATIONS OF DIFFERENT SPATIOTEMPORAL FUSION METHODS

In this section, we present several limitations that are interwoven but related to different spatiotemporal fusion methods, and we discuss two of the main ones.

THE INCONSISTENCY PROBLEM OF DIFFERENT-RESOLUTION LSTs

Spatiotemporal LST fusion is implemented on LST images from two or more different satellite sensors, such as *Landsat* and MODIS. These LST images have many differences in spatial resolution, atmospheric conditions, solar geometry, viewing angle, observation time, and retrieval methods. Although spatiotemporal LST fusion requires LSTs with different spatial resolution, a wide resolution gap will cause a loss of detail and lead to significant error [17]. Systematic bias among different types of LST images results from differences in atmospheric conditions and retrieval methods. This bias can be reduced through relative calibration, such as using a linear regression or histogram matching method. Fortunately, some operational algorithms for consistently generating satellite-based LSTs have been presented, such as MODIS LST products. Collection 2 scene-based consistent *Landsat* surface temperature data are currently offered by the U.S. Geological Survey free of charge.

Solar geometry, viewing angle, and observation time can play more important roles in generating the inconsistency. Usually, similar observation times and angles/geometry were selected for spatiotemporal fusion algorithms [17], [104], [105], [116]. Taking the MODIS LST as an example, however, the overpass time, which differs by as much as ±1.5 h, and the viewing zenith angle, which varies greatly, i.e., −65° to +65°, force a spatiotemporal fusion between MODIS LST and other LST measures conducted over a small area with similar observation times or geometry [17],
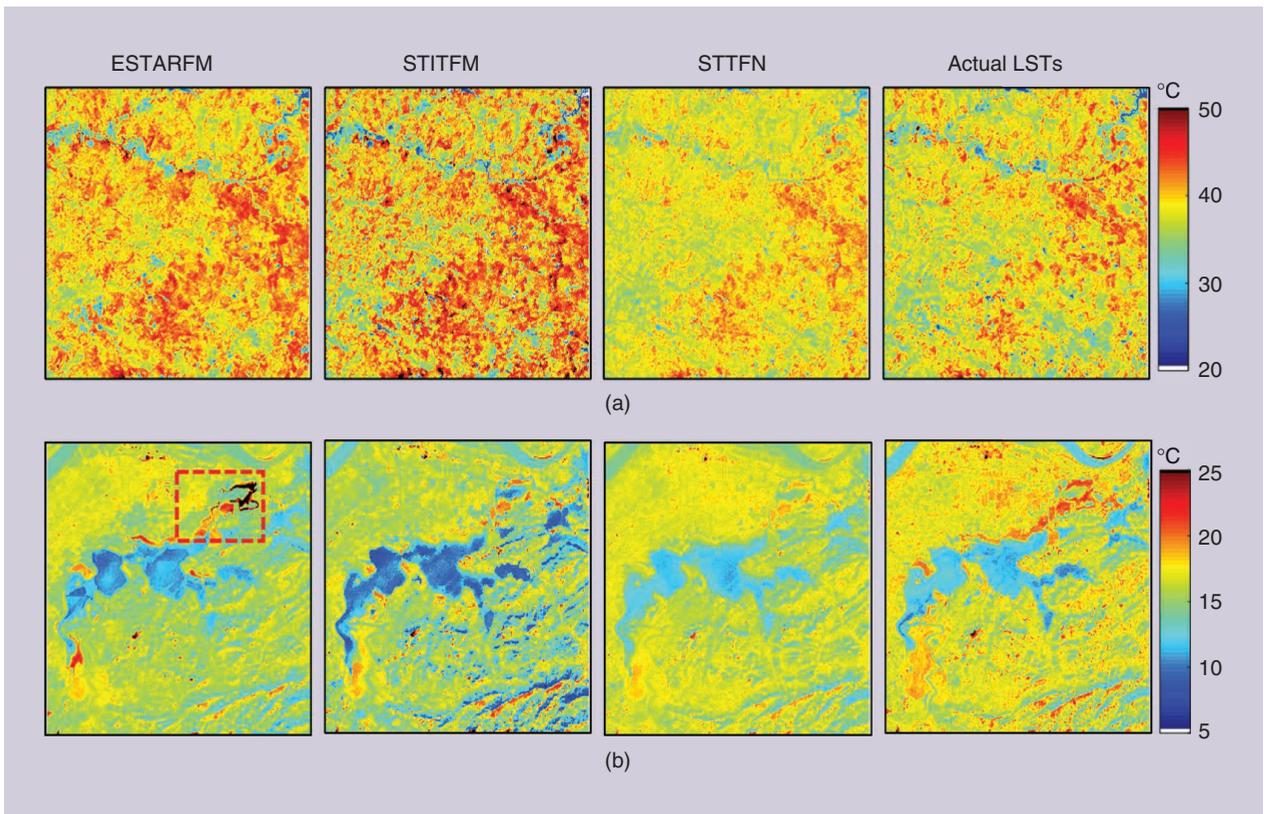
FIGURE 10. Examples of the fusion results from different methods and the actual *Landsat* LSTs of the predictions on (a) 24 June 2011 for area 1 and (b) 26 February 2017 for area 2.

[50], [105], [116]. However, correction of the inconsistency is essential when spatiotemporal LST fusion is conducted over a large area.

MODELING THE RELATIONSHIP BETWEEN OBSERVED COARSE–FINE IMAGE PAIRS

LSTs change rapidly in space as well as time owing to the heterogeneity of the land surface and changes in the surface type. Modeling the relationship between observed coarse–fine LST image pairs for spatiotemporal fusion algorithms is very challenging. In some weighted function-based LST fusion [17], [100]–[103] and hybrid methods [50], [105], it

is assumed that the output LST product can be expressed by a linear or simple nonlinear combination of inputs, which might not be appropriate if complex nonlinear temporal changes in LST are present.

Although some mathematical or physical models (i.e., ATC) have been designed to address the effect of temporal variations of LST [104], [105], these theoretical approaches are not always suitable because LST can be a highly changeable variable [116]. Moreover, the unmixing-based methods build a linear mixing equation system for linking the LST pixels in coarse and fine images even though LST is not linearly additive. Existing studies show that using a linear mixing model to link LST data at different scales might not introduce significant errors; thus, some existing methods also use the linear mixing model for fusing LST products [111].

Given their powerful nonlinear representation, learning-based methods, particularly those using DL, are advanced techniques for modeling the relations between the input and output LSTs. For example, a nonlinear relation between input and output LSTs was specified by STITN based on an integration of features extracted at different levels and fusion through a specially designed convolutional neural network

TABLE 5. A QUANTITATIVE EVALUATION FOR AREAS 1 AND 2 AND THE COMPUTATION EFFICIENCY COMPARISON FOR AREA 1.

AREA	ESTARFM		STITFM		STTFN	
	RMSE (°C)	SSIM	RMSE (°C)	SSIM	RMSE (°C)	SSIM
Area 1	2.59	0.955	2.75	0.872	**1.55**	**0.967**
Area 2	1.58	0.952	2.03	0.949	**1.08**	**0.977**
Training time (s)	—		—		170	
Predicting time (s)	718		155		8	
Total time (s)	718		155		178	

Values in bold represent the best qualitative scores.

(CNN) [116]. However, the performance of the learning-based methods relies greatly on training samples and several customized model parameters. To make the relationship modeling more reasonable, future trends should focus on modeling with small samples and adaptive parameters.

VALIDATION OF RECONSTRUCTED AND FUSED LST PRODUCTS

The validation of reconstructed and fused LST product accuracy relies mainly on simulated experimental results and cross validation with real experimental results [17]. When these findings are more similar to actual LSTs, the reconstructed or fused LST products are preferred over the original degraded LST data. Cross validation is performed mainly against ground-, other satellite-, and data assimilation-based LST data.

It remains difficult to accurately simulate LSTs under cloudy-sky conditions. For reconstructed LSTs of missing pixels and fused LSTs, simulation and cross validation can be conducted in parallel. Based on the type of validation data, four different methods are distinguished as follows. To validate the reconstructed and fused results, the recommended, feasible, and inapposite approaches are also listed in Table 6.

SIMULATED VALIDATION

Simulated validation (SiV) is the most commonly adopted validation technique. LSTs from reconstruction or fusion processing are directly compared with actual LST data [13]. The closer the agreement between the generated and actual LSTs, the better the performance of the reconstruction or fusion method. In practice, SiV is performed by first degrading the original LST image, e.g., by masking out some pixels, and then using the original LST image as a reference for the reconstructed result.

GROUND-BASED VALIDATION

The ground-based validation (GrV) approach is practicable if in situ measurements within the reconstructed or fused LST image are available. Except for a few studies that used air temperature from stations, this method usually compares the data with in situ LST observations, which have frequently been adopted to validate LST products [6], such as MODIS, *GOES*, SEVIRI, and VIIRS LST [122]–[127]. Usually, ground-based LSTs observed from in situ measurements are adopted directly to validate and evaluate the predicted LSTs obtained from spatiotemporal fusion or reconstruction methods [17], [81], [88], [102]. Notably, GrV can be used to validate reconstructed LSTs of cloudy pixels.

The three main limitations of GrV are regional restriction, spatial representativeness of the ground observations, and thermal radiation directional effects [5]. Although existing in situ networks or stations for LST reference measurements are commonly used to validate LST products—e.g., the Surface Radiation Network [123], [128] as well as validation stations of the Karlsruhe Institute of Technology [127], [129], [130] and Heihe Watershed Allied Telemetry

Experimental Research [131]—in situ reference observations remain sparse.

GrV can be performed only if in situ reference measurements are located within the reconstructed and fused LST images. Furthermore, LST measured from a station does not necessarily represent a coarser satellite sensor footprint [5]. Finally, most field measurements record data near the nadir, whereas satellite-based sensors usually have a wide field of view that can observe from the nadir to a view angle of about 60° (such as MODIS and VIIRS) [132]. Therefore, two strict criteria are usually considered for selecting valid in situ reference data: 1) networks of advanced instruments with good maintenance and upgrades and 2) strong spatial representativeness for the satellite-based remote sensing footprint [5].

OTHER SATELLITE-BASED VALIDATION

Other satellite-based validation (OsV) compares reconstructed or fused LST products with heritage LST products. This method is particularly valuable when no in situ reference measurements within the reconstructed or fused LST image are available. Moreover, this approach can also be used to validate the reconstructed LST of cloudy pixels, e.g., when LSTs from a microwave sensor are available.

However, the technique obtains only relative validation results, and intercomparisons among different satellite-based LST products are insufficient for validating reconstructed or fused LST products alone [5]. That is, LSTs retrieved from different algorithms based on similar assumptions or formulations could be in good agreement with each other but not consistent with the ground reference LST [126].

DATA ASSIMILATION-BASED VALIDATION

The data assimilation-based validation (DsV) approach compares reconstructed or fused LST products with LSTs from land data assimilation models [133]. This method is similar to OsV and can be used to validate reconstructed or fused LST products. Studies show that data assimilation-based LST products have good consistency with actual observations [91]. However, those with very low spatial resolution can cause errors when validating.

TABLE 6. THE VALIDATION METHODS AND QUANTITATIVE EVALUATION INDICATORS.

	SiV	GrV	OsV	DsV	QUANTITATIVE EVALUATION INDICATOR
Reconstruction of cloudy pixels	✕	✓✓	✓✓	✓	RMSE, MS, MAE, AAD, ME, CC, CD, SSIM, and SD
Reconstruction of missing pixels	✓✓	✓	✓	✓	
Spatiotemporal fusion	✓✓	✓	✓	✓	

AAD: average absolute difference; ME: mean error; SD: standard deviation; ✓: feasible; ✓✓: recommended; ✕: is inapposite.

QUANTITATIVE EVALUATION INDICES

It is significant to quantitatively evaluate the accuracy of the reconstruction and fusion results. Numerous quantitative indicators for evaluating reconstructed and fused results have been adopted [5], [17], [42], [78], [134]. The main evaluation indicators include the RMSE, MSE, MAE, average absolute difference, mean error, CC, CD, SSIM, and standard deviation.

It should be noted that, in this study, peer-reviewed research papers are considered; it is beyond the scope of this article to validate all of the findings of other researchers. However, experimental results from significant algorithms and their quantitative evaluation are provided here. Based on our evaluation results and the published literature, we conclude that, in addition to the reconstruction and fusion methods themselves, the validation results depend significantly on LST data as well as the distribution characteristics and number of missing LST pixels, degree of heterogeneity over the study area, and observation time [8], [78]. It is not unreasonable to assess the quality of various methods based on their individual validation results.

FUTURE PROSPECTS

Despite the achievements described, methods for obtaining SCHR-LST are currently hot research topics. However, LST reconstruction and spatiotemporal fusion are complicated and inherently ill-posed inverse problems; thus, research is still needed and provides the potential for further development. In this section, we list and discuss several topics that appear promising for improving SCHR-LST retrieval from space-based measurements.

EXPLOITATION OF GEOGRAPHICAL LAWS AND SIGNAL PROCESSING TECHNIQUES

Geographic surface parameters (e.g., LST) occur in the realm of space and time and, thus, follow geographical laws describing the behavior of their spatiotemporal autocorrelation, heterogeneity, and spatial similarity. The first two geographical laws state that geographic environmental variables 1) are spatiotemporally correlated with themselves and 2) vary in space and time. The recently proposed third law of geography concentrates more on the similarity of the geographic configuration (SGC) of locations [135]. Spatial prediction of the geographic environmental variables can be conducted on the basis of SGCs between a sample and its prediction point. However, signal processing techniques are usually employed for establishing the global numeric relationships among variables with or without insufficient consideration of geographical laws.

Nearly all current reconstruction and spatiotemporal fusion methods originate from signal processing. These approaches perform well over small and homogeneous geographic areas without significant land use or land cover changes. However, when facing large and complex geographic areas, auxiliary geographical information should be utilized. LST is a typical geographical parameter; thus, the integration of geographical laws and signal processing techniques into SCHR-LST algorithms appears promising.

INTEGRATION OF PHYSICAL PROPERTIES AND SIGNAL PROCESSING TECHNIQUES

Although input multitemporal and multisensor LST products are retrieved with physical models, obtaining reconstructed and fused LSTs often involves filtering, interpolation, regression, variational processing, and sparse representation methods. Previous studies have shown that the accuracy of LST reconstruction differs significantly between nighttime and daytime and varies with seasons [78]. Physical properties, such as high-dynamic-change characteristics, DTC, and ATC, might explain these results better than individual LST values.

Furthermore, almost all reconstruction and spatiotemporal fusion methods are implemented on LST images from multitemporal or multisensor products. However, as reviewed by Li et al., LSTs can vary between the viewing zenith angle and acquisition time (local solar time) [6]. The differences in LST measured at the nadir and off-nadir can be up to 5–10 K for different land cover types [6]. Because many polar-orbiting satellite sensors scan Earth's surface in the cross-track direction, their viewing zenith angle varies significantly (i.e., −65° to +65°); therefore, LST values of different pixels in the same orbit cannot be compared [136]. This effect must also be considered for multisource or multitemporal LST products.

In addition, LST data retrieved from the same sensors could be incomparable if their variation in observation time is significant [137], [138]. This effect is even more obvious in LST data from different satellites acquired at varying times. Inevitably, angular and temporal dissimilarities pose great challenges to reconstruction and spatiotemporal fusion. Therefore, it is necessary to integrate physical properties (i.e., DTC and ATC modeling as well as angular and temporal normalization) into reconstruction and spatiotemporal fusion methods.

NEW PROCESSING FRAMEWORK FOR RECONSTRUCTION AND SPATIOTEMPORAL FUSION

It is well known that satellite-based LSTs change rapidly in space and time [21]. The traditional signal processing framework (e.g., filtering, interpolation, regression, variational processing, and sparse representation) assumes that the input data and output results are linear or simple nonlinear transformations. It is difficult to conduct comprehensive feature mining for a complex nonlinear transformation process, nonstationary characteristics (such as high-dynamic-change characteristics), and large-scale differences between high- and low-resolution LST images.

In recent years, DL has gained the attention of the remote sensing community and been used for various image analysis and classification problems. In particular, overviews of DL for data fusion [140], [141] and reviews on DL in environmental remote sensing [142] have been published. It is

possible to simulate complex relationships among different LST images so that a trained model can be used to generate SCHR-LST [142]. Among the recently developed methods, reconstruction and spatiotemporal fusion based on the DL framework have already been reported.

For example, Zhang et al. [143] and Malek et al. [144] proposed effective CNN models to recover missing data in remote sensing images. Song et al., Tan et al., and Liu et al. proposed novel spatiotemporal fusion models using deep convolutional neural network [145], deep convolutional spatiotemporal fusion network [146] and StfNet [147] to fuse *Landsat* and MODIS reflectance data from different perspectives.

Despite its proven efficiency for reflectance images, DL has rarely been used for LST retrieval. As pioneering works, a multiscale feature connection reconstruction CNN for geostationary satellite LST images with large missing regions [148] and STTFN [116] have been proposed. These methods provide a new framework and advanced capabilities for obtaining spatially continuous and high-resolution LSTs or other remotely sensed data products. In the future, we will apply more DL studies to LST research owing to the availability of sample data of high quality.

However, imperfect knowledge of remotely sensed LSTs and low computing efficiency are key factors limiting the wide application of DL to reconstruction and spatiotemporal fusion. However, the available LST data will continue to grow and improve with the development of large data methods and further advances in remote sensing. In addition, more advanced processing strategies (i.e., parallel computing and graphics processing unit acceleration) and advanced cloud-based geospatial processing platforms (e.g., Google Earth Engine) have been put forward [149], [150], thereby addressing problems with computing efficiency. All of these developments will significantly improve the application potential of remotely sensed LST products, which is the ultimate goal of LST reconstruction and spatiotemporal fusion.

COMBINING DATA ASSIMILATION WITH RECONSTRUCTION AND SPATIOTEMPORAL FUSION

Data assimilation provides continuous information on variables and locations without direct observation by using proxy data, thereby filling gaps between sparsely distributed observations [151]. However, input uncertainty easily leads to the accumulation of errors by the model performing the assimilation, and substantial surface heterogeneity limits its application, e.g., over urban areas.

Remote sensing data can provide accurate spatial information and do not accumulate errors over time. Using data assimilation, LSTs can be combined organically with land surface process models to optimize the simulated

TABLE 7. A SUMMARY OF APPLICATION CONDITIONS, FEATURES, AND DATA SOURCES OF RECONSTRUCTION AND SPATIOTEMPORAL FUSION METHODS.

METHOD	APPLICABLE CONDITIONS AND FEATURES	APPLICABLE LST DATA SOURCES
Spatial reconstruction	Homogeneous landscapes with a small number of missing pixels High efficiency Hypothetical LST values provided under cloudless conditions	LST TIR-based from polar-orbiting sensors, such as MODIS and *Landsat* [60]
Temporal reconstruction	Regular temporal variation with larger missing regions Requires multitemporal data Provides hypothetical LST values under cloudless conditions	Polar-orbiting and geostationary satellites, such as *GOES*, SEVIRI, *FY-2F*, and others [62], [72]
Spatiotemporal reconstruction	Applicable to most cases of missing LST data High accuracy Provides hypothetical LST values under cloudless conditions	Both polar-orbiting and geostationary satellites [74], [75]
PWM-based reconstruction	Requires corresponding microwave data Estimates the LST values under actual cloudy conditions Requires consideration of differences in resolution, subsurface temperature, and LST	AMSR-E/2 data used to reconstruct MODIS LST [24], [79], [85] PWM data also have the potential to reconstruct geostationary satellite LST
SEB-based reconstruction	Makes some assumptions on environmental variables and requires specific meteorological observations Estimates the LST values under actual cloudy conditions	MODIS LST is popular [81], [88], [90] Applicable to almost all LSTs from TIR sensors
Weighted function-based fusion	Has gained considerable popularity and applications Requires cotemporal (reference time) pairs of fine- and coarse-spatial-resolution LST data in addition to coarse-spatial-resolution LST data at the prediction time	Fusion of MODIS and *Landsat* is popular [100], [103], [104] SEVIRI or *GOES* LST also used as coarse-spatial-resolution data [17]
Unmixing-based fusion	Requires coarse-spatial-resolution LST data at the prediction and reference times as well as high-spatial-resolution images near the predicted date	Fusion of MODIS and ASTER was tested [111]
Hybrid methods fusion	Integrates two or more techniques and contains weighted function- and unmixing-based methods High accuracy	Fusion of MODIS and *Landsat* was tested [50] *FY-2F* LST also used as coarse-spatial-resolution data [105]
Learning-based fusion	A promising approach with powerful nonlinear representation ability between input and output images Relies heavily on training samples and several customized model parameters	Fusion of MODIS and *Landsat* attempted [115], [116] Fusion data from satellite and ground sensors show potential

LSTs in time and space, which provides a new approach for obtaining continuous LST data. With this technique, a series of LST data can be simulated, and likely trends can be identified.

Such trend information provides a physical constraint on LST and can enhance the accuracy of LST reconstruction and spatiotemporal fusion methods. This point has recently attracted the attention of scholars. Based on the spatiotemporal fusion method and CLDAS LST data, Long et al. generated MODIS-like LSTs suitable for all-weather conditions [91]. It can be predicted that the generation of SCHR-LST using LST from data assimilation will become more common.

SYNERGIES BETWEEN RECONSTRUCTION AND SPATIOTEMPORAL FUSION

All of the reviewed LST reconstruction and spatiotemporal fusion methods have been implemented as separate, independent processes. In general, input data for spatiotemporal fusion should be spatially continuous or filtered with a common cloud mask [17], [43], [104], [105]. Unfortunately, cloud contamination, defective sensors, and scanning gaps between orbits cause numerous abnormal or missing values, which hinders LST spatiotemporal fusion [13]. Several applied studies on spatiotemporal fusion simply ignore spatially discontinuous images or use synthetic products (e.g., eight-day LST composite products and MOD11A2) [100], [139], [152], [153].

Although synthetic products are helpful for analyzing dynamic changes in a long time series, they increase the uncertainty in location- and time-specific quantitative studies. LST fusion results can be used as input data for reconstruction [91]. Therefore, from an application perspective, it is highly desirable to synergistically combine LST reconstruction and spatiotemporal fusion.

CONCLUSIONS

Missing information and data gaps are prevalent occurrences in satellite-retrieved LST values, and a tradeoff always exists between temporal and spatial resolutions. Obtaining SCHR-LST production is crucial to many fields of research and applications [18], e.g., Earth's energy balance, material and energy exchange on a global scale, subpixel wildfire temperature detection, drought assessment, and urban heat island monitoring on a local or fine scale.

The reconstruction and spatiotemporal fusion of LST products have become areas of abundant research that aims to address the limitations. Although reconstruction methods attempt to obtain spatially continuous LST fields, i.e., by filling existing data gaps, spatiotemporal fusion methods generate gap-free LST images at high spatial and temporal resolution simultaneously.

This study reviews the recent advances made in the reconstruction and spatiotemporal fusion of LST products and puts forward prospective research. The main contributions of this work are summarized in the following three aspects:

◗ Brief theoretical concepts of LST retrieval and processing are presented. Currently used standard LST satellite products along with their specifications and the limitations of satellite-derived LST products are summarized. The problems of reconstruction and spatiotemporal fusion are described, and the connections between them are discussed in the "Background" section.

◗ A thorough overview of the current advancements made in reconstruction and spatiotemporal fusion is provided in the "Reconstruction for Spatial Continuity" and "Spatiotemporal Fusion for High Resolution" sections. The survey covers the classifications of existing methods, provides detailed comparisons of significant algorithms from experimental and theoretical perspectives, presents limitations related to SCHR-LST, and introduces prevalent validation strategies. We present three types of reconstruction methods for missing pixels (the spatial, temporal, and spatiotemporal approaches), two groups of reconstruction methods for cloudy pixels (the PMW- and SEB-based techniques), three types of spatiotemporal fusion methods (weighted function-, unmixing-, hybrid, and learning-based approaches), and four validation methods (SiV, GrV, OsV, and DsV) for the reconstructed and fused LST products. In addition, we summarize the main evaluation indicators in the "Validation of Reconstructed and Fused LST Products" section.

◗ Several future prospects of reconstruction and spatiotemporal fusion are put forward in the "Future Prospects" section. The exploitation of geographical laws, integration of physical properties, application of DL-based frameworks, combination of assimilated data, and creation of synergies between reconstruction and spatiotemporal fusion are discussed. The application conditions and features as well as the data sources of these approaches are summarized in Table 7. This review is designed to provide researchers with guidelines on selecting the proper reconstruction or fusion method.

AUTHOR INFORMATION

Penghai Wu (wuph@ahu.edu.cn) received his Ph.D. degree in cartography and geographical information engineering from Wuhan University, Wuhan, China, in 2014. He is currently an associate professor with the School of Resources and Environmental Engineering, Anhui University, Hefei, 230601, China. He has authored more than 40 research articles in peer-reviewed international journals. He is a member of the editorial board of *Journal of Sensors* and has frequently served as a referee for more than 20 international journals for remote sensing

and ecoenvironment. His research interests include land surface temperature reconstruction and fusion, deep learning, and regional ecoenvironmental change. He is a Member of IEEE.

Zhixiang Yin (yinzhixiang0630@126.com) received his M.S. degree in geodesy and surveying engineering from Wuhan University, Wuhan, China, in 2015. He is currently pursuing his Ph.D. degree in physical geography from the Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China, and is a lecturer with the School of Resources and Environmental Engineering, Anhui University, Hefei, 230601, China. His research interests include deep learning and remotely sensed image fusion and superresolution.

Chao Zeng (zengchaozc@hotmail.com) received his Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2014. He was a postdoctoral researcher in the Department of Hydraulic Engineering, Tsinghua University, Beijing, China. Currently, he is with the School of Resources and Environmental Science, Wuhan University, Wuhan, 430079, China. His research interests include remote sensing image processing and hydrological remote sensing applications.

Si-Bo Duan (duansibo@caas.cn) received his Ph.D. degree in cartography and geographical information systems from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2014. He is currently an associate professor with the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, 100081, China. His research interests include the retrieval and validation of land surface temperature.

Frank-Michael Göttsche (frank.goettsche@kit.edu) received his M.Sc. degree in physics in 1993 and his Ph.D. degree in in geophysics in 1997 from the University of Kiel, Germany. From 1996 to 1997, he was a research scholar at the Department of Earth Sciences, Uppsala University, Sweden. Between 1997 and 2003, he worked as a research scientist at the Institute of Meteorology and Climate Research, Forschungszentrum Karlsruhe, 6980, Germany. He then became an assistant professor at the United Arab Emirates (UAE) University, UAE, where he taught physics and remote sensing classes and served as a scientific consultant to the European operational satellite agency for monitoring weather, climate and the environment from space Land Surface Analysis Satellite Application Facility (LSA-SAF). In 2007, he joined the Karlsruhe Institute of Technology, Karlsruhe, Germany, as senior researcher working for LSA-SAF. His current work focuses on the in situ determination of land surface temperature and its use for validating satellite-derived LST products. Since 2017, he has served as a focus area lead (Europe) of the Committee on Earth Observation Satellites Land Product Validation Subgroup on LST and Emissivity.

Xiaoshuang Ma (mxs.88@163.com) received his Ph.D. degree in cartography and geographic information engineering from Wuhan University, Wuhan, China, in 2016. He is currently an associate professor in the Resources and Environmental Engineering Department, Anhui University, Hefei, 230601, China. His research interests include synthetic aperture radar image processing and interpretation as well as remote sensing of the environment.

Xinghua Li (lixinghua5540@whu.edu.cn) received his B.S. degree in geographical information system and his Ph.D. degree in cartography and geographical information engineering from Wuhan University, China, in 2011 and 2016, respectively. He is currently an associate professor with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430079, China. He has published more than 40 research papers. He is currently a guest editor of *Journal of Applied Remote Sensing* and a member of the editorial board of three other journals. His current research interests include multitemporal remote sensing analysis and application, remote sensing image processing, deep learning, and sparse representation. He is a Member of IEEE.

Hui Yang (yanghui@whu.edu.cn) received his Ph.D. degree in cartography and geographical information engineering from Wuhan University, Wuhan, China, in 2019. He is currently a lecturer with the Institutes of Physical Science and Information Technology, Anhui University, Hefei, 230601, China. His research interests include remote sensing image semantic segmentation, artificial intelligence, and geographic information mining.

Huanfeng Shen (shenhf@whu.edu.cn) received his B.S. degree in surveying and mapping engineering and his Ph.D. degree in photogrammetry and remote sensing from Wuhan University, China, in 2002 and 2007, respectively. In 2007, he joined the School of Resource and Environmental Sciences (SRES), Wuhan University, Wuhan, 430079, China, where he is currently a Luojia Distinguished Professor and associate dean of the SRES. He was or is the principal investigator of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has authored more than 100 research papers in peer-reviewed international journals. He is currently a member of the editorial board of *Journal of Applied Remote Sensing* and *Geography and Geo-Information Science*. His research interests include remote sensing image processing, multisource data fusion, and intelligent environmental sensing. He is a Senior Member of IEEE, council member of the China Association of Remote Sensing Application, Education Committee member of the Chinese Society for Geodesy Photogrammetry and Cartography, and Theory Committee member of the Chinese Society for Geospatial Information Society.

REFERENCES

[1] J. Tierney, J. Russell, Y. Huang, J. Damste, E. Hopmans, and A. Cohen, "Northern hemisphere controls on tropical southeast African climate during the past 60,000 years," *Science*, vol. 322, no. 5899, pp. 252–255, 2008. doi: 10.1126/science.1160485.

[2] J. Hansen, R. Ruedy, M. Sato, and K. Lo, "Global surface temperature change," *Rev. Geophys.*, vol. 48, no. 4, 2010, Art. no. RG4004. doi: 10.1029/2010RG000345.

[3] M. Anderson, R. Allen, A. Morse, and W. Kustas, "Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources," *Remote Sens. Environ.*, vol. 122, pp. 50–65, July 2012. doi: 10.1016/j.rse.2011.08.025.

[4] P. Sellers et al., "Modeling the exchanges of energy, water, and carbon between continents and the atmosphere," *Science*, vol. 275, no. 5299, pp. 502–509, 1997. doi: 10.1126/science.275.5299.502.

[5] P. Guillevic et al., "Land surface temperature product validation best practice protocol. Version 1.0," Best practice satellite-derived land product validation," 2017. [Online]. Available: https://lpvs.gsfc.nasa.gov/PDF/CEOS_LST_PROTOCOL_Oct2017_v1.0.0.pdf

[6] Z. Li et al., "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013. doi: 10.1016/j.rse.2012.12.008.

[7] Z. Wan, Y. Zhang, Q. Zhang, and Z.-L. Li, "Quality assessment and validation of the MODIS global land surface temperature," *Int. J. Remote Sens.*, vol. 25, no. 1, pp. 261–274, 2004. doi: 10.1080/0143116031000116417.

[8] Z. Liu, P. Wu, S. Duan, W. Zhan, X. Ma, and Y. Wu, "Spatiotemporal reconstruction of land surface temperature derived from FengYun geostationary satellite data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 10, pp. 4531–4543, 2017. doi: 10.1109/JSTARS.2017.2716376.

[9] I. Trigo, I. Monteiro, F. Olesen, and E. Kabsch, "An assessment of remotely sensed land surface temperature," *J. Geophys. Res.. Atmos.*, vol. 113, no. D17, 2008, Art. no. D17108. doi: 10.1029/2008JD010035.

[10] M. Jin, "Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 2. Cloudy-pixel treatment," *J. Geophys. Res., Atmos.*, vol. 105, no. D3, pp. 4061–4076, 2000. doi: 10.1029/1999JD901088.

[11] W. Cornette and J. Shanks, "Impact of cirrus clouds on remote sensing of surface temperatures," in *Proc. Passive Infrared Remote Sens. Clouds Atmos.*, 1993, pp. 252–263. doi: 10.1117/12.154910.

[12] G. Yang, W. Sun, H. Shen, X. Meng, and J. Li, "An integrated method for reconstructing daily MODIS land surface temperature data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, pp. 1026–1040, 2019. doi: 10.1109/JSTARS.2019.2896455.

[13] H. Shen, L. Huang, L. Zhang, P. Wu, 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 Sens. Environ.*, vol. 172, pp. 109–125, Jan. 2016. doi: 10.1016/j.rse.2015.11.005.

[14] H. Shwetha and D. Kumar, "Prediction of high spatio-temporal resolution land surface temperature under cloudy conditions using microwave vegetation index and ANN," *ISPRS J. Photogram. Remote Sens.*, vol. 117, pp. 40–55, July 2016. doi: 10.1016/j.isprsjprs.2016.03.011.

[15] T. Holmes, C. Hain, M. Anderson, and W. Crow, "Cloud tolerance of remote-sensing technologies to measure land surface temperature," *Hydrol. Earth Syst. Sci.*, vol. 20, no. 8, pp. 3263–3275, 2016. doi: 10.5194/hess-20-3263-2016.

[16] X. Zhang, J. Zhou, S. Liang, L. Chai, D. Wang, and J. Liu, "Estimation of 1-km all-weather remotely sensed land surface temperature based on reconstructed spatial-seamless satellite passive microwave brightness temperature and thermal infrared data," *ISPRS J. Photogram. Remote Sens.*, vol. 167, pp. 321–344, Sept. 2020. doi: 10.1016/j.isprsjprs.2020.07.014.

[17] P. Wu, H. Shen, L. Zhang, and F. 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 Sens. Environ.*, vol. 156, pp. 169–181, Jan. 2015. doi: 10.1016/j.rse.2014.09.013.

[18] S. Yang, D. Zhang, L. Sun, Y. Wang, and Y. Gao, "Assessing drought conditions in cloudy regions using reconstructed land surface temperature," *J. Meteorol. Res.*, vol. 34, no. 2, pp. 1–16, 2020. doi: 10.1007/s13351-020-9136-4.

[19] W. Zhan et al., "Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats," *Remote Sens. Environ.*, vol. 131, pp. 119–139, Apr. 2013. doi: 10.1016/j.rse.2012.12.014.

[20] T. Hengl, G. B. M. Heuvelink, M. P. Tadi, and E. J. Pebesma, "Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images," *Theor. Appl. Climatol.*, vol. 107, nos. 1–2, pp. 265–277, 2012. doi: 10.1007/s00704-011-0464-2.

[21] A. J. Prata, V. Caselles, C. Coll, J. A. Sobrino, and C. Ottlé, "Thermal remote sensing of land surface temperature from satellites: Current status and future prospects," *Remote Sens. Rev.*, vol. 12, nos. 3–4, pp. 175–224, 1995. doi: 10.1080/02757259509532285.

[22] M. Jin and R. E. Dickinson, "Land surface skin temperature climatology: Benefitting from the strengths of satellite observations," *Environ. Res. Lett.*, vol. 5, no. 4, p. 44,004, 2010. doi: 10.1088/1748-9326/5/4/044004.

[23] Z. L. Li and S. B. Duan, "Land surface temperature," in *Comprehensive Remote Sensing*, vol. 5, S. Liang, Ed., Amsterdam, The Netherlands: Elsevier, 2018, pp. 264–283.

[24] S. Duan, Z. Li, and P. Leng, "A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data," *Remote Sens. Environ.*, vol. 195, pp. 107–117, June 2017. doi: 10.1016/j.rse.2017.04.008.

[25] J. Zhou, S. L. Liang, J. Cheng, Y. J. Wang, and J. Ma, "The GLASS land surface temperature product," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 2, pp. 493–507, Feb. 2019. doi: 10.1109/JSTARS.2018.2870130.

[26] M. Isaya Ndossi and U. Avdan, "Application of open source coding technologies in the production of land surface temperature (LST) maps from Landsat: A PyQGIS plugin," *Remote Sens.*, vol. 8, no. 5, p. 413, 2016. doi: 10.3390/rs8050413.

[27] Z. M. Zhang and G. J. He, "Generation of Landsat surface temperature product for China, 2000-2010," *Int. J. Remote Sens.*, vol. 34, no. 20, pp. 7369–7375, Oct. 2013. doi: 10.1080/01431161.2013.820368.

[28] J. Zhou, F. Dai, X. Zhang, S. Zhao, and M. Li, "Developing a temporally land cover-based look-up table (TL-LUT) method for estimating land surface temperature based on AMSR-E data over the Chinese landmass," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 34, pp. 35–50, Feb. 2015. doi: 10.1016/j.jag.2014.07.001.

[29] C. Huang et al., "A physically based algorithm for retrieving land surface temperature under cloudy conditions from AMSR2 passive microwave measurements," *Int. J. Remote Sens.*, vol. 40, nos. 5–6, pp. 1828–1843, 2019. doi: 10.1080/01431161.2018.1508920.

[30] F. Hong et al., "Comprehensive assessment of four-parameter diurnal land surface temperature cycle models under clear-sky," *ISPRS J. Photogram. Remote Sens.*, vol. 142, pp. 190–204, Aug. 2018. doi: 10.1016/j.isprsjprs.2018.06.008.

[31] D. Sun, R. T. Pinker, and J. B. Basara, "Land surface temperature estimation from the next generation of Geostationary Operational Environmental Satellites: GOES M-Q," *J. Appl. Meteorol.*, vol. 43, no. 2, pp. 363–372, 2004. doi: 10.1175/1520-0450(2004)043<0363:LSTEFT>2.0.CO;2.

[32] L. Pessanha, "Product user manual: Land surface temperature (LST)," LSA SAF, 2010. [Online]. Available: https://landsaf.ipma.pt/GetDocument.do?id=304

[33] Y. Yu, P. Yu, and J. Daniels, "Production, monitoring and evaluation of GOES-R series land surface temperature data (Conference Presentation)," in *Proc. Remote Sens. Agric., Ecosyst, Hydrol. XXI*, 2019, p. 111491A. doi: 10.1117/12.2532555.

[34] Z. Wan and Z. L. Li, "A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, pp. 980–996, July 1997. doi: 10.1109/36.602541.

[35] C. Gao et al., "Land surface temperature retrieval from FY-3C/VIRR data and its cross-validation with Terra/MODIS," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 11, pp. 4944–4953, 2017. doi: 10.1109/JSTARS.2017.2728082.

[36] J. Sobrino et al., "Synergistic use of MERIS and AATSR as a proxy for estimating Land Surface Temperature from Sentinel-3 data," *Remote Sens. Environ.*, vol. 179, pp. 149–161, June 2016. doi: 10.1016/j.rse.2016.03.035.

[37] G. Hulley, C. Hughes, and S. Hook, "Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data," *J. Geophys. Res., Atmos.*, vol. 117, no. D23, 2012, Art. no. D23113.. doi: 10.1029/2012JD018506.

[38] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle, "A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1113–1126, 1998. doi: 10.1109/36.700995.

[39] M. Cook, J. Schott, J. Mandel, and N. Raqueno, "Development of an operational calibration methodology for the Landsat thermal data archive and initial testing of the atmospheric compensation component of a Land Surface Temperature (LST) product from the archive," *Remote Sens.*, vol. 6, no. 11, pp. 11,244–11,266, 2014. doi: 10.3390/rs61111244.

[40] K. Laraby and J. Schott, "Uncertainty estimation method and Landsat 7 global validation for the Landsat surface temperature product," *Remote Sens. Environ.*, vol. 216, pp. 472–481, Oct. 2018. doi: 10.1016/j.rse.2018.06.026.

[41] X. L. He, T. R. Xu, S. M. Bateni, M. Ek, S. M. Liu, and F. Chen, "Mapping regional evapotranspiration in cloudy skies via variational assimilation of all-weather land surface temperature observations," *J. Hydrol.*, vol. 585, 2020 June, Art no. 124790. doi: 10.1016/j.jhydrol.2020.124790.

[42] H. Shen et al., "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, 2015. doi: 10.1109/MGRS.2015.2441912.

[43] 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 Trans. Geosci. Remote Sens.*, vol. 44, pp. 2207–2218, Aug. 2006. doi: 10.1109/TGRS.2006.872081.

[44] D. Sun, R. Pinker, and M. Kafatos, "Diurnal temperature range over the United States: A satellite view," *Geophys. Res. Lett.*, vol. 33, no. 5, p. L05705, 2006. doi: 10.1029/2005GL024780.

[45] A. 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," *J. Geophys. Res.*, vol. 113, no. D7, p. D07107, 2008. doi: 10.1029/2007JD009048.

[46] X. Zhu, F. Cai, J. Tian, and T. Williams, "Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions," *Remote Sens.*, vol. 10, no. 4, p. 527, 2018. doi: 10.3390/rs10040527.

[47] M. Stathopoulou and C. Cartalis, "Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation," *Remote Sens. Environ.*, vol. 113, no. 12, pp. 2592–2605, 2009. doi: 10.1016/j.rse.2009.07.017.

[48] B. Bechtel, K. Zakšek, and G. Hoshyaripour, "Downscaling land surface temperature in an urban area: A case study for Hamburg, Germany," *Remote Sens.*, vol. 4, no. 10, pp. 3184–3200, 2012. doi: 10.3390/rs4103184.

[49] D. Fasbender, D. Tuia, P. Bogaert, and M. Kanevski, "Support-based implementation of Bayesian data fusion for spatial enhancement: Applications to ASTER thermal images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 598–602, 2008. doi: 10.1109/LGRS.2008.2000739.

[50] H. Xia, Y. Chen, Y. Li, and J. Quan, "Combining kernel-driven and fusion-based methods to generate daily high-spatial-resolution land surface temperatures," *Remote Sens. Environ.*, vol. 224, pp. 259–274, Apr. 2019. doi: 10.1016/j.rse.2019.02.006.

[51] H. Shen, X. Meng, and L. Zhang, "An integrated framework for the spatio–temporal–spectral fusion of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, 2016. doi: 10.1109/TGRS.2016.2596290.

[52] Y. Lu, P. Wu, X. Zhu, Y. Jiang, Z. Yin, and X. Ma, "Comparison of Surface Urban Heat Island (SUHI) at Landsat scale in Hefei, China: Diurnal, seasons and drivers," in *Proc. 2018 5th Int. Workshop Earth Observ. Remote Sens. Appl. (EORSA)*, pp. 1–4. doi: 10.1109/EORSA.2018.8598630.

[53] D. Long et al., "Generation of spatially complete and daily continuous surface soil moisture of high spatial resolution," *Remote Sens. Environ.*, vol. 233, p. 111,364, Nov. 2019. doi: 10.1016/j.rse.2019.111364.

[54] H. Shen, J. Wu, Q. Cheng, M. Aihemaiti, C. Zhang, and Z. Li, "A spatiotemporal fusion based cloud removal method for remote sensing images with land cover changes," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, pp. 862–874, 2019. doi: 10.1109/JSTARS.2019.2898348.

[55] M. Metz, D. Rocchini, and M. Neteler, "Surface temperatures at the continental scale: Tracking changes with remote sensing at unprecedented detail," *Remote Sens.*, vol. 6, no. 5, pp. 3822–3840, 2014. doi: 10.3390/rs6053822.

[56] C. Zeng, H. Shen, M. Zhong, L. Zhang, and P. Wu, "Reconstructing MODIS LST based on multitemporal classification and robust regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 512–516, 2015. doi: 10.1109/LGRS.2014.2348651.

[57] C. Guillemot and O. L. Meur, "Image inpainting: Overview and recent advances," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 127–144, 2013. doi: 10.1109/MSP.2013.2273004.

[58] M. Kilibarda et al., "Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution," *J. Geophys. Res., Atmos.*, vol. 119, no. 5, pp. 2294–2313, 2014. doi: 10.1002/2013JD020803.

[59] M. Neteler, "Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data," *Remote Sens.*, vol. 2, no. 1, pp. 333–351, 2010. doi: 10.3390/rs1020333.

[60] X. Fan, H. Liu, G. Liu, and S. Li, "Reconstruction of MODIS land-surface temperature in a flat terrain and fragmented landscape," *Int. J. Remote Sens.*, vol. 35, no. 23, pp. 7857–7877, 2014. doi: 10.1080/01431161.2014.978036.

[61] J. Kang, J. Tan, R. Jin, X. Li, and Y. Zhang, "Reconstruction of MODIS land surface temperature products based on multitemporal information," *Remote Sens.*, vol. 10, no. 7, p. 1112, 2018. doi: 10.3390/rs10071112.

[62] W. Crosson, M. Al-Hamdan, S. Hemmings, and G. Wade, "A daily merged MODIS Aqua–Terra land surface temperature data set for the conterminous United States," *Remote Sens. Environ.*, vol. 119, pp. 315–324, Apr. 2012. doi: 10.1016/j.rse.2011.12.019.

[63] G. L. Zhang et al., "Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data," *ISPRS J. Photogram. Remote Sens.*, vol. 106, pp. 157–171, Aug. 2015. doi: 10.1016/j.isprsjprs.2015.05.011.

[64] Y. Xu and Y. Shen, "Reconstruction of the land surface temperature time series using harmonic analysis," *Comput. Geosci.*, vol. 61, pp. 126–132, Dec. 2013. doi: 10.1016/j.cageo.2013.08.009.

[65] J. Scharlemann et al., "Global data for ecology and epidemiology: A novel algorithm for temporal Fourier processing MODIS data," *PLoS One*, vol. 3, no. 1, p. e1408, 2008. doi: 10.1371/journal.pone.0001408.

[66] X. Lu, R. Liu, J. Liu, and S. Liang, "Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products," *Photogram. Eng. Remote Sens.*, vol. 73, no. 10, pp. 1129–1139, 2007. doi: 10.14358/PERS.73.10.1129.

[67] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1824–1832, 2002. doi: 10.1109/TGRS.2002.802519.

[68] H. R. G. Malamiri, I. Rousta, H. Olafsson, H. Zare, and H. Zhang, "Gap-filling of MODIS time series land surface temperature (LST) products using singular spectrum analysis (SSA)," *Atmosphere*, vol. 9, p. 334. Sept. 2018. doi: 10.3390/atmos9090334.

[69] H. Liu, N. Lu, H. Jiang, J. Qin, and L. Yao, "Filling gaps of monthly terra/MODIS daytime land surface temperature using discrete cosine transform method," *Remote Sens.*, vol. 12, no. 3, p. 361, 2020. doi: 10.3390/rs12030361.

[70] G. Udahemuka, F. Van den Bergh, B. Van Wyk, and M. Van Wyk, "Robust fitting of diurnal brightness temperature cycles: Pattern recognition special edition," *South African Comput. J.*, vol. 2008, no. 40, pp. 31–36, 2008.

[71] M. Jin and R. E. Dickinson, "Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle: 1. Without clouds," *J. Geophys. Res., Atmos.*, vol. 104, no. D2, pp. 2105–2116, 1999. doi: 10.1029/1998JD200005.

[72] X. Li, H. Shen, L. Zhang, H. Zhang, Q. Yuan, and G. Yang, "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7086–7098, 2014. doi: 10.1109/TGRS.2014.2307354.

[73] L. Sun et al., "Reconstructing daily clear-sky land surface temperature for cloudy regions from MODIS data," *Comput. Geosci.*, vol. 105, pp. 10–20, Aug. 2017. doi: 10.1016/j.cageo.2017.04.007.

[74] D. Weiss, P. Atkinson, S. Bhatt, B. Mappin, S. I. Hay, and P. Gething, "An effective approach for gap-filling continental scale remotely sensed time-series," *ISPRS J. Photogram. Remote Sens.*, vol. 98, pp. 106–118, Dec. 2014. doi: 10.1016/j.isprsjprs.2014.10.001.

[75] H. Pham, S. Kim, L. Marshall, and F. Johnson, "Using 3D robust smoothing to fill land surface temperature gaps at the continental scale," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 82, p. 101,879, Oct. 2019. doi: 10.1016/j.jag.2019.05.012.

[76] T. Wang et al., "Recovering land surface temperature under cloudy skies considering the solar-cloud-satellite geometry: Application to MODIS and Landsat-8 data," *J. Geophys. Res., Atmos.*, vol. 124, no. 6, pp. 3401–3416, 2019. doi: 10.1029/2018JD028976.

[77] X. Li, Y. Zhou, G. R. Asrar, and Z. Zhu, "Creating a seamless 1 km resolution daily land surface temperature dataset for urban and surrounding areas in the conterminous United States," *Remote Sens. Environ.*, vol. 206, pp. 84–97, Mar. 2018. doi: 10.1016/j.rse.2017.12.010.

[78] T. Pede and G. Mountrakis, "An empirical comparison of interpolation methods for MODIS 8-day land surface temperature composites across the conterminous Unites States," *ISPRS J. Photogram. Remote Sens.*, vol. 142, pp. 137–150, Aug. 2018. doi: 10.1016/j.isprsjprs.2018.06.003.

[79] S. Xu, J. Cheng, and Q. Zhang, "Reconstructing all-weather land surface temperature using the Bayesian maximum entropy method over the Tibetan plateau and Heihe River basin,"

IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 9, pp. 3307–3316, 2019. doi: 10.1109/JSTARS.2019.2921924.

[80] P. Fu et al., "A physical model-based method for retrieving urban land surface temperatures under cloudy conditions," *Remote Sens. Environ.*, vol. 230, pp. 111,191, Sept. 2019. doi: 10.1016/j.rse.2019.05.010.

[81] W. Yu, J. Tan, M. Ma, X. Li, X. She, and Z. Song, "An effective similar-pixel reconstruction of the high-frequency cloud-covered areas of Southwest China," *Remote Sens.*, vol. 11, no. 3, p. 336, 2019. doi: 10.3390/rs11030336.

[82] X. Kou, L. Jiang, Y. Bo, S. Yan, and L. Chai, "Estimation of land surface temperature through blending MODIS and AMSR-E data with the Bayesian maximum entropy method," *Remote Sens.*, vol. 8, no. 2, p. 105, 2016. doi: 10.3390/rs8020105.

[83] J. Zhou et al., "A thermal sampling depth correction method for land surface temperature estimation from satellite passive microwave observation over barren land," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4743–4756, 2017. doi: 10.1109/TGRS.2017.2698828.

[84] T. Holmes, W. Crow, C. Hain, M. Anderson, and W. Kustas, "Diurnal temperature cycle as observed by thermal infrared and microwave radiometers," *Remote Sens. Environ.*, vol. 158, pp. 110–125, Mar. 2015. doi: 10.1016/j.rse.2014.10.031.

[85] X. Zhang, J. Zhou, F.-M. Göttsche, W. Zhan, S. Liu, and R. Cao, "A method based on temporal component decomposition for estimating 1-km all-weather land surface temperature by merging satellite thermal infrared and passive microwave observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4670–4691, 2019. doi: 10.1109/TGRS.2019.2892417.

[86] L. Lu, V. Venus, A. Skidmore, T. Wang, and G. Luo, "Estimating land-surface temperature under clouds using MSG/SEVIRI observations," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 2, pp. 265–276, 2011. doi: 10.1016/j.jag.2010.12.007.

[87] W. Yu, M. Ma, X. Wang, and J. Tan, "Estimating the land-surface temperature of pixels covered by clouds in MODIS products," *J. Appl. Remote Sens.*, vol. 8, no. 1, p. 083525, 2014. doi: 10.1117/1.JRS.8.083525.

[88] C. Zeng, D. Long, H. Shen, P. Wu, Y. Cui, and Y. Hong, "A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud," *ISPRS J. Photogram. Remote Sens.*, vol. 141, pp. 30–45, July 2018. doi: 10.1016/j.isprsjprs.2018.04.005.

[89] 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 Sens. Environ.*, vol. 247, pp. 111931, Sept. 2020. doi: 10.1016/j.rse.2020.111931.

[90] J. Martins et al., "An all-weather land surface temperature product based on MSG/SEVIRI observations," *Remote Sens.*, vol. 11, no. 24, p. 3044, 2019. doi: 10.3390/rs11243044.

[91] D. Long et al., "Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach," *Remote Sens. Environ.*, vol. 246, p. 111,863, Sept. 2020. doi: 10.1016/j.rse.2020.111863.

[92] C. Prigent, C. Jimenez, and F. Aires, "Toward 'all weather,' long record, and real-time land surface temperature retrievals from microwave satellite observations," *J. Geophys. Res., Atmos.*, vol. 121, no. 10, pp. 5699–5717, 2016. doi: 10.1002/2015JD024402.

[93] J. Moncet et al., "Land surface microwave emissivities derived from AMSR-E and MODIS measurements with advanced quality control," *J. Geophys. Res., Atmos.*, vol. 116, no. D16, 2011, Art. no. D16104.. doi: 10.1029/2010JD015429.

[94] X. Zhu, J. Chen, F. Gao, X. Chen, and J. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2610–2623, 2010. doi: 10.1016/j.rse.2010.05.032.

[95] G. Yang et al., "Evaluation of ASTER-like daily land surface temperature by fusing ASTER and MODIS data during the HiWATER-MUSOEXE," *Remote Sens.*, vol. 8, no. 1, p. 75, 2016. doi: 10.3390/rs8010075.

[96] H. Liu and Q. Weng, "Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007," *Remote Sens. Environ.*, vol. 117, pp. 57–71, 2012. doi: 10.1016/j.rse.2011.06.023.

[97] M. Anderson et al., "Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery," *Hydrol. Earth Syst. Sci.*, vol. 15, no. 1, pp. 223–239, 2011. doi: 10.5194/hess-15-223-2011.

[98] Y. Li, C. Huang, J. Hou, J. Gu, G. Zhu, and X. Li, "Mapping daily evapotranspiration based on spatiotemporal fusion of ASTER and MODIS images over irrigated agricultural areas in the Heihe River Basin, Northwest China," *Agric. Forest Meteorol.* vols. 244–245, pp. 82–97, 2017. doi: 10.1016/j.agrformet.2017.05.023.

[99] Y. Ma et al., "Estimation of daily evapotranspiration and irrigation water efficiency at a Landsat-like scale for an arid irrigation area using multi-source remote sensing data," *Remote Sens. Environ.*, vol. 216, pp. 715–734, Oct. 2018. doi: 10.1016/j.rse.2018.07.019.

[100] K. Hazaymeh and Q. K. Hassan, "Fusion of MODIS and Landsat-8 surface temperature images: A new approach," *PLoS One*, vol. 10, no. 3, p. e0117755, 2015. doi: 10.1371/journal.pone.0117755.

[101] J. Kim and T. S. Hogue, "Evaluation and sensitivity testing of a coupled Landsat-MODIS downscaling method for land surface temperature and vegetation indices in semi-arid regions," *J. Appl. Remote Sens.*, vol. 6, no. 1, pp. 063569-1–063569-17, 2012. doi: 10.1117/1.JRS.6.063569.

[102] P. Wu, H. Shen, T. Ai, and Y. Liu, "Land-surface temperature retrieval at high spatial and temporal resolutions based on multi-sensor fusion," *Int. J. Digit. Earth*, vol. 6, no. sup1, pp. 113–133, 2013. doi: 10.1080/17538947.2013.783131.

[103] B. Huang, J. Wang, H. Song, D. Fu, and K. Wong, "Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1011–1015, 2013. doi: 10.1109/LGRS.2012.2227930.

[104] Q. Weng, P. Fu, and F. Gao, "Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data," *Remote Sens. Environ.*, vol. 145, pp. 55–67, Apr. 2014. doi: 10.1016/j.rse.2014.02.003.

[105] 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 Sens. Environ.*, vol. 206, pp. 403–423, Mar. 2018. doi: 10.1016/j.rse.2017.12.003.

[106] B. Zhukov, D. Oertel, F. Lanzl, and G. Reinhackel, "Unmixing-based multisensor multiresolution image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1212–1226, 1999. doi: 10.1109/36.763276.

[107] R. Zurita-Milla, J. G. Clevers, and M. E. Schaepman, "Unmixing-based Landsat TM and MERIS FR data fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 453–457, 2008. doi: 10.1109/LGRS.2008.919685.

[108] M. Wu, W. Huang, Z. Niu, and C. Wang, "Generating daily synthetic Landsat imagery by combining Landsat and MODIS data," *Sensors*, vol. 15, no. 9, pp. 24002–24025, 2015. doi: 10.3390/s150924002.

[109] 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," *J. Appl. Remote Sens.*, vol. 6, p. 063507, 2012.

[110] B. Huang and H. Zhang, "Spatio-temporal reflectance fusion via unmixing: Accounting for both phenological and land-cover changes," *Int. J. Remote Sens.*, vol. 35, no. 16, pp. 6213–6233, 2014. doi: 10.1080/01431161.2014.951097.

[111] M. Wu, H. Li, W. Huang, Z. Niu, and C. Wang, "Generating daily high spatial land surface temperatures by combining ASTER and MODIS land surface temperature products for environmental process monitoring," *Environ. Sci., Process. Impacts*, vol. 17, no. 8, pp. 1396–1404, 2015. doi: 10.1039/C5EM00254K.

[112] C. Gevaert and F. García-Haro, "A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion," *Remote Sens. Environ.*, vol. 156, pp. 34–44, Jan. 2015. doi: 10.1016/j.rse.2014.09.012.

[113] X. Zhu, E. Helmer, F. Gao, D. Liu, J. Chen, and M. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," *Remote Sens. Environ.*, vol. 172, pp. 165–177, Jan. 2016. doi: 10.1016/j.rse.2015.11.016.

[114] X. Li, F. Ling, G. Foody, Y. Ge, Y. Zhang, and Y. Du, "Generating a series of fine spatial and temporal resolution land cover maps by fusing coarse spatial resolution remotely sensed images and fine spatial resolution land cover maps," *Remote Sens. Environ.*, vol. 196, pp. 293–311, July 2017. doi: 10.1016/j.rse.2017.05.011.

[115] V. Moosavi, A. Talebi, M. H. Mokhtari, S. R. F. Shamsi, and Y. Niazi, "A wavelet-artificial intelligence fusion approach (WAI-FA) for blending Landsat and MODIS surface temperature," *Remote Sens. Environ.*, vol. 169, pp. 243–254, Nov. 2015. doi: 10.1016/j.rse.2015.08.015.

[116] Z. Yin et al., "Spatiotemporal fusion of land surface temperature based on a convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, early access. doi: 10.1109/TGRS.2020.2999943.

[117] N. E. Cerna Araya, "Wireless sensor network and satellite data fusion using a deep learning approach for spatio-temporal land surface temperature estimation and forecasting," Master's thesis, School Eng., Pontificia Universidad Católica de Chile, 2018.

[118] Y. Choe and J. Yom, "Improving accuracy of land surface temperature prediction model based on deep-learning," *Spatial Inform. Res.*, vol. 28, no. 3, pp. 377–382, 2019. doi: 10.1007/s41324-019-00299-5.

[119] A. Li, Y. Bo, Y. Zhu, P. Guo, J. Bi, and Y. He, "Blending multiresolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method," *Remote Sens. Environ.*, vol. 135, pp. 52–63, Aug. 2013. doi: 10.1016/j.rse.2013.03.021.

[120] J. Xue, Y. Leung, and T. Fung, "A Bayesian data fusion approach to spatio-temporal fusion of remotely sensed images," *Remote Sens.*, vol. 9, no. 12, p. 1310, 2017. doi: 10.3390/rs9121310.

[121] J. Xue, Y. Leung, and T. Fung, "An unmixing-based Bayesian model for spatio-temporal satellite image fusion in heterogeneous landscapes," *Remote Sens.*, vol. 11, no. 3, p. 324, 2019. doi: 10.3390/rs11030324.

[122] H. Li et al., "Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China," *Remote Sens. Environ.*, vol. 142, pp. 111–121, Feb. 2014. doi: 10.1016/j.rse.2013.11.014.

[123] K. Wang and S. Liang, "Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites," *Remote Sens. Environ.*, vol. 113, no. 7, pp. 1556–1565, 2009. doi: 10.1016/j.rse.2009.03.009.

[124] Y. Yu et al., "Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 704–713, 2012. doi: 10.1109/TGRS.2011.2162338.

[125] D. Sun and R. T. Pinker, "Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8)," *J. Geophys. Res.*, vol. 108, no. D11, p. 4326, 2003. doi: 10.1029/2002JD002422.

[126] P. Guillevic et al., "Validation of Land Surface Temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements," *Remote Sens. Environ.*, vol. 154, pp. 19–37, Nov. 2014. doi: 10.1016/j.rse.2014.08.013.

[127] F. Göttsche, F. Olesen, I. Trigo, A. Bork-Unkelbach, and M. Martin, "Long term validation of land surface temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa," *Remote Sens.*, vol. 8, no. 5, p. 410, 2016. doi: 10.3390/rs8050410.

[128] N. K. Malakar, G. C. Hulley, S. J. Hook, K. Laraby, M. Cook, and J. R. Schott, "An operational land surface temperature product for Landsat thermal data: Methodology and validation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5717–5735, 2018. doi: 10.1109/TGRS.2018.2824828.

[129] F. Göttsche, F. Olesen, and A. Bork-Unkelbach, "Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia," *Int. J.*

Remote Sens., vol. 34, nos. 9–10, pp. 3069–3083, 2013. doi: 10.1080/01431161.2012.716539.

[130] E. Kabsch, F. Olesen, and F. Prata, "Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements," *Int. J. Remote Sens.*, vol. 29, nos. 17–18, pp. 5329–5345, 2008. doi: 10.1080/01431160802036326.

[131] X. Li et al., "Heihe watershed allied telemetry experimental research (HiWATER): Scientific objectives and experimental design," *Bull. Amer. Meteor. Soc.*, vol. 94, no. 8, pp. 1145–1160, 2013. doi: 10.1175/BAMS-D-12-00154.1.

[132] S. Duan et al., "Validation of Collection 6 MODIS land surface temperature product using in situ measurements," *Remote Sens. Environ.*, vol. 225, pp. 16–29, May 2019. doi: 10.1016/j.rse.2019.02.020.

[133] M. Rodell et al., "The global land data assimilation system," *Bull. Amer. Meteorol. Soc.*, vol. 85, no. 3, pp. 381–394, 2004. doi: 10.1175/BAMS-85-3-381.

[134] M. Herrero-Huerta, S. Lagüela, S. M. Alfieri, and M. Menenti, "Generating high-temporal and spatial resolution TIR image data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 78, pp. 149–162, June 2019. doi: 10.1016/j.jag.2019.01.016.

[135] A. Zhu, G. Lu, J. Liu, C. Qin, and C. Zhou, "Spatial prediction based on Third Law of Geography," *Ann. GIS*, vol. 24, no. 4, pp. 225–240, 2018. doi: 10.1080/19475683.2018.1534890.

[136] S. L. Ermida, I. F. Trigo, C. C. Dacamara, F. M. G. Ttsche, F. S. Olesen, and G. Hulley, "Validation of remotely sensed surface temperature over an oak woodland landscape — The problem of viewing and illumination geometries," *Remote Sens. Environ.*, vol. 148, pp. 16–27, May 2014. doi: 10.1016/j.rse.2014.03.016.

[137] W. Zhao, H. Wu, G. Yin, and S.-B. Duan, "Normalization of the temporal effect on the MODIS land surface temperature product using random forest regression," *ISPRS J. Photogram. Remote Sens.*, vol. 152, pp. 109–118, June 2019. doi: 10.1016/j.isprsjprs.2019.04.008.

[138] S. Duan, Z. Li, B. Tang, H. Wu, and R. Tang, "Generation of a time-consistent land surface temperature product from MODIS data," *Remote Sens. Environ.*, vol. 140, pp. 339–349, Jan. 2014. doi: 10.1016/j.rse.2013.09.003.

[139] Y. Lu, P. Wu, X. Ma, H. Yang, and Y. Wu, "Monitoring seasonal and diurnal surface urban heat islands variations using Landsat-scale data in Hefei, China, 2000–2017," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6410–6423, Oct. 2020. doi: 10.1109/JSTARS.2020.3035040.

[140] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, 2016. doi: 10.1109/MGRS.2016.2540798.

[141] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, p. 042609, 2017. doi: 10.1117/1.JRS.11.042609.

[142] Q. Yuan et al., "Deep learning in environmental remote sensing: Achievements and challenges," *Remote Sens. Environ.*, vol. 241, p. 111,716, May 2020. doi: 10.1016/j.rse.2020.111716.

[143] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, 2018. doi: 10.1109/TGRS.2018.2810208.

[144] S. Malek, F. Melgani, Y. Bazi, and N. Alajlan, "Reconstructing cloud-contaminated multispectral images with contextualized autoencoder neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2270–2282, 2017. doi: 10.1109/TGRS.2017.2777886.

[145] H. Song, Q. Liu, G. Wang, R. Hang, and B. Huang, "Spatiotemporal satellite image fusion using deep convolutional neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 821–829, 2018. doi: 10.1109/JSTARS.2018.2797894.

[146] Z. Tan, P. Yue, L. Di, and J. Tang, "Deriving high spatiotemporal remote sensing images using deep convolutional network," *Remote Sens.*, vol. 10, no. 7, p. 1066, 2018. doi: 10.3390/rs10071066.

[147] X. Liu, C. Deng, J. Chanussot, D. Hong, and B. Zhao, "StfNet: A two-stream convolutional neural network for spatiotemporal image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6552–6564, 2019. doi: 10.1109/TGRS.2019.2907310.

[148] P. Wu, Z. Yin, H. Yang, Y. Wu, and X. Ma, "Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network," *Remote Sens.*, vol. 11, no. 3, p. 300, 2019. doi: 10.3390/rs11030300.

[149] B. Chen et al., "A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform," *ISPRS J. Photogram. Remote Sens.*, vol. 131, pp. 104–120, Sept. 2017. doi: 10.1016/j.isprsjprs.2017.07.011.

[150] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017. doi: 10.1016/j.rse.2017.06.031.

[151] H. Goosse, E. Crespin, A. de Montety, M. Mann, H. Renssen, and A. Timmermann, "Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation," *J. Geophys. Res., Atmos.*, vol. 115, no. D9, 2010, Art. no. D09108. doi: 10.1029/2009JD012737.

[152] Y. Lu, P. Wu, X. Ma, and X. Li, "Detection and prediction of land use/land cover change using spatiotemporal data fusion and the Cellular Automata–Markov model," *Environ. Monit. Assess.*, vol. 191, no. 2, p. 68, 2019. doi: 10.1007/s10661-019-7200-2.

[153] H. Jiang, H. Shen, X. Li, C. Zeng, H. Liu, and F. Lei, "Extending the SMAP 9-km soil moisture product using a spatio-temporal fusion model," *Remote Sens. Environ.*, vol. 231, p. 111,224, Sept. 2019. doi: 10.1016/j.rse.2019.111224.

GRS