A New Approach for Surface Urban Heat Island Monitoring Based on Machine Learning Algorithm and Spatiotemporal Fusion Model

YUAN YAO[1,2], CUN CHANG[3], FELIX NDAYISABA[4,5], AND SHUANG WANG[6]

[1]School of Architecture and Civil Engineering, Chengdu University, Chengdu 610106, China
[2]Key Laboratory of Pattern Recognition and Intelligent Information Processing, Institutions of Higher Education of Sichuan Province, Chengdu University, Chengdu 610106, China
[3]State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Science, Urumqi 830011, China
[4]Faculty of Bioscience Engineering, Ghent University, 9000 Ghent, Belgium
[5]Department of Geography, Protestant Institute of Arts and Social Sciences, Huye, Rwanda
[6]School of Management, Xinjiang Agricultural University, Urumqi 830000, China

Corresponding author: Cun Chang (changcun@ms.xjb.ac.cn)

ABSTRACT Land surface temperature (LST) is an important indicator for assessing the surface urban heat island (SUHI) effect. This paper presents a novel approach to derive LST estimates by integrating machine learning algorithm and spatiotemporal fusion model at high spatial and temporal resolution. The spatial resolutions of Landsat TM and Landsat 8 LST data were first downscaled using random forest (RF) algorithm from 120 m and 100 m, respectively, to 30 m. The resultant LST data were fused with MODerate-resolution Imaging Spectroradiometer (MODIS) LST data, by means of the Flexible Spatiotemporal Data Fusion method (FSDAF), in order to generate high spatiotemporal resolution summer daytime LST data covering the center of Chengdu city in China. The proposed new method was used to estimate the spatiotemporal variations of the summer daytime SUHI from 2009 to 2018 over Chengdu city. Results show that: (1) RF performs way better than the classical downscaling algorithm—thermal sharpening algorithm (TsHARP) for LST, and produces higher accuracy for different land covers; (2) the fused high spatiotemporal resolution summer daytime LST values were evaluated with in situ LST obtained from Chengdu Meteorological Office and the final validation results indicated that the proposed method, in generating LST dataset, can provide more details of urban thermal environment and produce higher accuracy than the traditional FSDAF; (3) significantly increasing trends of summer daytime SUHI intensity (SUHII) in the study area were observed. SUHII increased from 2.78 °C in 2009 to 4.04 °C in 2018. The highest and lowest summer daytime LST estimates were recorded over impervious surface area (ISA) and waters, respectively.

INDEX TERMS Machine learning algorithm, land surface temperature, spatiotemporal fusion, surface urban heat island (SUHI).

I. INTRODUCTION

The urban heat island (UHI) refers to the urban centers experiencing higher temperatures than their surrounding rural areas due to human activities, and is one of the most well-known negative impacts of rapid urbanization in local environment, climate, human health and energy consumption [1], [2]. Therefore, better monitoring and understanding the variation of UHI intensity (UHII) at various temporal (inter-annual, seasonal, diurnal) and spatial (from local to global) scales

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino.

is of critical importance for global environmental and urban climate research and impact studies, which has strong implications for designing effective measure to mitigate the UHI [3], [4]. Satellite remote sensing provides an unhindered tool for studying UHI, especially the Surface Urban Heat Island (SUHI), which represents the spatial temporal structures of land surface temperature (LST) differences between urban and suburban areas [5]. So far, various studies have been conducted using satellite derived LST to study SUHI for hundreds of cities around the world [6]. For instance, Peng *et al.* [7] analyzed the SUHI intensity (SUHII) of 419 largest cities around the world during 2003-2008 using

MODerate-resolution Imaging Spectroradiometer (MODIS) MYD11A2 LST data. Similarly, Clinton and Gong [8] used MODIS (MOD/MYD11A2) LST data to investigate the SUHI in every urban area between latitude 71 and −55 in 2010.

Currently, LST data available for use in the study of SUHI are derived from Thermal Infra-Red (TIR) remote sensing such as Landsat TM/ETM+/8 and MODIS. However, the low spatial temporal resolution of the currently available satellite LST products largely inhibit their potential application in SUHI studies [9]. For example, Landsat TM/ETM+/8 TIR data provide LST data at spatial resolution between 60 m and 120 m. Moreover, their temporal resolution, which is around half to one whole month, together with frequent cloud contamination may prohibit its application for monitoring SUHI [10]. By contrast, MODIS provide high LST temporal resolution of one day or half a day, but with low spatial resolution (1000 m), and this may limit detailed spatial analysis of SUHI [11]. In other words, due to the tradeoffs among these spatial and temporal resolutions [12], there is currently no single satellite system that can directly acquire LST data of the whole earth at high spatiotemporal resolutions. It is, thus, necessary to develop a new method that can integrate multi-sensors to generate LST data at high spatial and temporal resolution for monitoring SUHI dynamics.

In recent years, several spatiotemporal image fusion models were developed in order to produce high spatial and temporal surface reflectance images [13]–[16]. Although these fusion models were originally designed to fuse short-wave surface reflectance rather than LST fusion [17]–[20], they have also offered a new ground for effectively predicting fine LST time series. Liu and Weng [21] adopted the STARFM model to simulate ASTER-like LST images using MODIS for west Nile over Los Angeles, and LST prediction residual was less than 1°C. However, caution should still apply while directly using STARFM for LST prediction in urban environments since LST is usually affected by the surrounding features [9]. In order to generate diurnal Landsat-like LSTs, Wu et al. [22] used high spatial temporal LST data from Landsat, MODIS, and GEOS/SEVIRI and proposed the spatiotemporal integrated temperature fusion model (STITFM), which successfully predicted LST with an accuracy of 2.5 K compared with in situ data. Also, Zhu et al. [23] put forward a Flexible Spatiotemporal Data Fusion method (FSDAF), which can minimize the chance for error to predict dense time series LST data and has shown some advantages beyond other fusion methods [24].

At present, many of the existing methods were developed to fuse LST data from Landsat LST and MODIS LST in order to predict daily or 8-day LST at the Landsat TIR spatial resolution. Although the United States Geological Survey (USGS) distributes resampled data from Landsat TIR channel, it is important to recall that the actual spatial resolution of these datasets is still about 100 m (Landsat TM 120 m, Landsat 8 100 m). Therefore, directly applying the spatiotemporal fusion model in generating high

temporal LST data with data from Landsat TIR channel remains far from meeting the needs of SUHI monitoring. For example, Sobrino et al. [25] found the spatial resolution of LST greater than 50 m can assess SUHI effect at district level, otherwise cannot distinguish the urban thermal environment between the different urban areas. Also, Stathopoulou and Cartalis [11] demonstrated that LST data are easily affected by regional landscape characteristics, and that high spatial resolution LST data could improve the SUHII estimation accuracy. Therefore, an improved LST data fusion approach should be developed to derive a sequence of LST datasets at 30 m spatial resolution for SUHI quantitative monitoring.

Chengdu city in Southwestern China enjoys a subtropical monsoon humid climate with many cloudy and rainy days throughout the whole year. Despite the well-established fact that poor weather conditions in this area result in very low probability of obtaining one cloud-free Landsat LST image throughout the year, little to no attention has been paid to addressing the problem. In this paper, in order to monitor the spatiotemporal variations of the summer daytime SUHII in Chengdu during 2009—2018, random forest (RF) was firstly adopted to downscale the Landsat TM and Landsat 8 LST data from 120 m and 100 m spatial resolution, respectively, to 30 m. The FSDAF was then adopted to generate high spatiotemporal resolution LST images from downscaled Landsat TM, Landsat 8 and MODIS LST data, in order to analyze the spatiotemporal variations of summer daytime SUHI in Chengdu.

II. MATERIALS AND METHODS
A. STUDY AREA
Chengdu city (Figure 1a), the capital of Sichuan province, is in Southwestern China, between 102° 54′ E and 104° 53′ E longitudes, and 30° 05′ N and 31° 26′ N latitudes. The urban area lies on the western slope of the Longmen Mountain, it is low to the southeast and high to the northwest. With a typical subtropical monsoon humid climate,
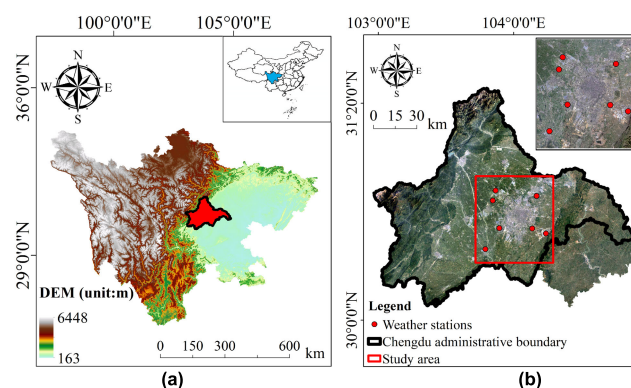

FIGURE 1. Study area: (a) The location of Chengdu city in Sichuan province, China, and (b) Geolocation of the study area and the distribution of weather stations in Chengdu city, with Landsat OLI true color image.

the average annual precipitation is 918.2 mm, and annual average number of sunny days does not exceed 25 days. Rainfall is mostly abundant in the period from June to August, and the annual average number of 340 days experience cloudy and rainy weather. Although much rain is found in summer (June - August), it also important to note that this season remains the hottest during the year, with a seasonal mean temperature of 28 °C [26]. High spatial resolution remote sensing data are thus vulnerable to poor atmospheric conditions over the area, largely limiting their potential for summer daytime SUHI monitoring.

Chengdu city is one of the most densely populated and economically vibrant cities in the world, the population of the city exceeded 16.33 million in 2018. In recent years, Chengdu has experienced rapid urbanization and industrialization growth rate, while large areas of cropland and grasslands being converted into urban and built up areas. The center of Chengdu city, covering an area of approximately 3200 km^2 was selected for the present study (Figure 1b).

B. DATA AND PREPROCESSING

1) IN SITU LST

In situ LST data were collected from Chengdu Meteorological Office covering the summer (from June to August) of 2009 and 2018. Data were acquired from 7 weather stations distributed across the study area (Figure 1b).

2) LST RETREIEVAL FROM LANDSAT DATA

Landsat satellites' products constitute an important dataset for LST retrieval because of the relatively high resolution of Landsat. Landsat TM collects TIR channel data at 120 m spatial resolution while Landsat 8 has two TIR bands at 100 m spatial resolution. Cloud free Landsat TM and Landsat 8 (path/row: 129/39) covering the study area were download from the U.S. Geological Survey (http://earthexplorer. usgs.gov/). However, Landsat has a 16-day revisit cycle, which limit the use of Landsat data in SUHI studies. Data were geographically corrected and rectified to the Universal Transverse Mercator system (UTM/WGS84). Radiometric correction was applied to convert the digital number (DN) for Landsat TM and Landsat 8 into surface reflectance values using ENVI 5.3 software. Fast line-of-sight atmospheric analysis spectral hypercubes (FLAASH) model was adopted to remove the atmospheric influence [27]. We retrieved LST data from Landsat TM TIR band 6 and Landsat 8 TIRS band 10 using the mono-window algorithm, because it is a highly effective method for retrieving LST [28]. For details of the mono-window algorithm, please refer to Qin *et al.* (2001) [29].

3) PREPROCESSING OF MOD11A1 AND MOD11A2

MODIS daily LST product (MOD11A1) and 8 Day average LST product (MOD11A2) at the local time, around 10:30 am, covering the summer of 2009 and 2018 were obtained from

NASA's Earth Observing System Data and Information System (https://earthdata.nasa. gov). These are 1 km spatial resolution LST products covering the study area (h26v05). MOD11A1 and MOD11A2 LST data were retrieved by means of the generalized split window algorithm [30]. Many research results show that the accuracy of the MODIS LST data is high in most large cities around the world [31]–[33]. We used MODIS Reprojection Tool to re-project data to the same coordinate system as Landsat. Then, we reduced the potential geometric errors by co-registering MOD11A1 and MOD11A2 to Landsat LST data. Finally, we used the quality control band in the MOD11A1 and MOD11A2 to determine cloud contaminated pixels in order to exclude them from further analysis.

C. IMPLEMENTATION FOR LST PREDICTION AND MONITORING SUMMER DAYTIME SUHI

In this study, the implementation consists of two parts that are, testing the proposed approach and monitoring summer daytime SUHI (Figure 2).
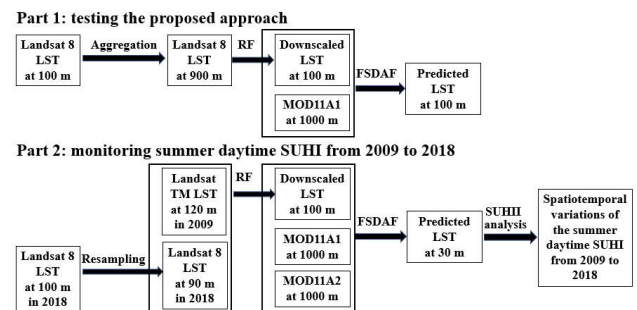

FIGURE 2. Flowchart of testing the proposed approach and monitoring summer daytime SUHI.

In the first part, three dates were chosen, 2 April 2018, 18 April 2018, and 5 June 2018. Both the downscaled Landsat 8 LST image using the RF model and MOD11A1 images captured on 2 April 2018 at 03:32 and 03:00 UTC, respectively, were used in this study as the start time (t_1) input base data for the proposed method (Table 1), while the other MOD11A1 data captured at 03:00 UTC on 18 April 2018 and 5 June 2018 were used as the predicted time (t_m) input base data for generating the high spatial resolution LST predictions at t_m. In the following parts, we call LST derived from the proposed method or FSDAF as ''predicted LST'', and LST directly from MODIS and Landsat as ''observed LST''. To test the performance of our proposed method, the observed Landsat 8 LST images at t_m were used as the referenced data for validation. In order to validate the accuracy of the predicted LST data, the coefficient of determination (R^2), the absolute average difference (AAD), and the root mean square error (RMSE) were computed between the predicted LST images and the observed LST images.

If the performance of the proposed method in the first part is better, we can conduct the next part. We selected the pairs of downscaled Landsat TM LST image acquired on

TABLE 1. The selected downscaled Landsat 8 LST data and MOD11A1 for testing the proposed method.

Satellite data	Acquisition date	Acquisition time (UTC)	Input LST data at t_1	Input LST data at t_m	Evaluate the Predicted LST data at t_m
Landsat 8	2 April 2018	03:32	√		
MOD11A1	2 April 2018	03:00	√		
MOD11A1	18 April 2018	03:00		√	
Landsat 8	18 April 2018	03:32			√
MOD11A1	5 June 2018	03:00		√	
Landsat 8	5 June 2018	03:32			√

24 March 2009, downscaled Landsat 8 LST image acquired on 2 April 2018, and the MOD11A1 data in the same period as the input base data at t_1 (Table 2). Afterwards, MOD11A2 at t_m were used to fusion the predicted LST data at 30 m spatial resolution and temporal resolution of 8 days at t_m in 2009 and 2018. Finally, the predicted summer daytime LSTs were averaged and used to assess the spatiotemporal variations of the summer daytime SUHI in the study area.

TABLE 2. The selected downscaled Landsat LST data and MODIS LST for monitoring summer daytime SUHI.

Satellite data	Acquisition date	Day of year	Input LST data at t_1	Input LST data at t_m
Landsat TM	24 March 2009	83	√	
MOD11A1	24 March 2009	83	√	
MOD11A2	2-9 Jun 2009	153-160		√
MOD11A2	29 Aug-5 Sep 2009	241-248		√
Landsat 8	2 April 2018	92	√	
MOD11A1	2 April 2018	92	√	
MOD11A2	2-9 Jun 2018	153-160		√
MOD11A2	10-17 Jun 2018	161-167		√
MOD11A2	20-27 Jul 2018	201-208		√
MOD11A2	13-20 Aug 2018	225-232		√
MOD11A2	21-28 Aug 2018	233-240		√

1) DOWNSCALING LST USING A RANDOM FOREST MODEL

RF is an integrated extension of decision tree. The basic concept of RF is to construct a set of uncorrelated classification and decision regression trees. RF is capable to unite a big number of binary decision trees constructed using bootstrap samples from the training dataset and randomly select a subset of independent variables and dependent variables at each node [34]. The results of RF training turn out to be the voting output for all decision trees. Recently, LST downscaling with machine learning algorithms, such as RF [35], support vector machine (SVM) [36], artificial neural network (ANN) [37], genetic algorithm techniques (GA) [38] have been proposed and gained more recognition with high computing precision and fast operation. RF is an effective machine learning method with the advantage of high accuracies in fitting the nonlinear statistical relationship between LST and other variables. Compared to common statistics-based downscaling LST methods, RF is more efficient in complex regions, especially in urban areas. Bartkowiak *et al.* [34] evaluated three

LST downscaling methods in Alpine vegetated areas, RF was found capable of modelling non-linear relationships between LST and variables in a very robust way. It has been proven that RF can improve LST downscaling performance in arid regions (especially in deserts) [39]. Li *et al.* [40] compared 3 machine learning algorithms, and indicated that RF yielded accurate results both in urban and rural areas.

Choosing the predictor variables in RF downscaling approach should reflect the spatial variation of LSTs over different areas. Therefore, in this paper, four predictor variable groups were selected according to the correlations between LST and different biophysical variables. Table 3 provides a brief description of the predictor variables employed in this study. From Table 3, biophysical indices were calculated from the reflectance bands of Landsat TM and Landsat 8 with the spatial resolution of 30 m. The land classification maps were obtained using SVM based on cloud-free Landsat image in summer. A DEM was acquired from the NASA's Shuttle Radar Topography Mission (SRTM), at approximately 90 m resolution. DEM derivatives such as aspect, slope and hill shade were established using the spatial analysis tool in ArcGIS.

A detailed procedure for LST downscaling is presented in Figure 3.

In the first part (testing the proposed approach), LST retrieval from Landsat 8 with 100 m resolution and the predictor variables were registered to the UTM projection over the study area. In this study, the Landsat 8 LST was aggregated to 900 m as the original low resoluton LST (LST_{OLR}), the predictor variables including the biophysical indices, terrain factors, and reflectance band were aggregated to 900 m as well. On this low spatial resolution level, RF model the statistical relationship between low resolution LST_{LR} at 900 m and the predictor variables, it can obtain as follows:

$$LST_{LR} = f(\rho_L + b_L + t_L) \qquad (1)$$

where the subscript LR means variable at low resolution, L indicates the aggregated the high spatial resolution predictors to 900 m, ρ represents the reflectance band, b is the biophysical indices, the t is the terrain factor.

The residual LST (Δ) was the difference between the LST_{LR} and LST_{OLR}, as shown in Equation (2):

$$\Delta = LST_{LR} - LST_{OLR} \qquad (2)$$

Therefore, the training model was applied to the predictor variables with high resolution (100 m) given the scale invariance, which assume a unique statistical relationship between LST and other variables exists within a sensor scene at multiple spatial resolutions [35]. The downscaled LST with 100 m resolution (LST_H) can be expressed as follows:

$$LST_H = f(\rho_H + b_H + t_H) + \Delta \qquad (3)$$

where the subscript H indicates the high resolution (100 m) variables.

In part 2 (monitoring summer daytime SUHI), LST estimates were derived from Landsat TM and Landsat 8 at their

TABLE 3. Predictor variables used for downscaling LST.

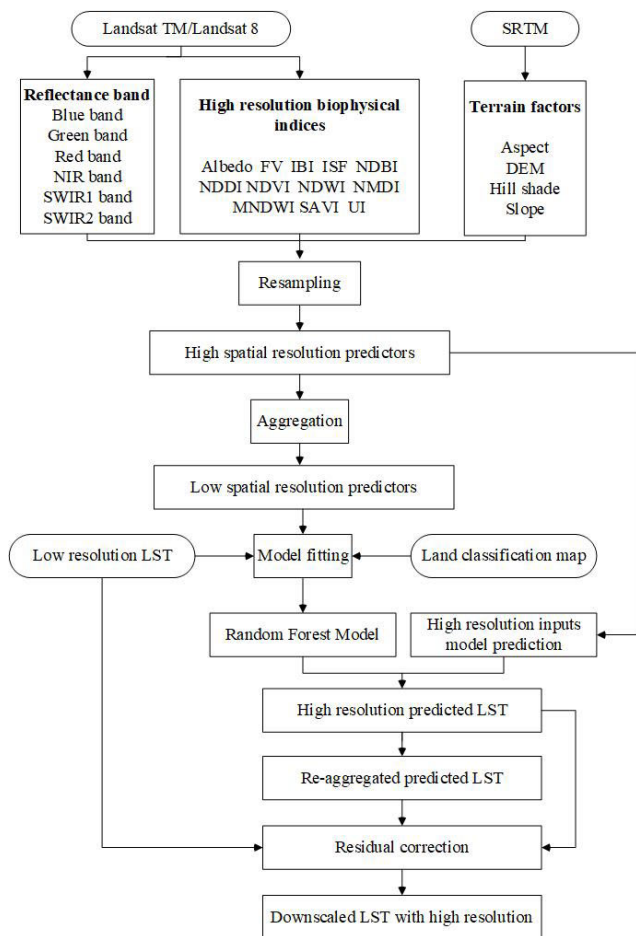| Predictor variable group	Predictor variable
Surface reflectance band	①Blue band; ②Green band; ③Red band; ④Near infrared (NIR) band; ⑤ Short wave infrared 1 (SWIR1) band; ⑥Short wave infrared 2 (SWIR2) band
Biophysical indices [42], [43]	① Vegetation-related indices: Normalized Difference Vegetation Index (NDVI); Soil Adjusted Vegetation Index (SAVI); Fraction cover (FV) ②Water indices: Normalized Difference Water Index (NDWI); Modified Normalized Difference Water Index (MNDWI) ③Built up-related indices: Normalized Difference Built-up index (NDBI); Impervious Surface Fraction (ISF); Index-based Built-up Index (IBI); Urban Index (UI) ④Soil-related indices: Normalized Difference Dust Index (NDDI); Normalized Multi-band Drought Index (NMDI) ⑤Albedo
Digital Elevation Model (DEM) and its derivatives	①DEM; ②Aspect; ③Slope; ④Hill shade
Land classification map	Land cover information

FIGURE 3. Schematic representation of LST downscaling procedure based on Random Forest algorithm.

original low resolution (120 m and 100 m, respectively) LST (LST$_{OLR}$). The biophysical indices, terrain factors, and reflectance band were also resampled to 120 m and 100 m, respectively. On this low spatial resolution level, RF model calculated the statistical relationship between low resolution LST$_{LR}$ and the predictor variable using Equation (1). Using

Equation (2) and Equation (3), the training model was applied to the predictor variables with high resolution (30 m). RF was applied to downscale Landsat TM LST and Landsat 8 LST from 120m and 100 m, respectively, to 30 m.

2) IMPLEMENTATION OF THE PROPOSED APPROACH

Figure 4 presents a detailed procedure of the proposed approach. In this study, we used FSDAF to fusion high spatiotemporal resolution LST data in order to monitor summer daytime SUHI covering the study area by combining the spatiotemporal characteristics of MOD11A1, MOD11A2 and downscaled RF LST image. FADAF was originally developed by Zhu *et al.* [23] to fuse high spatial and temporal data from one high resolution image and two low resolution images. Zhang *et al.* [24] fused MODIS and Landsat image by the FSDAF method to generate dense time series LST data. FSDAF integrated the advantages of STARFM, spectral unmixing analysis and spatial interpolation to minimize the chance for error, while capturing the LST change with time and increase the availability of high spatial and temporal LST data. The principle of FSDAF is that data from the different sensors are comparable and consistent.

In this study, FSDAF mainly includes the following steps [23]: (1) classifying downscaled LST using RF at t_1; (2) detecting the changes in each class of MODIS LST from t_1 to t_m; (3) predicting the high spatial resolution LST at t_m according to predicted temporal changes and calculating the residuals of MODIS LST; (4) interpolating the MODIS LST at t_m by the thin plate spline (TPS) function to predict the high spatial resolution LST; (5) distributing the calculated residual in step (3) to the predicted high spatial resolution of LST; (6) generating final high resolution predicted LST at t_m according to the assigned weight by neighborhood information.

The calculation process can be expressed as follows:

$$\widehat{\mathrm{LST}}_{t_m}\left(x_{ij}, y_{ij}\right) = \mathrm{LST}_{t_1}\left(x_{ij}, y_{ij}\right) + \sum_{k-1}^{n} W_k \times \Delta \mathrm{LST}(x_k, y_k) \tag{4}$$

$$\Delta \mathrm{LST}\left(x_{ij}, y_{ij}\right) = \varepsilon_{\mathrm{high}}\left(x_{ij}, y_{ij}\right) + \Delta \mathrm{LST}(c) \tag{5}$$
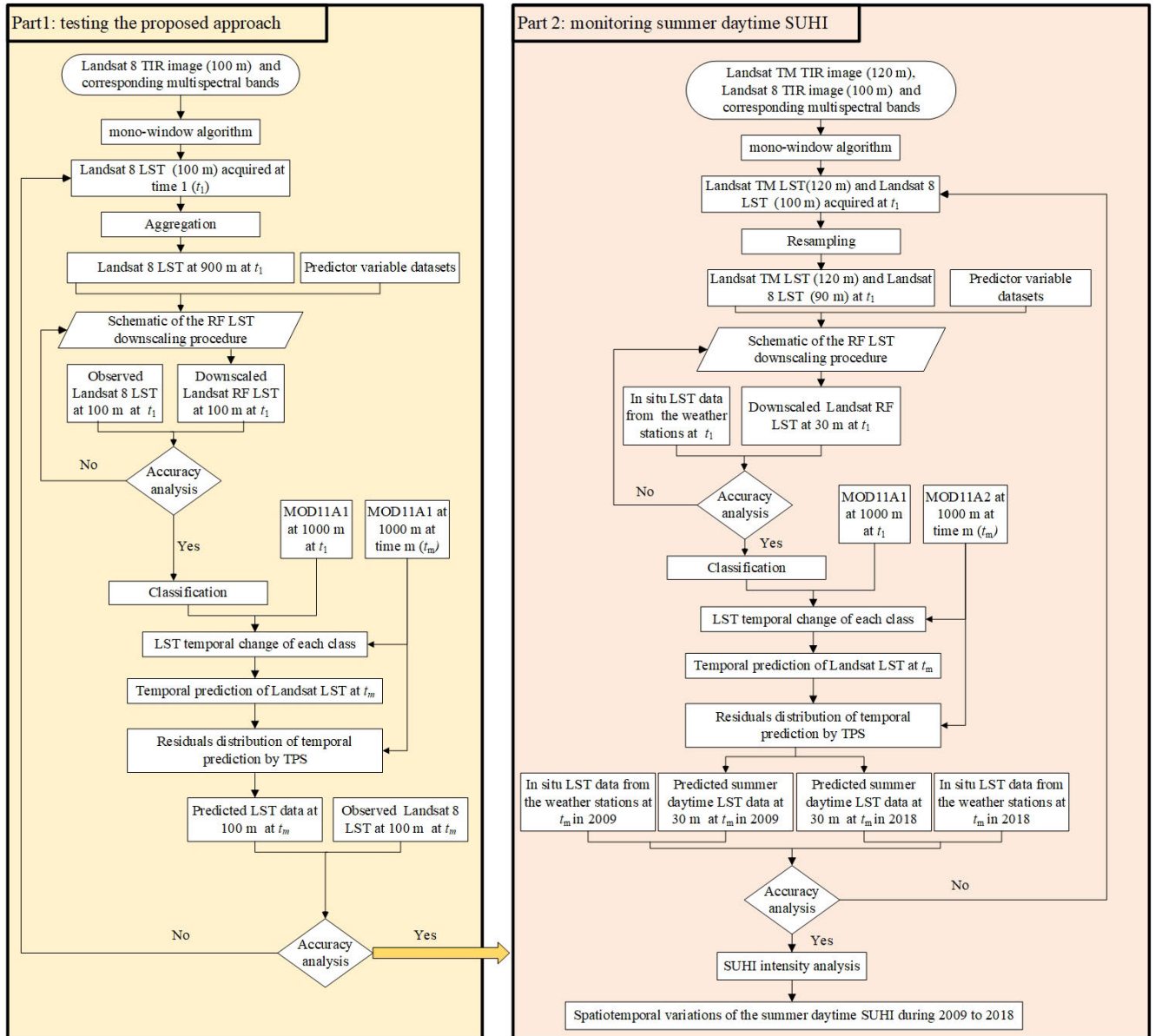
FIGURE 4. Flowchart of the proposed method for monitoring summer daytime SUHI in the study area during 2009 to 2018.

in which $\widehat{\mathrm{LST}}_{t_m}(x_{ij}, y_{ij})$ is the predicted high spatial resolution LST at time t_m, (x_{ij}, y_{ij}) is the coordinate index of the j_{th} high spatial resolution pixel within the i_{th} low spatial resolution pixel. $\mathrm{LST}_{t_1}(x_{ij}, y_{ij})$ is the downscaled RF LST data at time t_1, W_k is the weight of the k_{th} similar pixel. $\Delta\mathrm{LST}(x_k, y_k)$ represents the LST change from t_1 to t_m of similar pixels. In Equation (5), $\Delta\mathrm{LST}(x_{ij}, y_{ij})$ is the prediction of the total change of the target pixel (x_{ij}, y_{ij}) between t_1 and t_m. $\varepsilon_{\mathrm{high}}(x_{ij}, y_{ij})$ is the residual which is allocated to the j_{th} high spatial resolution pixel in the i_{th} low spatial resolution pixel. $\Delta\mathrm{LST}(c)$ is the change in LST over a certain class in the high spatial resolution data from t_1 to t_m. For more detailed steps of the FSDAF model and the calculation of the weight, kindly refer to previous studies [23]. TPS function can be used to

compute the residual by the following Equation:

$$\varepsilon_{high}(x_{ij}, y_{ij}) = m \times L(x_i, y_i) \times W(x_{ij}, y_{ij}) \tag{6}$$

$$L(x_i, y_i) = \Delta L(x_i, y_i) - \frac{1}{m}\sum_{j=1}^{m} \mathrm{LST}_{t_m}^{TP}(x_{ij}, y_{ij})$$
$$- \sum_{j=1}^{m} \mathrm{LST}_{t_1}(x_{ij}, y_{ij}) \tag{7}$$

$$LW(x_{ij}, y_{ij}) = E_{h0}(x_{ij}, y_{ij}) \times HI(x_{ij}, y_{ij}) + E_{he}(x_{ij}, y_{ij})$$
$$\times [1 - HI(x_{ij}, y_{ij})] \tag{8}$$

$$E_{h0}(x_{ij}, y_{ij}) = \mathrm{LST}_{t_m}^{SP}(x_{ij}, y_{ij}) - \mathrm{LST}_{t_m}^{TP}(x_{ij}, y_{ij}) \tag{9}$$

$$\mathrm{LST}_{t_m}^{SP}(x_{ij}, y_{ij}) = f_{TPS-b}(x_{ij}, y_{ij}) \tag{10}$$

in which $L(x_i, y_i)$ is the residual values between high spatial resolution observed LST and predicted LST, m is the number

TABLE 4. The accuracy of SVM classification in 2018.

Classified Data	Reference Data				Reference Totals	Classified Totals	Number Correct	Producer's Accuracy	User's Accuracy
	Vegetation	Water	Bare soil	ISA					
Vegetation	227	0	14	16	267	257	227	85.02%	88.32%
Water	0	20	0	0	21	20	20	95.24%	100%
Bare soil	17	1	176	31	220	225	176	80.00%	78.22%
ISA	23	0	30	945	992	998	945	95.26%	94.69%

of the sub-pixels in low spatial resolution pixel. $\Delta L(x_i, y_i)$ represents the change in LST of low spatial images from time t_1 to time t_m, $W(x_{ij}, y_{ij})$ is the normalized weight of $LW(x_{ij}, y_{ij})$. $LST_{t_m}^{TP}$ is the LST at time t_m predicted based on temporal change. $LW(x_{ij}, y_{ij})$ presents the weights of residual distribution, $E_{h0}(x_{ij}, y_{ij})$ is the temporal prediction error, $HI(x_{ij}, y_{ij})$ is the homogeneous coefficient, $E_{he}(x_{ij}, y_{ij})$ represents the equal error within a low spatial pixel. $LST_{t_m}^{SP}(x_{ij}, y_{ij})$ is the LST at time t_m of each pixel in high spatial resolution image based on TPS after optimizing the parameter. $f_{TPS-b}(x_{ij}, y_{ij})$ is the TPS functions.

3) SUHI INTENSITY ANALYSIS

SUHII indicator was used to assess summer daytime SUHII in the study area from 2009 to 2018. The urban and rural area difference in the average summer daytime LST was adopted to identify summer daytime SUHII variations by Equation 11:

$$SUHII = LST_{urban} - LST_{rural} \qquad (11)$$

III. RESULTS

A. LAND COVER CLASSIFICATION WITH LANDSAT TM AND LANDSAT 8

In order to analyze the spatiotemporal variations of the summer daytime SUHII in Chengdu, the first step was defining urban and rural areas in 2009 and 2018. Cloud-free Landsat TM and Landsat 8 data of 2009 and 2018 were acquired in summer. The land cover maps of Chengdu in 2009 and 2018 were classified using SVM [43], [44]. The urban land cover classes have been identified as follows: (i) vegetation; (ii) water; (iii) bare soil; (iv) impervious surface area (ISA). A Google Earth image acquired in the year 2018 at 1m spatial resolution was used as the reference data to evaluate the accuracy of classification results. A total of 1500 sample points from the Google Earth image were used as validation points to assess the accuracy of the classified results. Table 4 illustrates the producer's and user's accuracies.

Urban area is defined as a high intensity and densely occupied by ISA. Rural area is defined as the buffer zone that includes a rural area (as 150% of the urban area) around the urban area (Figure 5).

B. TESTING THE PROPOSED APPROACH

In order to test the performance of RF in downscaling LST data, thermal sharping algorithm (TsHARP) as the most common statistics-based LST downscaling algorithm [45], [46] was also used to downscale LST. Firstly, Landsat 8 LST
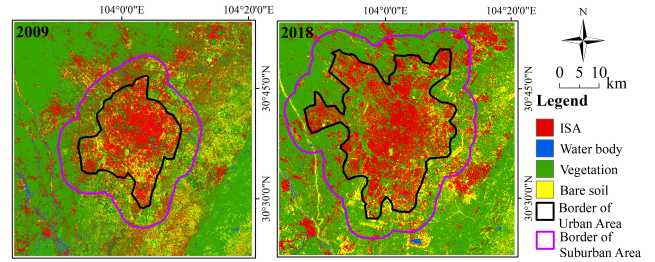

FIGURE 5. The delineation of urban and rural areas of the study area based on the land cover classification map in the years 2009 and 2018.
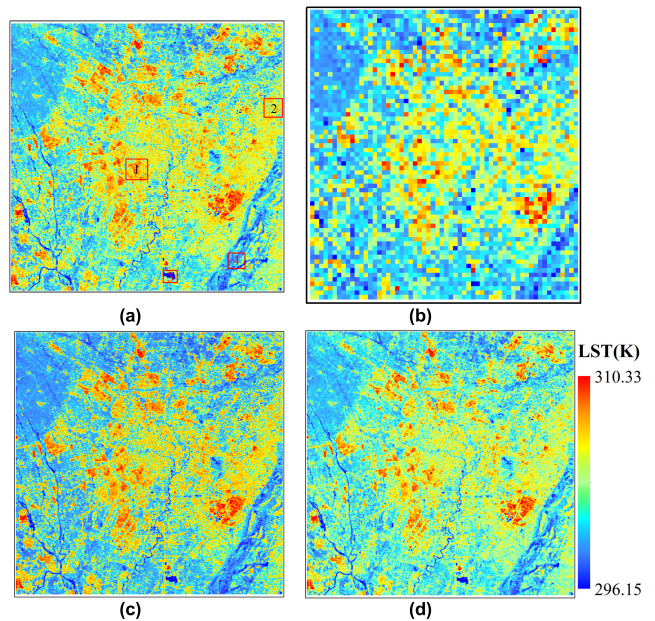

FIGURE 6. Spatial distributions of (a) 100-m observed Landsat 8 LST on 2 April 2018, (b) 900-m aggregated LST on 2 April 2018, (c) 100-m downscaled RF LST on 2 April 2018, and (d) 100-m downscaled TsHARP LST on 2 April 2018.

observed on 2 April 2018 (Figure 6a) was aggregated to 900 m as the low resolution LST (Figure 6b). The LST downscaling results of RF and TsHARP from 900 m to 100 m are shown in Figure 6c and Figure 6d, respectively. Table 5 shows the variable importance (VI) scores of RF in the study area. The contributions of DEM, FV, Red band are higher than other variable because Longquan Mountain is to the east of the study area, the spatial distribution of summer daytime LSTs is controlled by DEM, vegetation fraction and the red band that is sensitive to vegetation changes. Besides,

TABLE 5. Random forest variable importance scores in the study area.

Surface reflectance band	Biophysical indices	DEM and its derivatives	Land classification map
①Blue band (0.01); ②Green band (0.02); ③Red band (0.12); ④Near infrared (0.07) (NIR) band; ⑤SWIR1 band (0.02); ⑥SWIR2 band (0.01)	①NDVI (0.02); ②SAVI (0.01); ③FV (0.11); ④NDWI (0.03); ⑤MNDWI (0.01); ⑥NDBI (0.03); ⑦ISF (0.01); ⑧IBI (0.09); ⑨UI (0.03); ⑩NDDI (0.01); ⑪NMDI (0.02); ⑫Albedo (0.05)	① DEM (0.19); ② Aspect (0.01); ③Slope (0.06); ④Hill shade (0.02)	Land cover information (0.05)

IBI shows high importance proves LSTs is also affected by impervious surface cover in the study area.

As shown in Figure 6a, four kinds of subareas notably subarea 1 in the ISA, subarea 2 in the bare soil area, sub-area 3 in the water area, and subarea 4 in the vegetation area were used to further test the LST downscaling performance for different land cover types. By computing the RMSE and the mean error (ME) between the 100 m downscaled LST and 100 m observed Landsat 8 LST, the statistical results on the performance of RF in LST downscaling and TsHARP model for different land cover types are shown in Table 6.

TABLE 6. Downscaling statistics for RF and TsHARP method over the study area and different land cover types.

Region	RF		TsHARP	
	RMSE	ME	RMSE	ME
Study area	2.15	-0.07	3.88	-0.33
ISA	1.57	0.31	3.63	2.98
Bare soil area	1.44	0.42	3.43	2.37
Water area	1.79	-0.60	6.93	-2.77
Vegetation area	2.03	0.53	4.01	-1.64

Figure 7 shows the spatial distributions of LST error over ISA, bare soil, water body, and vegetation by using the RF and TsHARP.

From Table 6, we can see that RF produces higher LST downscaling accuracy than TsHARP. Over the study area, the RMSE of RF is 2.15 K, which indicates better perfor-mance compared to TsHARP. The ME of the downscaled LST for the RF is -0.07 K, about 0.26 K lower than TsHARP. In Figure 7, the downscaling performances of RF is better than TsHARP in different land cover types of the study area. From Figure 7b, 7d, 7f, and 7h, obvious underestimations and over estimations of the different land cover types when using TsHARP can be observed. The RMSEs of four land cover types range from 3.43 to 6.93 K by using TsHARP, while the MEs are also not close to 0. This result shows that TsHARP, that is only based on NDVI, has unsatisfactory results in downscaling LST over various and complex land cover types. This is because NDVI has great sensitivity to differences in land cover categories. The RMSEs of RF ranged from 1.44 to 2.03 K, while the MEs were close to 0, which proves RF based on the relationship between LST and multitype variables can produce the satisfactory LST downscaling results.

Therefore, the downscaled RF LST image could be used as the start time LST base data of FSDAF for generat-ing the high spatial resolution LST predictions at predict time. The next step is testing the proposed method in the
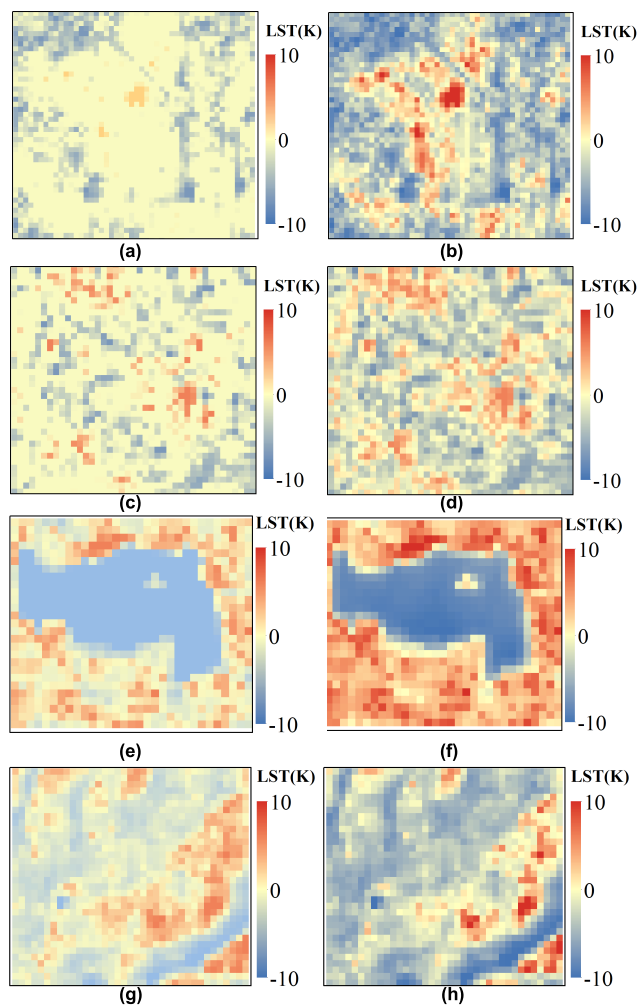

FIGURE 7. Spatial distributions of LST differences over (a) ISA by using RF, (b) ISA by using TsHARP, (c) bare soil by using RF, (d) bare soil by using TsHARP, (e) water body by using RF, (f) water body by using TsHARP, (g) vegetation by using RF, and (h) vegetation by using TsHARP.

first part (testing the proposed approach) of this study. Figure 6c and 8a were the downscaled RF LST at 100 m spatial resolution and MOD11A1 on 2 April 2018. Figure 8b and 8e are the observed MOD11A1 data on 18 April 2018 and 5 June 2018 for predicting the LST data at 100 m spatial resolution on the same date (Figure 8c and 8f). The observed Landsat 8 LST (Figure 8d and 8g), recorded on 18 April 2018 and 5 June 2018 can be used to evaluate the predicted LST results at the similar time.

Figure 9 shows scatter plots of correlation between the observed Landsat 8 LST at 03:32 UTC and predicted LSTs at
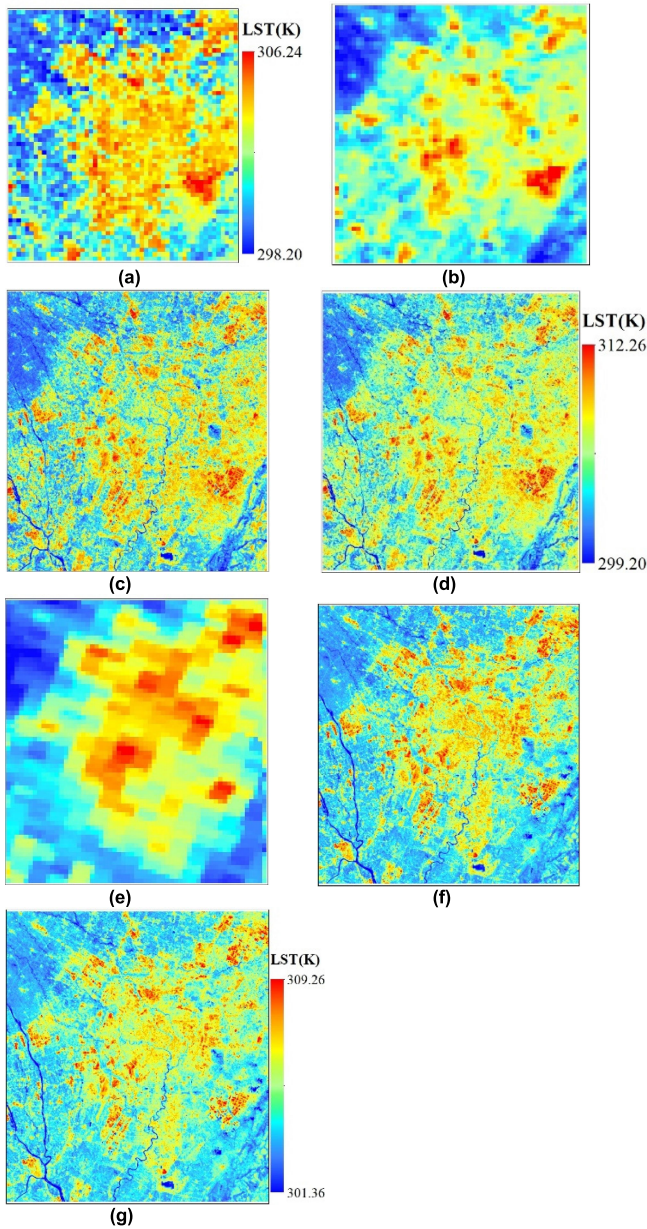
FIGURE 8. Spatial distributions of LST: (a) 1000-m MOD11A1 acquired on 2 April 2018, (b) 1000-m MOD11A1 acquired on 18 April 2018, (c) 100-m predicted LST on 18 April 2018 using the proposed method, (d) 100-m observed Landsat 8 LST on 18 April 2018, (e) 1000-m MOD11A1 acquired on 5 June 2018, (f) 100-m predicted LST on 5 June 2018 using the proposed method, and (g) 100-m observed Landsat 8 LST on 5 June 2018.

03:00 UTC on 18 April 2018 and 5 June 2018, respectively. From Figure 9, it can be observed that the R^2 between the predicted LST and the observed Landsat LST are 0.9025 and 0.9415, and AAD values are 0.031 K and 0.027 K, respectively. RMSE values are 0.089 K and 0.082 K on the same dates. These accuracy assessment results show that the proposed method has a better performance to generate the high spatiotemporal resolution LST data.

C. MONITORING SUMMER DAYTIME SUHI

The accuracy assessment results in the first part (testing the proposed approach) proves that the performance of the
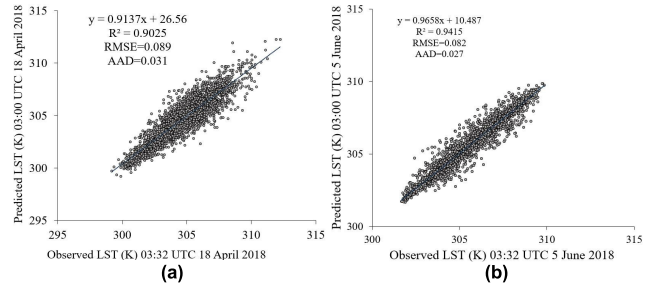

FIGURE 9. Scatter plots of the relation between observed Landsat 8 LST and predicted LST image for: (a) 18 April 2018, and (b) 5 June 2018.
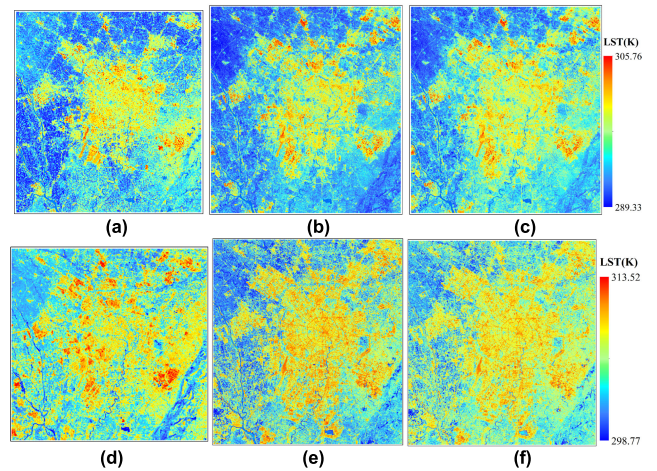

FIGURE 10. Spatial distributions of LST (a) 120-m observed LST on 24 March 2009, (b) 30-m downscaled RF LST on 24 March 2009, (c) 30-m downscaled TsHARP LST on 24 March 2009, (d) 90-m resampled Landsat 8 LST on 2 April 2018, (e) 30-m downscaled RF LST on 2 April 2018, and (f) 30-m downscaled TsHARP LST on 2 April 2018.

proposed method is better, which paves way to the next step that consist of predicting 8 day summer day summer daytime LST data at 30 m spatial resolution.

Firstly, Landsat TM LST at 120 m spatial resolution observed on 24 March 2009 (Figure 10a) and Landsat 8 LST observed on 2 April 2018 were resampled to 90 m spatial resolution (Figure 10d) and downscaled by RF to 30 m (Figure 10b and 10e), and further compared with the TsHARP as shown in Figure 10c and 10f.

Since there is no real satellite derived LST data at 30 m spatial resolution, we validated the downscaled LST data using in situ LST data. Figure 11a and 11b shows the scatterplots of correlations between the downscaled RF LST and downscaled TsHARP LST versus in situ LST retrieved on the 24th March 2009. Figure 11c and 11d shows scatterplots of downscaled RF LST and downscaled TsHARP LST versus in situ LST acquired on 2nd April 2018. From Figure 11, we can see the R^2 of the downscaled LSTs by RF exceed 90%, thus an accuracy improvement of approximately 10%, compared to TsHARP method.

Therefore, high spatiotemporal resolution summer daytime LSTs covering the study area in 2009 and 2018 were
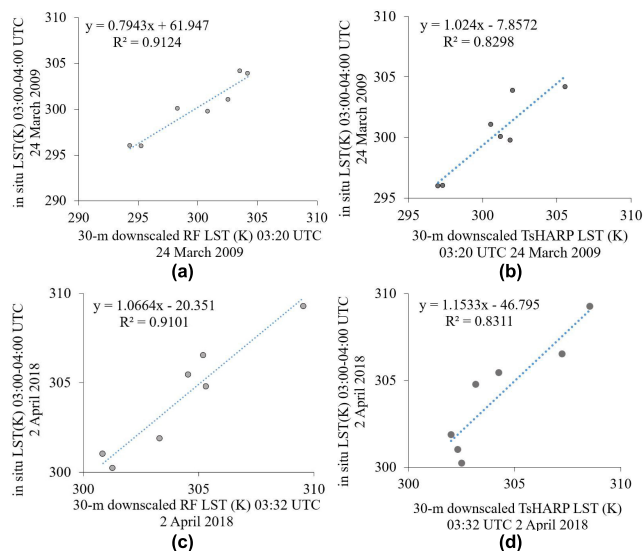
FIGURE 11. In situ LST versus (a) 30-m downscaled RF LST on 24 March 2009, (b) 30-m downscaled TsHARP LST on 24 March 2009, (c) 30-m downscaled RF LST on 2 April 2018, and (d) 30-m downscaled TsHARP LST on 2 April 2018, respectively.

predicted based on the downscaled RF LST and MODIS LST. The predicted 8 day LST data at 30 m spatial resolution and their correspondingly observed MOD11A1 are shown in Figure 12. Figure 12a and Figure 10b, Figure 8b and Figure 10e were the two pairs of MOD11A1 and downscaled RF LST on 24 March 2009 and 2 April 2018, respectively. The downscaled LST was utilized as the input base LST data at t_1 to the proposed method. Then, the second input base LST at t_m was the MOD11A2 at the predicted summer daytime on the following days of the year: 153-160 and 241-248 in 2009, 153-160, 161-167, 201-208, 225-232, and 233-240 in 2018. This permitted to fusion predicted LST data at 30 m spatial resolution and LST data at 8 days temporal resolution.

In addition, the traditional FSDAF model was also used to evaluate the performance of the predicted LST results based on the proposed method. Figure 10a and Figure 12a were the observed Landsat TM LST at 120 m spatial resolution and MOD11A1 on 24 March 2009, Figure 6a and Figure 8b were the observed Landsat 8 LST at 100 m spatial resolution and MOD11A1 on 2 April 2018 as the input base LST at t_1 to FSDAF. Then, the second input was MOD11A2 at t_m to fusion 120 m predicted summer daytime LST in 2009, and 100 m predicted summer daytime LST in 2018. It is noted that, although the Landsat TIR channel (Landsat TM 120 m, Landsat 8 100 m) is first resampled to 30 m by USGS before distribution to the users, the actual spatial resolution of observed LST retrieved form Landsat TIR channel and the predicted LST results based on FSDAF remains approximately 100 m.

Comparing the predicted LST form the proposed method and the traditional FSDAF at the same period, one notes that the general spatial distribution of the predicted LST results from the proposed method is consistent with FSDAF results.

Hence, the new approach could provide much better spatial resolution and provide important details of urban thermal environment such as continuous river, urban inner road and ring road. The R^2 between the predicted LST and in situ LST (Table 7) also show that the proposed method can produce high spatiotemporal resolution LST data with higher accuracy than traditional FSDAF.

TABLE 7. Comparison of the predicted LSTs using the proposed method and traditional FSDAF, respectively, with in situ LSTs during the summer daytime (03:00-04:00 UTC) in 2009 and 2018.

In situ LST acquisition date	R^2	
	The proposed method	FSDAF
2-9 Jun 2009	0.9433	0.8844
29 Aug-5 Sep 2009	0.9526	0.9017
2-9 Jun 2018	0.9537	0.8923
10-17 Jun 2018	0.9142	0.8741
20-27 Jul 2018	0.9341	0.8525
13-20 Aug 2018	0.9688	0.8933
21-28 Aug 2018	0.9523	0.8341

Both summer averaged daytime LST in 2009 and 2018 were computed (Figure 13). The summer mean daytime LST in 2009 were averaged on days 153-160 and 241-248. The summer mean daytime LST in 2018 were similarly averaged on days 153-160, 161-167, 201-208, 225-232, and 233-240. Figure 13a and 13d are the summer mean daytime LST at 1000 m resolution using MOD11A2 in 2009 and 2018, respectively. Figure 13b and 13e are the summer mean daytime LST at 30 m resolution using the proposed method in 2009 and 2018, respectively. Figure 13c is the summer mean daytime LST at 120 m spatial resolution using FSDAF in 2009, Figure 13f is the summer mean daytime LST at 100 m spatial resolution using FSDAF in 2018.

As shown in Figure 13, high summer daytime LSTs were mainly concentrated in the urban center and gradually extended to the urban surroundings from 2009 to 2018. The reason for this might be related to the anticipated rapid urbanization driven by population increase and fast economic growth in Chengdu. The local government deployed large industrial enterprises to the suburban of the city.

In order to further study the impact of urbanization on the urban thermal environment, we investigated the changes of summer mean LST data derived from the new method, the traditional FSDAF method and MOD11A2, respectively, over four dominant land cover types, such as ISA, vegetation, bare soil, and water body during 2009 to 2018. The results are shown in Table 8. Firstly, compared to the traditional FSDAF, the summer mean daytime LST of each land cover types using the new method is closely relate to the results derived from MOD11A2 suggesting significant consistency. Besides, similar trends were identified in all three LST product sets. Highest summer daytime LSTs were observed over ISA, followed by bare soil, vegetation and water body. ISA exhibited highest summer daytime LST because it is more readily dissipated as sensible heat and less energy is stored effectively, which means low thermal inertia resulting in an analogous thermal
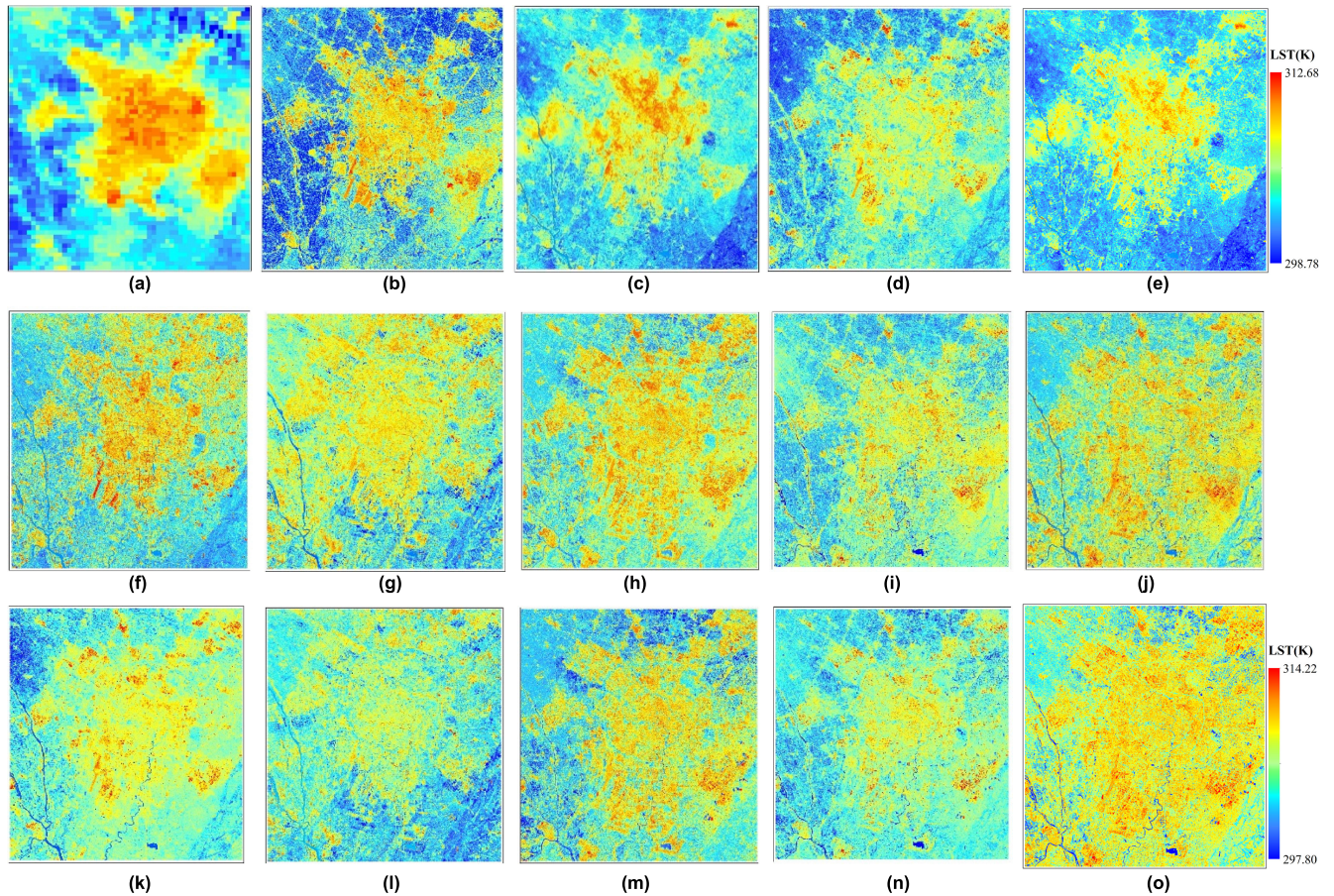
FIGURE 12. Spatial distributions of (a) 1000-m MOD11A1 on 24 March 2009, (b) 30-m predicted LST on 2-9 Jun 2009 using the proposed method, (c) 30-m predicted LST on 29 Aug-5 Sep 2009 using the proposed method, (d) 120-m predicted LST on 2-9 Jun 2009 using FSDAF, (e) 120-m predicted LST on 29 Aug-5 Sep 2009 using FSDAF, (f) 30-m predicted LST on 2-9 Jun 2018 using the proposed method, (g) 30-m predicted LST on 10-17 Jun 2018 using the proposed method, (h) 30-m predicted LST on 20-27 Jul 2018 using the proposed method, (i) 30-m predicted LST on 13-20 Aug 2018 using the proposed method, (j) 30-m predicted LST on 21-28 Aug 2018 using the proposed method, (k) 100-m predicted LST on 2-9 Jun 2018 using FSDAF, (l) 100-m predicted LST on 10-17 Jun 2018 using FSDAF, (m) 100-m predicted LST on 20-27 Jul 2018 using FSDAF, (n) 100-m predicted LST on 13-20 Aug 2018 using FSDAF, and (o) 100-m predicted LST on 21-28 Aug 2018 using FSDAF.

TABLE 8. Mean summer daytime LST over different land cover types in 2009 and 2018.

Method	Mean summer daytime LST in 2009 (°C)				Mean summer daytime LST in 2018 (°C)			
	ISA	Water body	Vegetation	Bare Soil	ISA	Water body	Vegetation	Bare Soil
The proposed method	36.53	27.99	31.21	34.62	38.89	27.76	33.72	37.67
MOD11A2	36.25	27.84	31.36	35.01	39.21	27.37	33.43	37.51
FSDAF	35.42	28.03	30.89	33.85	40.07	26.56	33.90	37.82

response behavior like bare soil. Vegetation had a lower summer daytime LST since vegetation's thermal capacities easily release heat stored through canopy transpiration. The water had the lowest summer daytime LST because water areas possess specifically large heat conservation capabilities. Therefore, in Chengdu, changing spatial patterns' control will be more useful in handling SUHI. The strategy of embedding high density vegetation, such as planting more shade trees and other vegetation species has been proved as an effective way to mitigate SUHI. In addition, no further urbanization should be envisaged in areas where the urban thermal environment is seriously deteriorating. Instead, actions to improve the local

thermal environment need to be undertaken such as using highly reflective, porous and light-colored paving materials and building. Improving energy efficiency would also halt the expansion of SUHI.

Figure 14 shows the summer daytime SUHII over the period from 2009 to 2018. Significantly increasing trends of summer daytime SUHII in the study area were observed using all four LST data. Using the new method, as proposed in this paper, it has been found that SUHII increased from 2.78 °C in 2009 to 4.04 °C in 2018 in Chengdu city. Compared to the traditional FSDAF and MOD11A2, these results are more closely related to the SUHII acquired from in situ LST data,
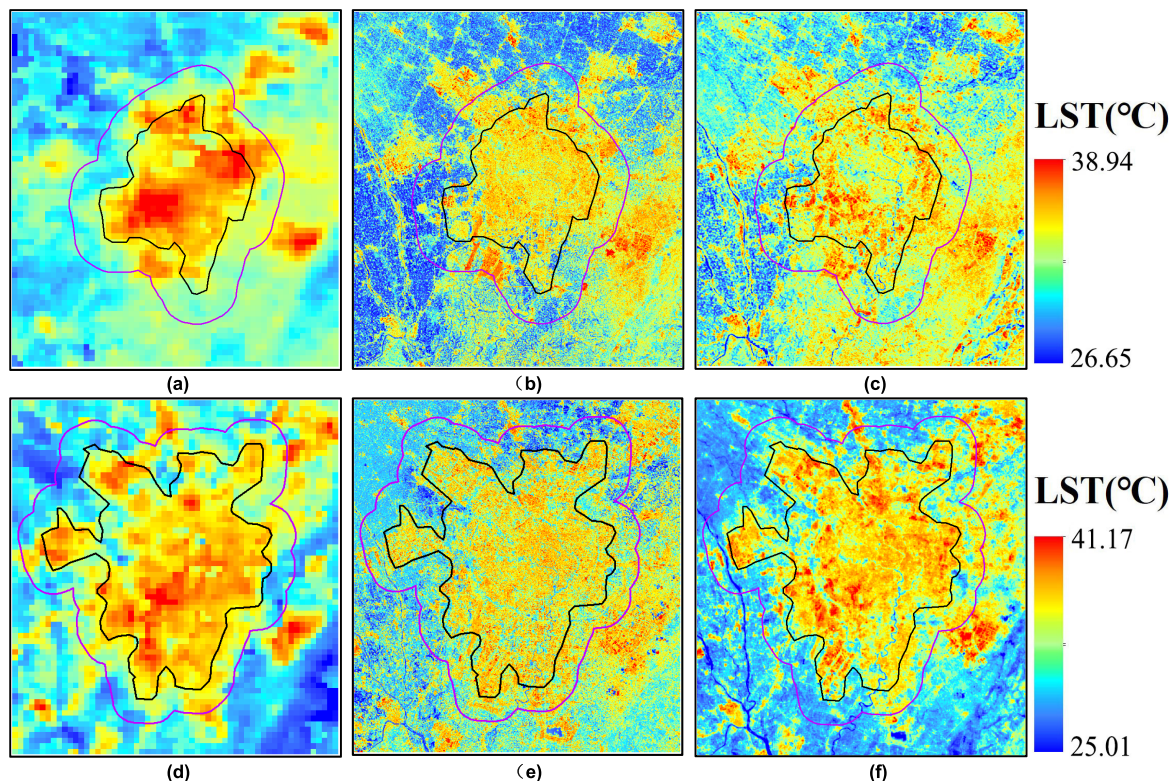
FIGURE 13. Spatial distribution of the summer daytime LSTs. (a) summer mean daytime 1000-m LST using MOD11A2 in 2009; (b) summer mean daytime 30-m predicted LST using the proposed method in 2009; (c) summer mean daytime 120-m predicted LST using FSDAF in 2009; (d) summer mean daytime 1000-m LST using MOD11A2 in 2018; (e) summer mean daytime 30-m predicted LST using the proposed method in 2018; (f) summer mean daytime 100-m predicted LST using FSDAF in 2018.
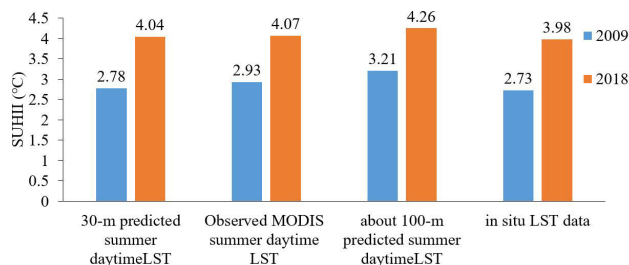

FIGURE 14. Temporal changes of SUHII in the study area for the year 2009 and 2018.

thus signaling that the proposed method is susceptible to provide reliable SUHI monitoring results.

IV. DISCUSSION

Considering the limitations of the existing spatiotemporal fusion models, this paper has presented a spatiotemporal fusion method by integrating RF and FSDAF to generate high spatiotemporal resolution summer daytime LST data. The results show that the proposed method can generate high resolution LST images with high accuracy (the RMSE and AAD value are both less than 0.089 k). Compared to the traditional FSDAF, the proposed method can increase the spatial resolution of the predicted LST image from 100 m

and 120 m to 30 m. In addition, results of this method for monitoring summer daytime SUHII are more correlated with the SUHII derived from in situ LST data.

However, the proposed method also has shortcomings that cannot be neglected. One issue relates to the fact that predicted LST data can be largely influenced by the quality of the input LST data at start time and predicted time. In this study, due to limitations of cloudy and rainy weather, we only selected Landsat and MODIS LST data for only 2 years as the input base data to fusion high spatiotemporal resolution LST data. It could be helpful in the future, to find an approach that tackles this problem by integrating some other high spatial resolution remote sensing data, in order to detect the diurnal and seasonal variations of daytime SUHIs. Secondly, since there is no actual satellite-derived LST data at 30 m resolution to evaluate the accuracy of predicted LST at 30 m spatial resolution yet, we evaluated the accuracy by in situ LST obtained from 7 weather stations, but the number of weather stations is relatively limited. The reliability of the proposed method may be influenced by the big differences in spatial resolutions between 30 m LST data and 1000 m MODIS LST data for LST data fusion. Future studies may utilize more in situ LSTs and in situ sensors mounted on fixed meteorological stations or traverses of vehicles to evaluate the predicted results. Thirdly, future studies may use the spatiotemporal fusion model to generate high spatiotemporal urban surface

biophysical variables covering longer and continued time period to allow for genuine quantitative analysis of the relationship between SUHI and the urban surface biophysical indicators.

V. CONCLUSION

In this study, the proposed method was successfully used to generate high spatial resolution (30 m) and temporal resolution (8 days) LST images from available Landsat TM, Landsat 8 and MODIS LST data in the summer of 2009 and 2018. Then, we generated the predicted summer daytime LST over the study period, analyzing the spatiotemporal change of summer daytime SUHII in Chengdu. Several conclusions were drawn from this research: (1) the performance of the proposed method could predict LST with relatively strong accuracies (R^2, AAD, RMSE values were in the range of 0.9025-0.9415; 0.027-0.037; 0.082-0.089, respectively). (2) Compared to the traditional FSDAF, the predicted summer daytime LST images derived from the proposed method can increase the spatial resolution of the predicted LST image from about 100 m to 30 m and provide more details of urban thermal environment. The predicted high spatiotemporal resolution summer daytime LST can be used for quantitatively monitoring the variations of SUHI. (3) The significantly increasing trends of SUHII in the study area were observed by using the proposed method, rising from 2.78 °C in 2009 to 4.04 °C in 2018. As discussed above, the limitations of the proposed method should be addressed, and the quantitative analysis of the relationship between SUHII and high spatiotemporal urban surface biophysical variables should be thoroughly studied in future works.

ACKNOWLEDGMENT

REFERENCES

[1] X. Zhou and H. Chen, "Impact of urbanization-related land use land cover changes and urban morphology changes on the urban heat island phenomenon," *Sci. Total Environ.*, vol. 635, pp. 1467–1476, Sep. 2018.

[2] R. Yao, L. Wang, X. Gui, Y. Zheng, H. Zhang, and X. Huang, "Urbanization effects on vegetation and surface urban heat islands in China's Yangtze River Basin," *Remote Sens.*, vol. 9, no. 6, p. 540, May 2017.

[3] D. Zhou, J. Xiao, S. Bonafoni, C. Berger, K. Deilami, Y. Zhou, S. Frolking, R. Yao, Z. Qiao, and J. Sobrino, "Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives," *Remote Sens.*, vol. 11, no. 1, p. 48, Dec. 2018.

[4] I. D. Stewart, "A systematic review and scientific critique of methodology in modern urban heat island literature," *Int. J. Climatol.*, vol. 31, no. 2, pp. 200–217, Jan. 2011.

[5] P. Ren, X. Zhang, H. Liang, and Q. Meng, "Assessing the impact of land cover changes on surface urban heat islands with high-spatial-resolution imagery on a local scale: Workflow and case study," *Sustainability*, vol. 11, no. 19, p. 5188, Sep. 2019.

[6] M. Santamouris, "Analyzing the heat island magnitude and characteristics in one hundred Asian and Australian cities and regions," *Sci. Total Environ.*, vols. 512–513, pp. 582–598, Apr. 2015.

[7] S. Peng, S. Piao, P. Ciais, P. Friedlingstein, C. Ottle, F.-M. Bréon, H. Nan, L. Zhou, and R. B. Myneni, "Surface urban heat island across 419 global big cities," *Environ. Sci. Technol.*, vol. 46, no. 2, pp. 696–703, Dec. 2012.

[8] N. Clinton and P. Gong, "MODIS detected surface urban heat islands and sinks: Global locations and controls," *Remote Sens. Environ.*, vol. 134, pp. 294–304, Jul. 2013.

[9] 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, Jan. 2013.

[10] Y. Bai, M. Wong, W.-Z. Shi, L.-X. Wu, and K. Qin, "Advancing of Land Surface Temperature retrieval using extreme learning machine and spatio-temporal adaptive data fusion algorithm," *Remote Sens.*, vol. 7, no. 4, pp. 4424–4441, Apr. 2015.

[11] 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, Dec. 2009.

[12] W. Zhan, Y. Chen, J. Zhou, J. Wang, W. Liu, J. Voogt, X. Zhu, J. Quan, and J. Li, "Disaggregation of remotely sensed Land Surface Temperature: Literature survey, taxonomy, issues, and caveats," *Remote Sens. Environ.*, vol. 131, pp. 119–139, Apr. 2013.

[13] 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, no. 8, pp. 2207–2218, Aug. 2006.

[14] X. Zhu, J. Chen, F. Gao, X. Chen, and J. G. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2610–2623, Nov. 2010.

[15] T. Hilker, M. A. Wulder, N. C. Coops, N. Seitz, J. C. White, F. Gao, J. G. Masek, and G. Stenhouse, "Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model," *Remote Sens. Environ.*, vol. 113, no. 9, pp. 1988–1999, Sep. 2009.

[16] B. Huang and H. Song, "Spatiotemporal reflectance fusion via sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3707–3716, Oct. 2012.

[17] J. Ma, W. Zhang, A. Marinoni, L. Gao, and B. Zhang, "An improved spatial and temporal reflectance unmixing model to synthesize time series of Landsat-like images," *Remote Sens.*, vol. 10, no. 9, p. 1388, Aug. 2018.

[18] C. Kwan, B. Budavari, F. Gao, and X. Zhu, "A hybrid color mapping approach to fusing MODIS and Landsat images for forward prediction," *Remote Sens.*, vol. 10, no. 4, p. 520, Mar. 2018.

[19] K. Hazaymeh and Q. K. Hassan, "Spatiotemporal image-fusion model for enhancing the temporal resolution of Landsat-8 surface reflectance images using MODIS images," *J. Appl. Remote Sens.*, vol. 9, no. 1, Jan. 2015, Art. no. 096095.

[20] C. Kwan, X. Zhu, F. Gao, B. Chou, D. Perez, J. Li, Y. Shen, K. Koperski, and G. Marchisio, "Assessment of spatiotemporal fusion algorithms for planet and worldview images," *Sensors*, vol. 18, no. 4, p. 1051, Mar. 2018.

[21] 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, Feb. 2012.

[22] P. Wu, H. Shen, L. Zhang, and F.-M. Göttsche, "Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution Land Surface Temperature," *Remote Sens. Environ.*, vol. 156, pp. 169–181, Jan. 2015.

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

[24] L. Zhang, Q. Weng, and Z. Shao, "An evaluation of monthly impervious surface dynamics by fusing Landsat and MODIS time series in the Pearl River Delta, China, from 2000 to 2015," *Remote Sens. Environ.*, vol. 201, pp. 99–114, Nov. 2017.

[25] J. A. Sobrino, R. Oltra-Carrió, G. Sòria, R. Bianchi, and M. Paganini, "Impact of spatial resolution and satellite overpass time on evaluation of the surface urban heat island effects," *Remote Sens. Environ.*, vol. 117, pp. 50–56, Feb. 2012.

[26] H. Wang and G. Han, "Chemical composition of rainwater and anthropogenic influences in Chengdu, Southwest China," *Atmos. Res.*, vol. 99, no. 2, pp. 190–196, Feb. 2011.

[27] S. M. Adler-Golden, P. K. Acharya, A. Berk, M. W. Matthew, and D. Gorodetzky, "Remote bathymetry of the littoral zone from AVIRIS, LASH, and QuickBird imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 337–347, Feb. 2005.

[28] L. Liu and Y. Zhang, "Urban heat island analysis using the Landsat TM data and ASTER data: A case study in Hong Kong," *Remote Sens.*, vol. 3, no. 7, pp. 1535–1552, Jul. 2011.

[29] Z. Qin, A. Karnieli, and P. Berliner, "A mono-window algorithm for retrieving Land Surface Temperature from Landsat TM data and its application to the Israel–Egypt border region," *Int. J. Remote Sens.*, vol. 22, no. 18, pp. 3719–3746, Jan. 2001.

[30] Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 892–905, Jul. 1996.

[31] N. Schwarz, S. Lautenbach, and R. Seppelt, "Exploring indicators for quantifying surface urban heat islands of European cities with MODIS Land Surface Temperatures," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3175–3186, Dec. 2011.

[32] M. L. Imhoff, P. Zhang, R. E. Wolfe, and L. Bounoua, "Remote sensing of the urban heat island effect across biomes in the continental USA," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 504–513, Mar. 2010.

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

[34] P. Bartkowiak, M. Castelli, and C. Notarnicola, "Downscaling Land Surface Temperature from MODIS dataset with random forest approach over alpine vegetated areas," *Remote Sens.*, vol. 11, no. 11, p. 1319, Jun. 2019.

[35] C. Hutengs and M. Vohland, "Downscaling Land Surface Temperatures at regional scales with random forest regression," *Remote Sens. Environ.*, vol. 178, pp. 127–141, Jun. 2016.

[36] J. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS) Taking Pulse Planet, Role Remote Sens. Manag. Environ.*, Jul. 2000, pp. 813–815.

[37] F. S. Mpelasoka, A. B. Mullan, and R. G. Heerdegen, "New Zealand climate change information derived by multivariate statistical and artificial neural networks approaches," *Int. J. Climatol.*, vol. 21, no. 11, pp. 1415–1433, Sep. 2001.

[38] M.-D. Yang and Y.-F. Yang, "Genetic algorithm for unsupervised classification of remote sensing imagery," *Image Process., Algorithms Syst. III*, vol. 5298, pp. 395–402, May 2004.

[39] Y. Yang, C. Cao, X. Pan, X. Li, and X. Zhu, "Downscaling Land Surface Temperature in an arid area by using multiple remote sensing indices with random forest regression," *Remote Sens.*, vol. 9, no. 8, p. 789, Jul. 2017.

[40] W. Li, L. Ni, Z.-L. Li, S.-B. Duan, and H. Wu, "Evaluation of machine learning algorithms in spatial downscaling of MODIS Land Surface Temperature," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 7, pp. 2299–2307, Jul. 2019.

[41] C. Yang, Q. Zhan, Y. Lv, and H. Liu, "Downscaling Land Surface Temperature using multiscale geographically weighted regression over heterogeneous landscapes in Wuhan, China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5213–5222, Dec. 2019.

[42] H. Wu and W. Li, "Downscaling Land Surface Temperatures using a random forest regression model with multitype predictor variables," *IEEE Access*, vol. 7, pp. 21904–21916, Jan. 2019.

[43] J. R. Otukei and T. Blaschke, "Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 12, pp. S27–S31, Feb. 2010.

[44] J. Knorn, A. Rabe, V. C. Radeloff, T. Kuemmerle, J. Kozak, and P. Hostert, "Land cover mapping of large areas using chain classification of neighboring Landsat satellite images," *Remote Sens. Environ.*, vol. 113, no. 5, pp. 957–964, May 2009.

[45] W. Ha, P. H. Gowda, and T. A. Howell, "Downscaling of Land Surface Temperature maps in the Texas high plains with the TsHARP method," *GIScience Remote Sens.*, vol. 48, no. 4, pp. 583–599, May 2011.

[46] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and C. M. U. Neale, "A vegetation index based technique for spatial sharpening of thermal imagery," *Remote Sens. Environ.*, vol. 107, no. 4, pp. 545–558, Apr. 2007.

YUAN YAO received the B.S. degree from Xinjiang Agriculture University, Urumqi, China, in 2010, the M.S. degree from Xinjiang University, Urumqi, in 2013, and the Ph.D. degree from the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, in 2019. He is currently a Lecturer with the School of Architecture and Civil Engineering, Chengdu University. His research interest includes scale transformation of land surface temperature.

CUN CHANG received the Ph.D. degree from the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China, in 2010. From 2010 to 2017, he was an Assistant Professor with the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences. He is currently an Associate Professor with the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences. He has published more than 20 peer-reviewed articles and one book. He has been involved in quantitative remote sensing and natural hazards in Central Asia.

FELIX NDAYISABA received the B.S. degree from the University of Lay Adventists of Kigali, Rwanda, in 2014, and the M.S. degree from the Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China, in 2017. He is currently appointed as a Permanent Lecturer of geography with the Faculty of Education, Protestant Institute of Arts and Social Sciences, Rwanda. He is also affiliated with Ghent University and Agrocampus Ouest France. He has published five research articles in internationally accredited journals.

SHUANG WANG received the B.S. degree from the China University of Geosciences, Beijing, China, in 2012, and the M.S. degree from Xinjiang University, Urumqi, China, in 2015. She is currently a Lecturer with Xinjiang Agriculture University, Urumqi. Her research interests include inter-calibration and land surface temperature retrieval from satellite data.

• • •