Cloud-Free Land Surface Temperature Reconstructions Based on MODIS Measurements and Numerical Simulations for Characterizing Surface Urban Heat Islands

Fengjiao Zhang, Xuepeng Zhang, Wei Chen [ID], Bin Yang [ID], Zhenting Chen, Hongzhao Tang, Zhe Wang, Pengshuai Bi, Lan Yang, Guangchao Li, and Zhe Jia

Abstract—Land surface temperature (LST) data in the thermal infrared (TIR) band measured by the moderate-resolution imaging spectroradiometer (MODIS) instrument are critical for studying surface urban heat islands (SUHIs); however, these acquired TIR LST data are contaminated by clouds, so it is crucial to develop a method to generate cloud-free LST products. In this article, employing Tianjin as the research area, we combined the weather research and forecasting model with a random forest and a spatial optimization algorithm to propose a cloud-free MODIS-like model (WRFFM). The model can reconstruct cloud-free MODIS-like LSTs and SUHIs are studied. The spatial patterns of the WRFFM LSTs and the MODIS LSTs are consistent; the correlation coefficients in July and December range from 0.8 to 0.91 and 0.8 to 0.93, respectively, and the root mean square errors range from 0.5 to 3.8 K and 0.4 to 1.8 K, respectively, indicating that the modeled results are accurate. We use these WRFFM LSTs to study SUHIs and evaluate the deviations between the MODIS SUHIs and WRFFM SUHIs. When the proportion of clear-sky pixels is below 30%, the deviation is above 3 K, and when the proportion of clear-sky pixels is above 80%, the deviation is below 0.6 K. The results indicate that the developed model can be applied to improve the study of SUHIs and that the number of clear-sky pixels for a city is an important factor that affects the bias relative to the actual SUHI.

Fengjiao Zhang, Wei Chen, Zhe Wang, Pengshuai Bi, Lan Yang, and Guangchao Li are with the School of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing 100083, China (e-mail: jiao970924@126.com; chenw@cumtb.edu.cn; wangzhe_cumtb@163.com; bps 19981102@163.com; ylannn2020@163.com; lgc911201@163.com).

Xuepeng Zhang is with the Research Center of Big Data Technology, Nanhu Laboratory, Jiaxing 314000, China (e-mail: 996588861@qq.com).

Bin Yang is with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: binyang@hnu.edu.cn).

Zhenting Chen is with the School of Information Engineering, Kunming University, Kunming 650208, China (e-mail: cztgis@qq.com).

Hongzhao Tang is with the Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing 100048, China (e-mail: tanghz@pku.edu.cn).

Zhe Jia is with the College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: 16071310@qq.com).

Digital Object Identifier 10.1109/JSTARS.2022.3199248

Index Terms—Cloud contamination, land surface temperature (LST), moderate-resolution imaging spectroradiometer (MODIS), surface urban heat island (SUHI), weather research and forecasting (WRF).

I. INTRODUCTION

WITH the rapid growth of populations and economies, impervious surfaces continue to replace natural surfaces, thus leading to large-scale urbanization [1]. From 2001 to 2018, the global urban area expanded by 1.68 times, reaching a total area of 802 233 km^2 [2]. The thermal properties of impervious surfaces and ground radiation can change significantly in response to the enhanced heat storage and thermal conductivity of urban building materials, changes in urban canopy structures, and the release of heat by human activities [3], [4]. The heat balance differences between urban and suburban areas lead to higher surface and atmospheric temperatures in urban areas than in the surrounding suburbs, resulting in a phenomenon known as "urban heat sources" [5] or "urban heat islands" (UHIs). Changes in urban thermal environments can lead to extremely high temperatures, increase climate risks, affect the human quality of life and well-being in urban areas [2], [6], [7], [8], [9], affect the quality of the ecological environment, and alter ecological functions [10], [11]. Therefore, long-term monitoring of UHIs is necessary to understand their characteristics and the urban climatology that underlies them.

Urban land surface temperature (LST) is an essential quantitative indicator of UHIs [12]. In remote sensing studies, LST is often used to define the surface urban heat island (SUHI) effect [13]. In recent years, LSTs retrieved from satellite thermal infrared (TIR) sensors have strongly promoted SUHI research [14], [15], [16], [17], [18]. Although remote sensing data cannot fully represent surface temperatures, these two metrics are related [19], [20], and SUHI data can be used to clearly distinguish between the heat island effects measured by surface models and those measured by air temperature models [13]. However, in the remote sensing data retrieval process, the TIR signals from the surface cannot penetrate the cloud layer; thus, when a pixel is contaminated by clouds, the resulting satellite observation information may be mixed with or even completely obscured by

clouds, resulting in missing values in the obtained data [21]. It is difficult to obtain spatiotemporally continuous LST data [22]. Li et al. [23] found that in the contiguous urban areas of the USA, the valid data in moderate-resolution imaging spectroradiometer (MODIS) LST products accounted for 28% of the total data. Thus, the severe absence of valid LST products limits their scientific application potential.

The retrieval accuracies of remote sensing are high, their data sources and coverages are wide, and their periods are relatively stable [24]. Therefore, remote sensing data have great advantages for studying the spatiotemporal heterogeneity of surface physical processes. However, the indirect detection of ground objects by satellites is easily affected by cloud contamination, and the spatial distribution of the obtained data is poorly continuous. Therefore, the LSTs obtained by remote sensing depend heavily on the utilized cloud-screening algorithm [25]. For MODIS LST product data, an important cloud-masking algorithm has been established that uses a series of visible and infrared thresholds to determine whether the observed region of the Earth's surface was obscured by clouds. If a pixel is covered completely or partly by clouds, its LST is not available (http://modis-atmos.gsfc.nasa.gov/). For this reason, many scholars have used a variety of reconstruction techniques to recover missing LST information resulting from cloud coverage. These methods can be divided into five categories [26]. One category is space-based gap-filling methods; LST has good spatial correlation and continuous spatial distribution characteristics. Through some spatial interpolation methods [27], [28], and geostatistics to fill in cloud-covered pixels, space-based gap-filling methods can be used to reconstruct surface temperature; however, as these methods consider only spatial information, the accuracy is unsatisfactory in cases of concentrated or large clouds. The second category is time-based gap-filling methods. These methods fill cloud-contaminated pixels with the data from valid pixels at adjacent times; examples include linear time interpolation methods [29], time-series and analysis methods [30], and harmonic fitting methods [31], [32], [33]. These methods have high reconstruction performance and are suitable for large cloud coverage areas, but they ignore adjacent geographic information and are very sensitive to land cover changes and sudden changes in natural disaster conditions [34]; the filling accuracy is susceptible. The third category is gap-filling methods based on spatiotemporal information. These methods integrate the spatiotemporal information of TIR remote sensing LST, and can offer the advantages of both time and space gap-filling methods. For example, some spatiotemporal fusion techniques and machine learning models have been introduced to estimate continuous LST in cloudy conditions [35], [36]. The accuracy of these methods mainly depends on the quality of the LST data, and they are suitable for filling large-scale data, but they are still not fully utilized with regard to the available spatiotemporal information. The fourth category is multisource fusion-based gap-filling methods, such as passive microwave measurement (PMW). Compared with TIR remote sensing measurements, these methods entail wider orbital scanning gaps and lower spatial resolutions [37], which create limitations in explaining the spatiotemporal variation in microwave emissivity [38]. Wu

et al. [39] proposed to generate gapless all-weather LST by combining PMW and TIR measurements, but this method is more costly than traditional methods and has limited applicability. The fifth category is surface energy balance based (SEB) gap-filling methods, which calculate the difference between clear- and cloudy-sky pixels based on the land SEB equation to estimate the actual LST under cloudy conditions [40]. For example, Jin [41] proposed an SEB-based theory to calculate missing LST information by interpolating adjacent pixels; however, this method does not take into account the terrain, which affects the accuracy of surface temperature reconstruction in mountainous areas. Yu et al. [42] introduced satellite radiation products into SEB theory, reducing the error caused by input variables, and thereby improving the generality and reliability of the SEB method. However, this method makes assumptions about environmental variables and requires specific meteorological and hydrological observations to calculate the LST difference between clear-sky and cloudy areas, which introduces errors [43].

The majority of the methods discussed above have a flaw in that the LST under cloudy conditions is represented by the LST theoretical value under clear-sky conditions. Although the filling methods based on multisource fusion can generate the actual LST, their verification accuracy may not be high. PMW can be regarded as one of the better methods of this kind, but it needs to overcome the rough spatial resolution and the differences among different LST data sources. Therefore, the technology is not quite mature at present. The weather research and forecasting (WRF) model can also obtain the LST under cloudy conditions, and WRF technology is more mature than PMW technology. WRF has five main modular physical schemes, namely, microphysics, shortwave radiation, longwave radiation, near-surface layer, and boundary layer schemes, which can simulate meteorological values under all weather conditions and surface skin temperature (TSK) with high temporal and spatial resolution [44], similar to MODIS. The LST retrieved from remote sensing observation data characterizes the integrated temperature of the surface skin with the surface thickness equal to the penetration depth (0.1–10 times the wavelength, detection of thermal radiation at a depth of several micrometers on the surface of the TIR observation spectrum) [45], [46]; the values of TSK and LST are very close, and can even be regarded as the same kind of data (TSK is treated as WRF LST in this article). Although the LST data simulated by WRF have spatial continuity and high spatiotemporal resolution, the simulation accuracy of this method is limited. In this article, considering that MODIS LST has high retrieval accuracy but is contaminated by clouds and has poor spatial continuity, numerical simulation with the WRF model, MODIS retrieval data, and the random forest (RF) model are integrated to give full play to the advantages of the WRF model and remote sensing data and generate a spatiotemporally complete LST. The advantage of the research method and source data in this article is the simulation of LST by WRF based on real historical observation data (final analysis data), which allows the method to simultaneously obtain LST with high accuracy in both clear-sky and cloudy conditions. First, to ensure an adequate number of valid pixels entering the RF model and improve the training accuracy of the model, we combine Terra/MODIS and AQUA/MODIS data, make full use

Fig. 1. Geographic location of Tianjin and the three one-way nested domains (a) in the WRF model used herein to simulate surface temperatures in Tianjin (b); land use cover map in panel (b) was derived from MODIS (MCD12Q1) product data. The two weather stations identified in the figure were used to verify the performance of the WRF model.

of the spatiotemporal information of both, and create a "merged" dataset (Terra-Aqua/MODIS) by filling missing Aqua/MODIS data on the premise that Terra/MODIS data are available [47]. We then generate as many MODIS pixels as possible and use the RF model as a bridge to build a model of the relationship between MODIS LST and WRF LST. This procedure can improve not only the continuous availability of retrieval data but also the accuracy of the simulated data. We then obtain the actual LST in cloudy conditions to study SUHIs in Tianjin at night and during the day and analyze the deviations between cloudy- and clear-sky MODIS SUHIs. The purpose of this article is to combine the advantages of remote sensing and meteorological techniques, consider the spatial properties of meteorological simulations, and compensate for the disadvantage of remote sensing data, i.e., its limitation by cloud contamination, which has hindered its scientific application.

This article is structured as follows: "Study area and data collection" introduces the study area and data sources; "Methods" introduces the data processing methods and technical routes; and the section "Results" presents the WRF model simulation results, the RF optimization results, and the SUHI results. "Discussion" describes the strengths of this article, problems with the model simulation and optimization methods, and potential future improvements; and the section "Conclusion" presents a summary of the findings of this article.

II. STUDY AREA AND DATA COLLECTION

A. Study Area

The study area in this article is Tianjin, a provincial administrative area of China that covers a total area of 11966.45 km^2 and has a population of approximately 13.86 million people (Fig. 1). Tianjin is located between 116°43′ and 118°04′E longitude and between 38°34′ and 40°15′N latitude, with flat topography, a temperate continental monsoon climate, high temperatures,

rainy summers, and cold and dry winters. The annual precipitation in this area is approximately 550 mm, and the annual average temperature is 13.7 °C. During the urbanization process, great temperature differences have arisen between urban and suburban areas of Tianjin, leading to an obvious heat island effect [48]. Studying the UHI effect in the area will help policy-makers formulate corresponding climate change measures and promote sustainable urban development.

B. Data on LSTs

MODIS is a 36-band sensor onboard the Terra [Earth Observing System (EOS) AM] and Aqua (EOS PM) satellites, which cover the Earth's surface four times a day, twice during the day (at 10:30 and 13:30 local time) and twice at night (at approximately 22:30 and 1:30 local time) [24]. Two kinds of daily satellite-derived MODIS LST products were used in this article, namely, the MOD11A1 data of the Terra/MODIS satellite collected in summer and winter from 2013 to 2018 and the MYD11A1 data of the Aqua/MODIS satellite. The MODIS dataset was derived from the National Aeronautics and Space Administration provided EOS data (EOSDIS, https://earthdata.nasa.gov/) at a spatial resolution of 1 km. The MODIS LSTs were generated from two TIR bands, bands 31 (10.78–11.28 μm) and 32 (11.77–12.27 μm), using a generalized split-window algorithm [49], [50]; the accuracy of these data on a uniform surface was above 1 K [25], [50]. The obtained MODIS data were spliced and reprojected to the world geodetic system-84 coordinate system and resampled to a resolution of 0.01° × 0.01° using the Arcpy tool.

C. Meteorological Data

The final operational global analysis data used in this article (http://rda.ucar.edu/datasets/ds083.2/) were reanalysis data produced and provided by the National Center for Environmental

Prediction (NCEP); these data can be used to initialize the WRF model. The utilized dataset had the characteristics of multiple phases, strong continuity, and a high resolution [51]; the temporal resolution was 6 h, the spatial resolution was 1°, and the format was gridded binary (GRIB2). The weather station-recorded data used to verify the WRF simulation results were obtained from the National Oceanic and Atmospheric Administration (NOAA) and National Center for Environmental Information website (https://www.ncei.noaa.gov/), and the temporal resolution was 30 min.

D. Data on Land Cover and Regions

Land use/cover changes (LUCCs) are considered to be the main factor affecting the climate system and are among the key parameters used in coupled atmosphere–land surface models. Land–atmosphere interactions are closely related to the regional climatic environment background, topography, vegetation, and so on [51], [52]. The built-in MODIS land-use MCD12Q1 data (2001) in the WRF model with a spatial resolution of 500 m were used in this article. UHI research requires the division between urban areas and suburbs. To improve the accuracy of the division scheme, we used Landsat 8 land use data obtained from the website of the Resource and Environmental Science and Data Center of the Chinese Academy of Sciences at a spatial resolution of 30 m.

III. METHODS

A. Technical Process

The technical process of this article is shown in Fig. 2, and is divided into the simulation of LST with the WRF model, the amalgamation of valid MODIS data, and the retrieval of cloud-free MODIS-like LST by the WRFFM model and characterization of Tianjin's SUHIs based on the model results. First, the WRF model was used to simulate the generation of surface temperature data in Tianjin. For the acquired AQUA/MODIS data, if there were not enough pixels in the study area due to cloud contamination, the pixels were filled by Terra/MODIS data from the same day. The next step was to combine the WRF model, RF model, and spatial optimization algorithm to generate cloud-free MODIS-like LST data in the case of partial cloud contamination in MODIS; if the MODIS image was completely cloud-contaminated or Terra/MODIS also had insufficient pixels, the time-adjacent model was used. Finally, the SUHIs in Tianjin were quantitatively analyzed based on the cloud-free MODIS-like LST data generated by the above steps.

B. WRF Model and Its Parameter Configuration

In this article, we used version 4.2.1 of the regional WRF model; this model was developed as a joint effort among several agencies, including the National Center for Atmospheric Research, NCEP, and NOAA. A new generation of small- and medium-scale weather forecast models has been developed [53]. The main purpose of this model development is to study and predict weather conditions [54]. The utilized model has many modular schemes that can be used to describe land surface

TABLE I
PHYSICAL PARAMETERIZATION SCHEME ADOPTED BY THE WRF MODEL

Scheme\nested domains	d01	d02	d03
Microphysical process	WSM6	WSM6	WSM6
Longwave radiation	RRTMG	RRTMG	RRTMG
Shortwave radiation	RRTMG	RRTMG	RRTMG
Surface-layer	Monin-Obukhov	Monin-Obukhov	Monin-Obukhov
Land surface process	Noah	Noah	Noah
Boundary layer	YSU	YSU	YSU
Cumulus parameterization	Kain-Fritsch	Default	Default

processes during the regional climate simulations [55], [56] and has been widely used in the meteorological field. The urban canopy model (UCM) can use the morphology of urban buildings to calculate the SEB [57], [58]; this model can accurately simulate LSTs at spatial resolutions of 1–10 km under cloudy and clear-sky conditions and surface energy fluxes [59]. In this article, we combined the WRF and UCM models using three-layer unidirectional nested domains with spatial resolutions of 9, 3, and 1 km (d01, d02, and d03, respectively, in Fig. 1); the pixel dimensions of these three domains were 97 × 97, 160 × 160, and 223 × 223, respectively. The WRF model has a physical scheme containing five modules: cloud, microphysics, radiation (longwave and shortwave), boundary layer, and surface modules [43]. Herein, we adopted the physical scheme described in Table I. According to the data recorded at the two automatic weather stations listed in Fig. 1(b), the WRF simulation results were verified, and the modeling performance was evaluated.

C. Amalgamation of Valid MODIS Data

The overpass times of the Aqua and Terra satellites are temporally close (within approximately 3 h). Assuming that most of the data are cloud contaminated, the information from these two data sources can be combined to create a "merged" dataset that complements the existing Aqua LST product and thereby reduce data loss. First, an adjustment method was applied to account for the different overpass times of the two platforms. Based on multiyear (2013–2020) MODIS LST data, the average Aqua LST and Terra LST values on Tianjin's image grid were determined according to seasons (winter: December, January and February; summer: June, July, and August). The mean Aqua-Terra LST differences were added to the daily Terra LSTs to remove the mean deviations related to the transit times, resulting in a merged Terra-Aqua/MODIS dataset. With independent offsets for each grid cell representing each season, it was possible to control the first-order statistics for factors, such as land cover, elevation, terrain slope and aspect, latitude, season, and snow cover during the diurnal surface temperature cycle [47]. The specific calculation formula can be expressed as follows:

$$M_{\text{difference}} = \frac{\sum_{i=1}^{n} T_i}{n} - \frac{\sum_{j=1}^{n} A_j}{n} \quad (1)$$
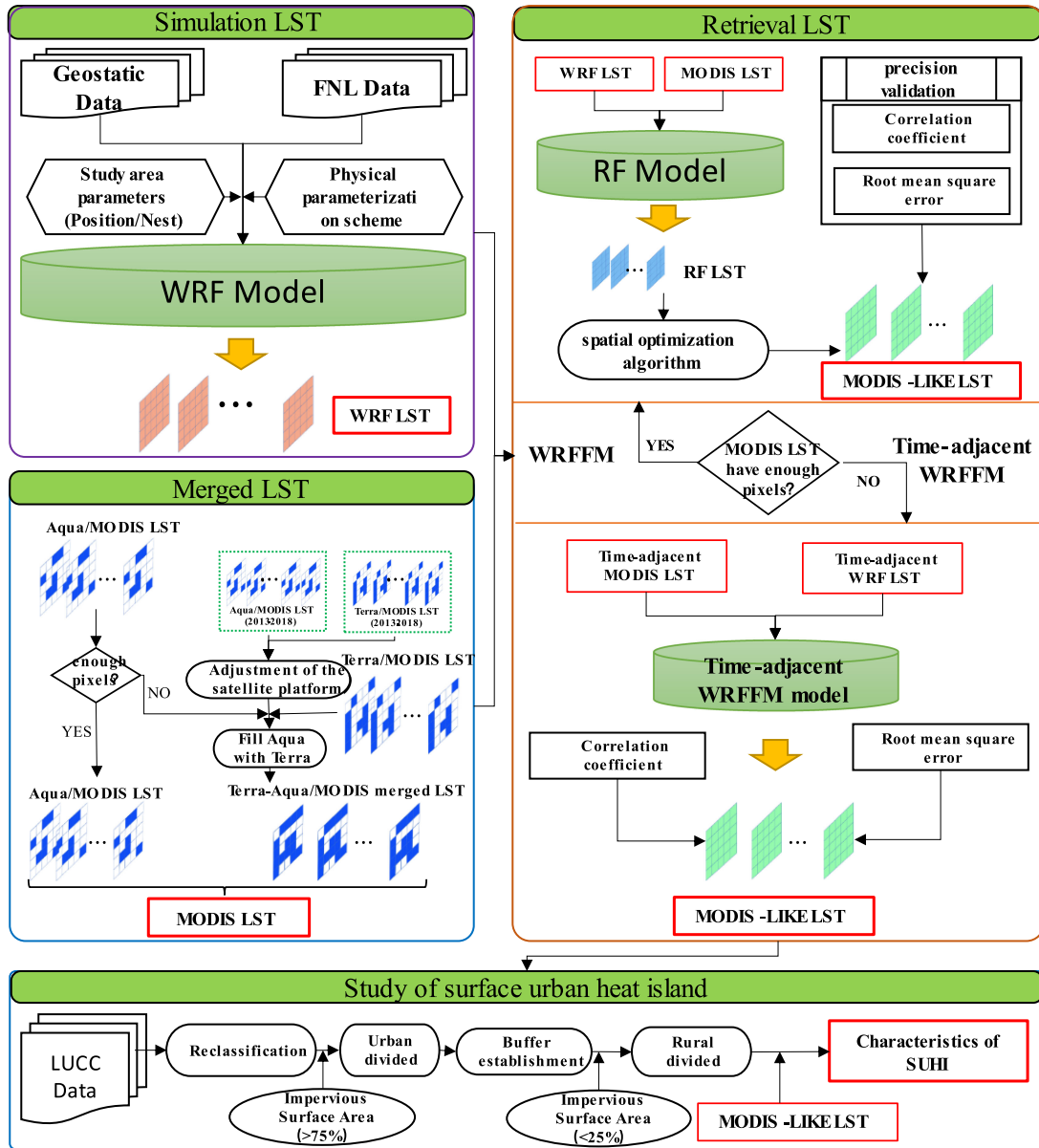
Fig. 2. Technical flow chart of the methodology followed in this article.

$$(\text{Terra} - \text{Aqua})_i = T_i + M_{\text{difference}_i} \tag{2}$$

where T_i is the summer or winter mean LST of the pixel grid of Terra/MOD11A1 data, A_j is the summer or winter mean LST of the pixel grid of the Aqua/MYD11A1 data, n is the number of years (2013–2020), $M_{\text{difference}}$ is the 8-year summer or winter mean bias, and $(\text{Terra} - \text{Aqua})_i$ is the pixel LST value of the combined Terra and Aqua datasets.

D. Retrieval and Optimization of LSTs Under Partial Cloud Conditions

The traditional RF model is a modern and highly flexible machine learning method that integrates multiple decision trees and can handle large amounts of missing data to provide high prediction accuracy [60]. In this article, two strategies were used for MODIS-like LST reconstruction under cloudy conditions based on the number of effective pixels in MODIS images (including Aqua/MODIS and the amalgamation Terra-Aqua/MODIS when cloud contamination was severe).

The reconstruction process when the MODIS LST is partially cloud contaminated (0% < cloud coverage < 99%) is shown in Fig. 3. After the simulations with the WRF model were obtained, the RF model was combined with the WRF model, where the input variables of the RF model are MODIS LST and WRF LST, the MODIS LSTs derived in July and December 2020 and the corresponding WRF LST data were taken as test data, and the image pixels input into the model were randomly divided into a training set and verification set at a ratio of 7:3. A model of the relationship between the MODIS and WRF data under clear-sky pixels was constructed, and RF LST images were
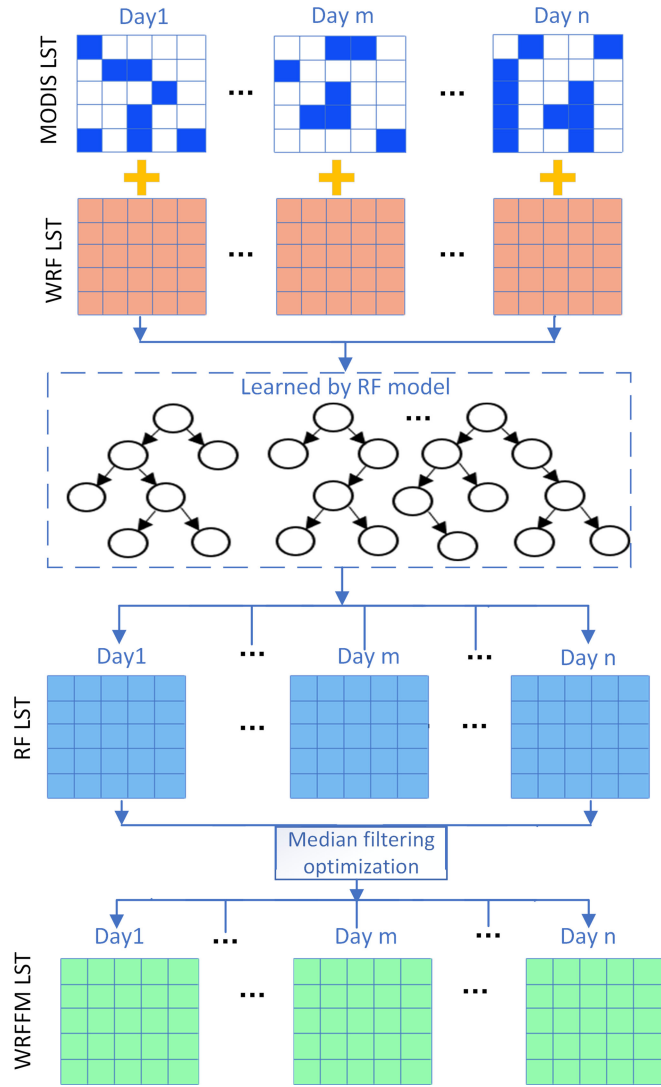
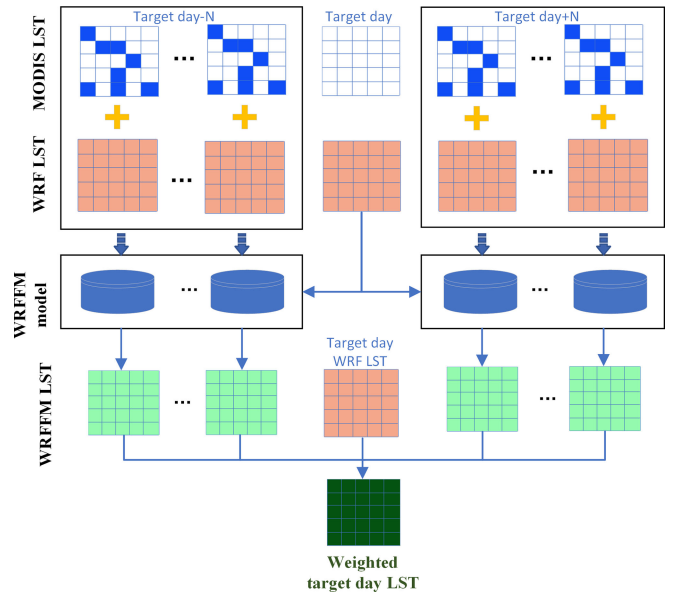Fig. 3. Reconstruction of MODIS-like LSTs under cloudy conditions.

Fig. 4. Technical process of the method used to obtain cloud-free retrievals similar to MODIS LSTs when MODIS images were completely cloud contaminated; the target day is the date when the MODIS image was completely cloud contaminated ($N = 5$), and the solid arrow represents the image input.

image of the day were used to retrieve the LST. Fig. 4 presents a flowchart of this method. The idea is to incorporate the WRF simulated images of the current day into the WRFFM model of adjacent dates and introduce the Cor value of each adjacent date verified by the WRFFM model to obtain a MODIS-like LST with weight parameters for the corresponding date. The calculation formula can be expressed as follows:

$$\text{LST}_c = \sum_{i=1}^{N} \omega_i f_i \left(\text{LST}_{c\text{WRF}}\right) + \text{CorLST}_{c\text{WRF}} \tag{3}$$

$$M_i = \frac{\frac{1}{(1-R_i) \times T_i}}{\sum_{j=1}^{n} \frac{1}{(1-R_j) \times T_j}} \tag{4}$$

$$\omega_i = M_i \times (1 - \text{cor}) \tag{5}$$

where LST_c is the MODIS-like LST image, f_i is the temporal-neighbor training model, and n is the number of adjacent image dates used to estimate LST; in this article, n was set to 10. R is the correlation coefficient between the MODIS LST image and the simulated image, T is the number of days between image acquisition dates, $\text{LST}_{c\text{WRF}}$ is the WRF image of the current day, Cor is the Pearson correlation coefficient describing the relationship between the MODIS image and the $\text{LST}_{c\text{WRF}}$ of the current day, and ω is the weight of the image.

obtained after the model was trained using the RF model. As the RF model is affected by the "regression mean" [61], salt and pepper noise arose; thus, it was necessary to use a spatial optimization algorithm to reduce this noise and smooth the images. Here, we combined the image prediction and optimization methods into a model named the WRFFM model and then used the root mean square error (RMSE) and Pearson correlation coefficient (Cor) to evaluate the accuracy of the WRFFM model outputs. Images with low Pearson correlations (Cor < 0.8) were regarded as exhibiting complete cloud contamination; in this way, high-quality cloud-free LSTs similar to those derived by MODIS could be obtained under cloudy conditions.

E. Retrieval and Optimization of LSTs Under Completely Cloud Conditions

In the case where the MODIS LST data were completely contaminated by clouds or the correlation of RF model training was low (Cor<0.8), the time-adjacent WRFFM model and the WRF

F. Calculated Indicators for SUHI and Deviations Between MODIS and WRFFM SUHIs

1) Calculated Indicators for SUHI: Based on the spatial analysis capabilities of geographic information systems, in this study, we used high-spatial-resolution LUCC data from Landsat 8, reclassified the data, and extracted the impervious surface

TABLE II
HEAT ISLAND INTENSITY CLASSIFICATION SCHEME

Grade	Heat island intensity (K)	Heat island class
1	(-∞, -5.0]	High-intensity SUCI
2	(-5.0, -3.0]	Moderate -intensity SUCI
3	(-3.0, -1.0]	Weak SUCI
4	(-1.0, 1.0]	No SUHI
5	(1.0, 3.0]	Weak SUHI
6	(3.0, 5.0]	Moderate -intensity SUHI
7	(5.0, +∞)	High-intensity SUHI

area (ISA) of Tianjin to identify urban areas. We then combined these data with the established buffer zone, which stratified the interior of the city according to ISA density; a 25% ISA threshold was used to define polygons based on thematic Landsat data, as studies have shown that the 25% ISA threshold can be used as a suitable boundary to determine the distance between a city and low-density residential land [62]. In this article, Heping District, Hedong District, Hexi District, Nankai District, Hebei District, and Hongqiao District were found to be the main urban areas in Tianjin with ISA densities >75%; 15- and 20-km buffer zones were built around these main urban areas [63]. Three concentric polygons were established, and the ISA density of the outermost polygon, which included Beichen District, Ninghe District, Jinnan District, Xiqing District, and Dongli District of Tianjin, was <25%; this delineation is similar to that utilized by Zhang et al. [64], who studied the continental USA. The indicators used to study the SUHI effect were consistent with the following simple UHI calculation formula:

$$\text{SUHI} = \text{Urban}_{\text{LST}} - \text{Suburban}_{\text{LST}} \tag{6}$$

where SUHI is a surface urban heat island, $\text{Urban}_{\text{LST}}$ is the mean value of the urban LST, and $\text{Suburban}_{\text{LST}}$ is the mean value of the suburban LST. If an SUHI was negative, it was considered to be a surface urban cold island (SUCI). According to Chen and Wang [65], by using the equal spacing method to define the intensity grade of a UHI, UHIs were divided into seven grades, as shown in Table II .

2) Deviation Between MODIS and WRFFM SUHI: Due to the influence of cloud contamination, the use of MODIS LSTs causes a lack of clear-sky pixels in the city and suburban areas in the resulting images. If the number of pixels in the area is too small, the estimated SUHI is unreliable due to the large potential deviations from the actual situation [66]. To quantify the bias (BMS) between the SUHI intensities calculated based on MODIS LST and WRFFM LST under different weather conditions and its influencing factors, in this article, we calculated the urban and suburban clear-sky coverage ratio (CSCR) and the biases of the ratio of urban clear-sky pixels to suburban clear-sky pixels under cloud contamination conditions and the ratio of urban pixels to suburban pixels under cloud-free conditions

(BCF)

$$\text{CSCR} = \frac{P_{\text{clear}-\text{sky}}}{P_{\text{all}}} \tag{7}$$

$$\text{BCF} = \frac{P_{\text{fre}-\text{urban}}}{P_{\text{fre}-\text{suburbans}}} - \frac{P_{\text{con}-\text{urban}}}{P_{\text{con}-\text{suburbans}}} \tag{8}$$

where $P_{\text{clear}-\text{sky}}$ is the number of clear-sky pixels in the city and suburban areas, P_{all} is the total number of pixels in the city and suburban areas without cloud contamination, $P_{\text{fre}-\text{urban}}$ and $P_{\text{fre}-\text{suburbans}}$ represent cases without cloud contamination, and $P_{\text{con}-\text{urban}}$ and $P_{\text{con}-\text{suburbans}}$ are the numbers of clear-sky pixels in cities and suburban areas, respectively, under cloud contamination conditions.

IV. RESULTS

A. WRF LST Image Results and Validation

Fig. 5 shows the WRF LST images and MODIS LST images obtained at 13:30 local time on July 6 (a) and 29 (d), and the nighttime images include WRF LST images and MODIS LST images taken at 1:30 local time on December 20 (c) and 31 (d), 2020. The high-temperature areas of these WRF LST images all corresponded to the construction lands in the land use type data [Fig. 1(b)], and the temperatures around urban areas were relatively low; that is, the WRF LSTs effectively captured the characteristics of UHIs, and they exhibited a similar spatial pattern to the MODIS LSTs of the corresponding date. To verify the difference between the WRF LSTs and the measured data, we compared the WRF LSTs obtained on July 6, July 29, December 20, and December 31, 2020, with the hourly temperatures recorded at the automatic weather stations to obtain the daily average temperature variation time series (Fig. 6). The daily maximum temperature occurred at 2:30 pm local time (UTC+8) and the lowest temperature occurred at approximately 6:00 pm local time. The Pearson correlation coefficient (Cor) values of the two data sources on July 6, July 29, December 20, and December 31, 2020 were above 0.8, with values of 0.818, 0.964, 0.809, and 0.816, respectively, and the corresponding RMSEs were 1.388, 0.896, 1.637, and 1.658, respectively, indicating that the WRF model optimally simulated the diurnal variations in air temperature, which is reasonable to some extent.

B. Generating Aqua-Terra/MODIS LSTs

Due to the influence of cloud contamination, the obtained Aqua/MODIS LST series had discontinuous data spaces due to missing pixels. However, sufficient effective pixels are required when fitting the WRF LSTs and MODIS LSTs using the RF model. In this article, we used Terra/MODIS clear-sky pixels to fill in the cloud-polluted pixels of the Aqua/MODIS data. Fig. 7 shows the average summertime 1) and wintertime 2) deviations in the Terra-Aqua data from 2013 to 2020. In Heping District, Hedong District, Hexi District, Nankai District, Hebei District, and Hongqiao District, the summertime surface temperature deviations were more obvious than those in winter. Taking the daytime Aqua/MODIS data obtained on July 5, 2020, as an example (Fig. 8), the pixels that filled the blank areas in the
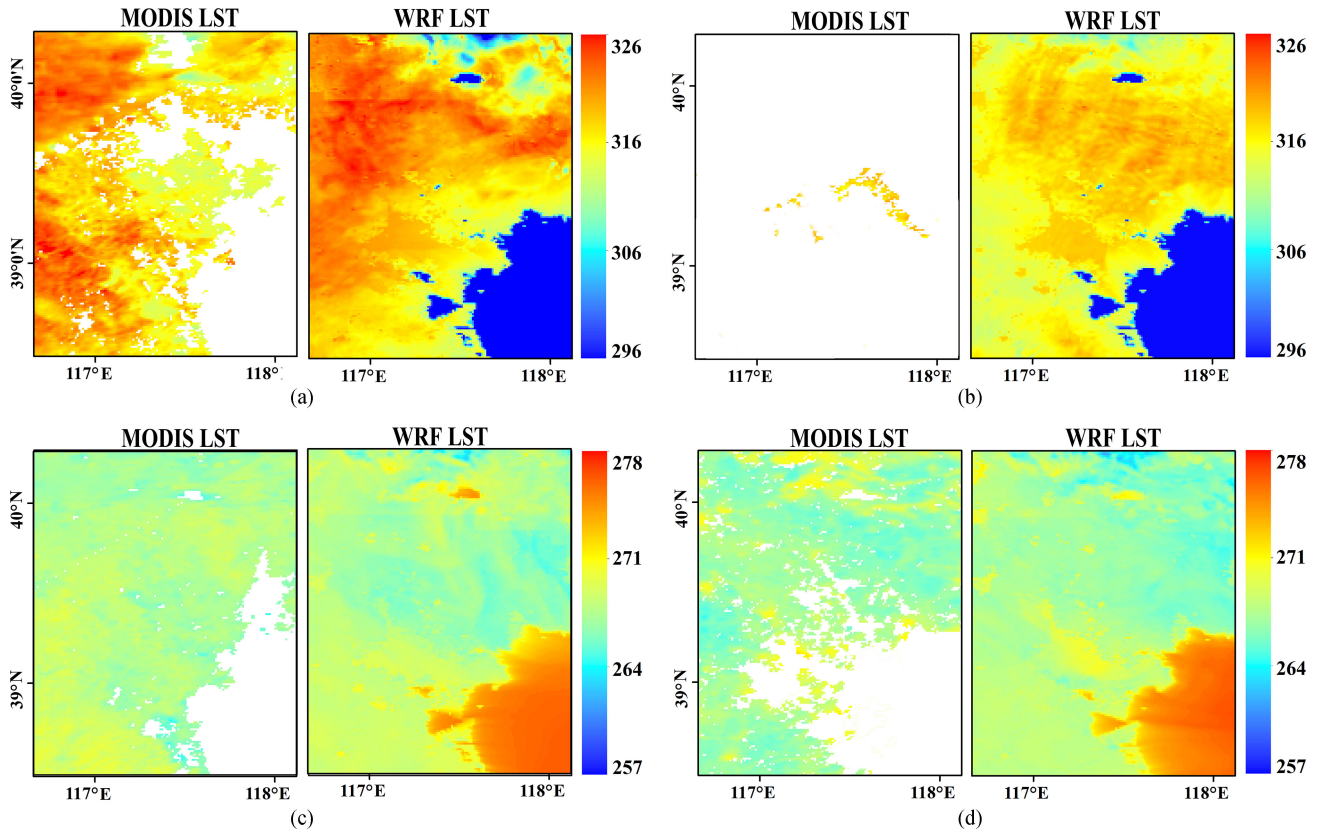
Fig. 5. WRF LST images and MODIS LST images on July 6 (a) and 29 (b), 2020 (13:30 pm), and on December 20 (c) and 31 (d), 2020 (1:30 am).

TABLE III
PERCENTAGES OF CLEAR-SKY LST PIXELS BEFORE AND AFTER THE DAYTIME AND NIGHTTIME TERRA AND AQUA DATA MERGERS IN TIANJIN IN JULY AND DECEMBER 2020

Month	Daytime		Nighttime	
	Aqua (13:30)	Terra (10:30)	Aqua (1:30)	Terra (22:30)
July (before)	25.18%		34.44%	
July (after)	29.45%	28.5%	35.56%	29.16%
December (before)	46.02%		46.46%	
December (after)	49.73%	48.9%	49.59%	47.48%

Aqua data with Terra data were scattered, and the effective pixel coverage of the Aqua data was subsequently greatly increased, especially in the main urban area. Table III shows the percentages of clear-sky LST pixels before and after the merging of the Terra and Aqua data in July and December 2020 in Tianjin. Generally, there were more clear-sky pixels in the MODIS data in December than in July. By supplementing with Terra data obtained at the corresponding times, the effective pixels in the LST data obtained by the Aqua satellite in daytime and nighttime in July increased from 25.18% to 29.45% and from 34.44% to 35.56%, respectively, exhibiting effective clear-sky

pixel coverage increases of 4.27% and 3.71%, respectively. In December, the daytime and nighttime values improved from 46.02% to 49.73% and from 46.46% to 49.59%, respectively, representing increases of 3.59% and 3.13%, respectively. These improvements played an important role in improving the fitting accuracy of the RF model.

C. WRFFM Model Results and Validation

In this section, we presented the simulation results of WRFFM model of this article. To further illustrate the results of our article, we presented simulations results of RF model, TRIMS LST proposed by Zhou et al. [67], and MKF-M LST images generated by Xu et al. [68]. TRIMS LST is generated by the satellite TIR remote sensing reanalysis data integration method based on the new surface temperature time decomposition model [69], [70], [71]. MKF-M LST product is generated based on MKF fusion TIR and PMW LST. Both TRIMS LST product and MKF-M LST product were obtained from the National Qinghai-Tibet Plateau Data Center http://data.tpdc.ac.cn/ [72], [73], and we compared them in spatial mode and quantitative accuracy (Fig. 9). First, under the condition that the MODIS data (collectively referred to as the Aqua and Terra-Aqua merged data in this article) were partially cloud-contaminated, the MODIS LST and WRF LST pixels were randomly divided into a training set and a verification set at a proportion of 7:3 for the RF model regression to fill in the blank MODIS LST area. Fig. 9(a), (f),
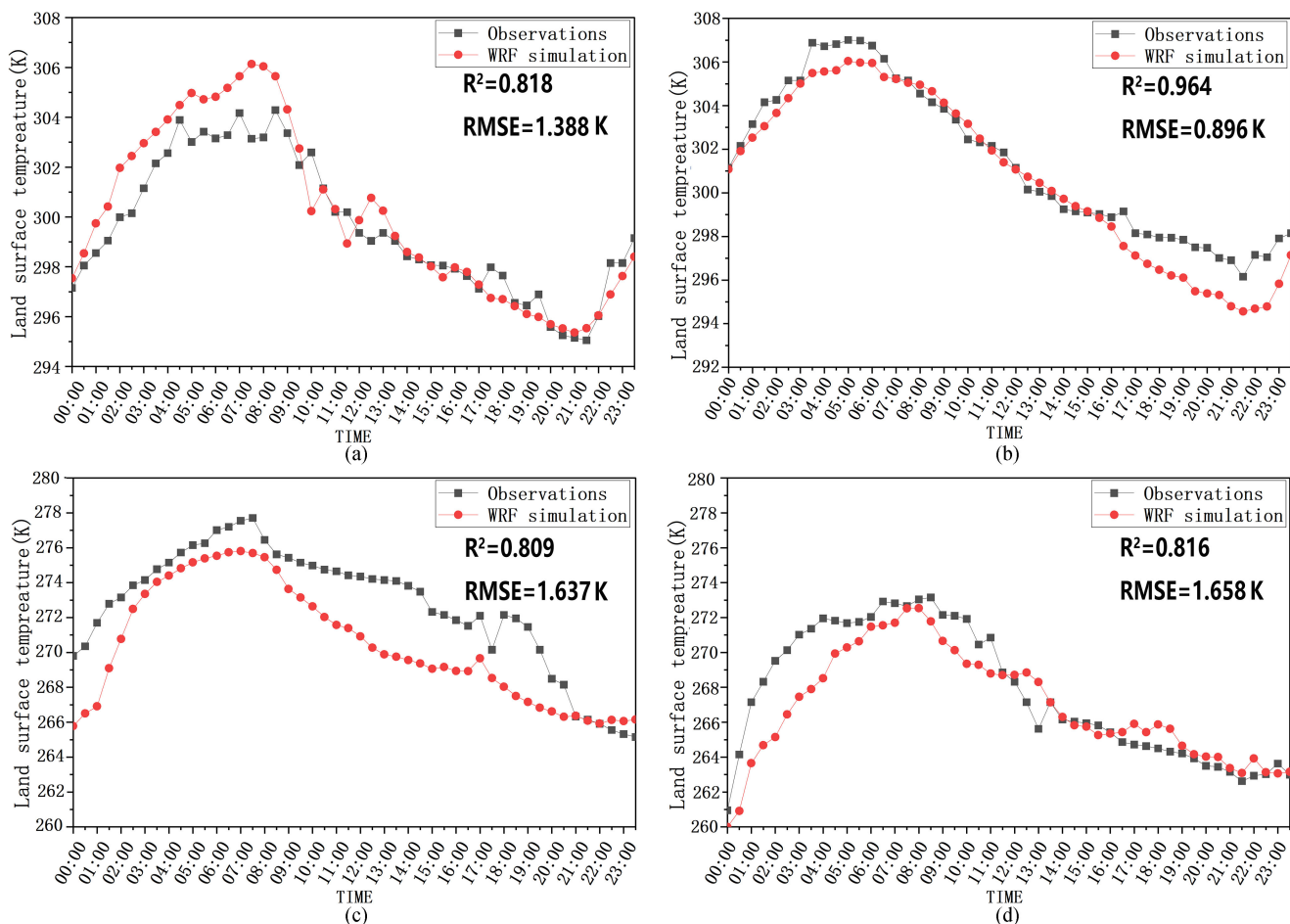
Fig. 6. Scatter density map of weather station and WRF simulation data obtained on July 6 (a), July 29 (b), December 20 (c), and December 31 (d), 2020. The weather station number is 5452732 (117.3462°E, 39.1243°N, and the elevation is 3.04 m).
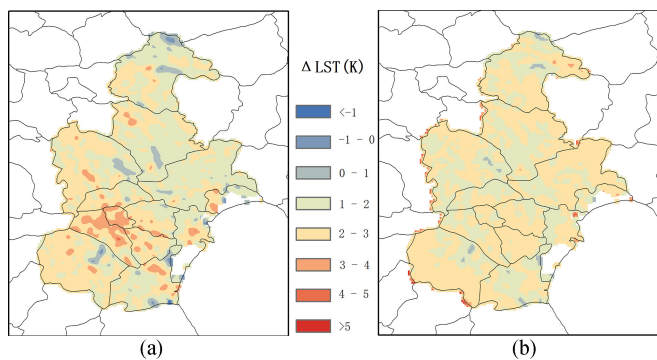

Fig. 7. Average summertime offset of the Terra-Aqua data collected in Tianjin from 2013 to 2020 (a) and average wintertime offset of the Terra-Aqua data collected in Tianjin from 2013–2020 (b).
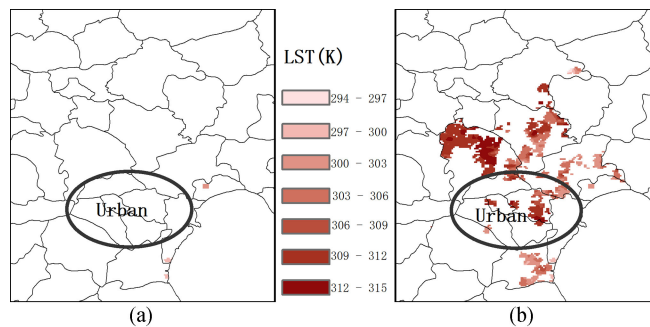

Fig. 8. Aqua satellite MYD11A1 data collected over Tianjin on July 5, 2020 (a) and the Terra-Aqua merged data (b).

(k), and (p) shows the MODIS LST images representing July 6, July 22, December 12, and December 18, 2020. Due to cloud contamination, the pixels in these images are spatially discontinuous and exhibit blank areas. After the RF model was regressed, the blank areas were filled in, but a "salt and pepper noise" phenomenon was apparent; therefore, we used the spatial optimization algorithm to optimize the results. Fig. 9 shows

that the WRFFM LSTs were relatively smooth following this filtering and exhibited very similar spatial distribution characteristics as the MODIS LSTs. Fig. 9(d), (i), (n), and (t) shows that MKF-M LST images are relatively "hazy." Fig. 9(e), (j), (o), and (u) shows that TRIMS LST images have relatively clear spatial distribution details which were similar to the MODIS LST.

In order to quantitatively compare the accuracy of WRFFM LST, MKF-M LST, and TRIMS LST images, a performance
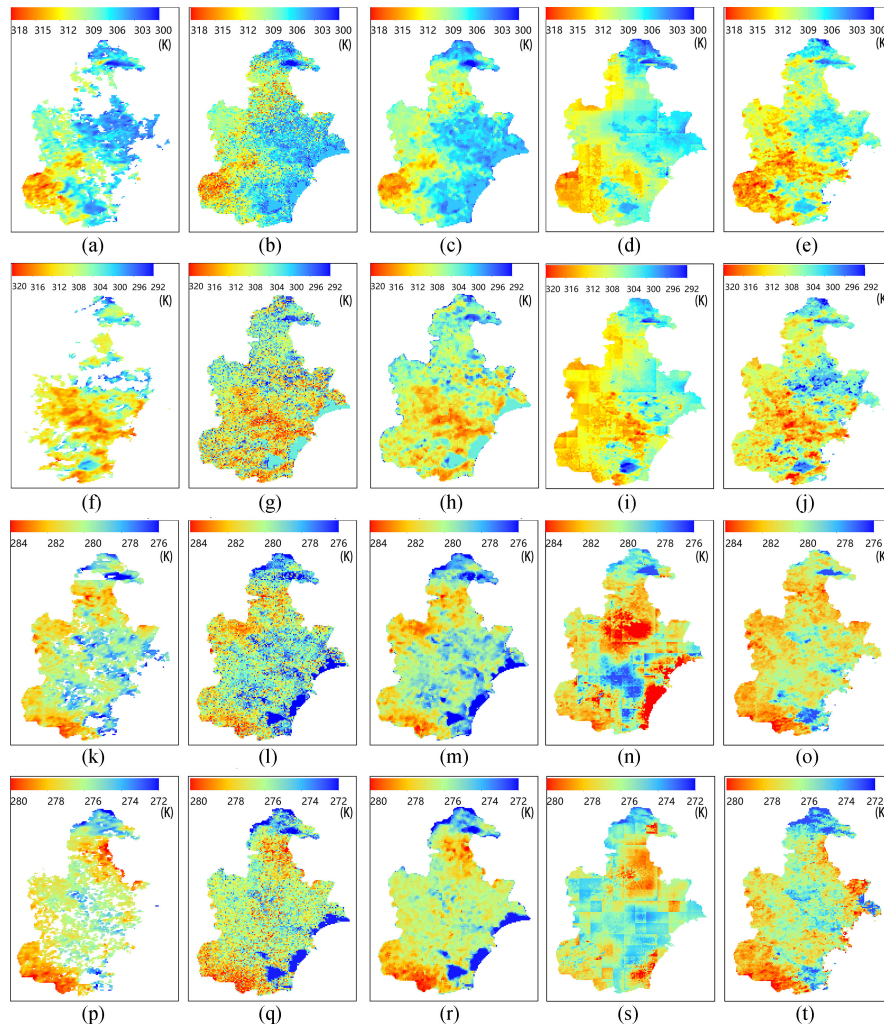
Fig. 9. MODIS LST images obtained on July 6 (a), July 22 (f), December 12 (k), and December 18 (p), 2020; RF LST images obtained for the corresponding dates [(b), (g), (l), and (r)]; WRFFM LST images obtained for the corresponding dates [(c), (h), (m), and (s)]; MKF-M LST images obtained for the corresponding dates [(d), (i), (n), (t)]; and TRIMS LST images obtained for the corresponding dates [(e), (j), (o), (u)].

measure RMSE and Pearson correlation coefficient Cor were used to verify them (Fig. 10). The relationship between the MODIS original images and three product LSTs can be seen directly from the scatter density map of the numerical. The correlation coefficients corresponding to July 6, July 22, December 12, and December 18, 2020 were ~0.875, ~0.841, and ~0.848. The correlation coefficients between TRIMS LST and MODIS LST are above 0.75, and the correlation coefficients of MKF-M LST are smaller than the former two. The RMSEs of WRFFM LSTs were ~1.82, ~3.06, ~0.90, and ~0.83 K, respectively. Compared with MKF-M LSTs and TRIMS LSTS, RMSE is smaller, indicating that the WRFFM LSTs were more consistent with the MODIS LSTs. The observed deviations can be explained by the upward sensible heat flux provided by the WRF model and the offset of solar shortwave radiation [74].

The MODIS images contaminated by clouds in July and December 2020 were filled, and the numbers of days in which valid data were input to the WRFFM model during the daytime in July and December were 22 and 28, respectively. Fig. 11 shows the changes in the Cor and RMSE values between the LSTs and

MODIS LSTs from before to after the regression and spatial optimization algorithm were applied. The July RMSEs 1) were generally higher than those of December 2). The RMSEs of July were between 0.5 and 3.8 K, while the December RMSEs were between 0.4 and 1.8 K. The RMSEs of the optimized images were reduced compared to those obtained before optimization. The distribution ranges of the Cor values derived for the two months before and after optimization were similar. The Cor values derived before optimization were almost entirely below 0.8; after optimization, Cor in July and December ranged from 0.8 to 0.91 and 0.8 to 0.93, respectively, indicating that the model results were similar to the MODIS LSTs and could thus be used as MODIS-like cloud-free images to further research the SUHI effect.

D. SUHI Intensities Under Cloudy Conditions

Based on the cloud-free MODIS-like LST images optimized by the WRFFM model, the daytime and nighttime SUHIs derived in July (a) and December (b), 2020, were obtained, as
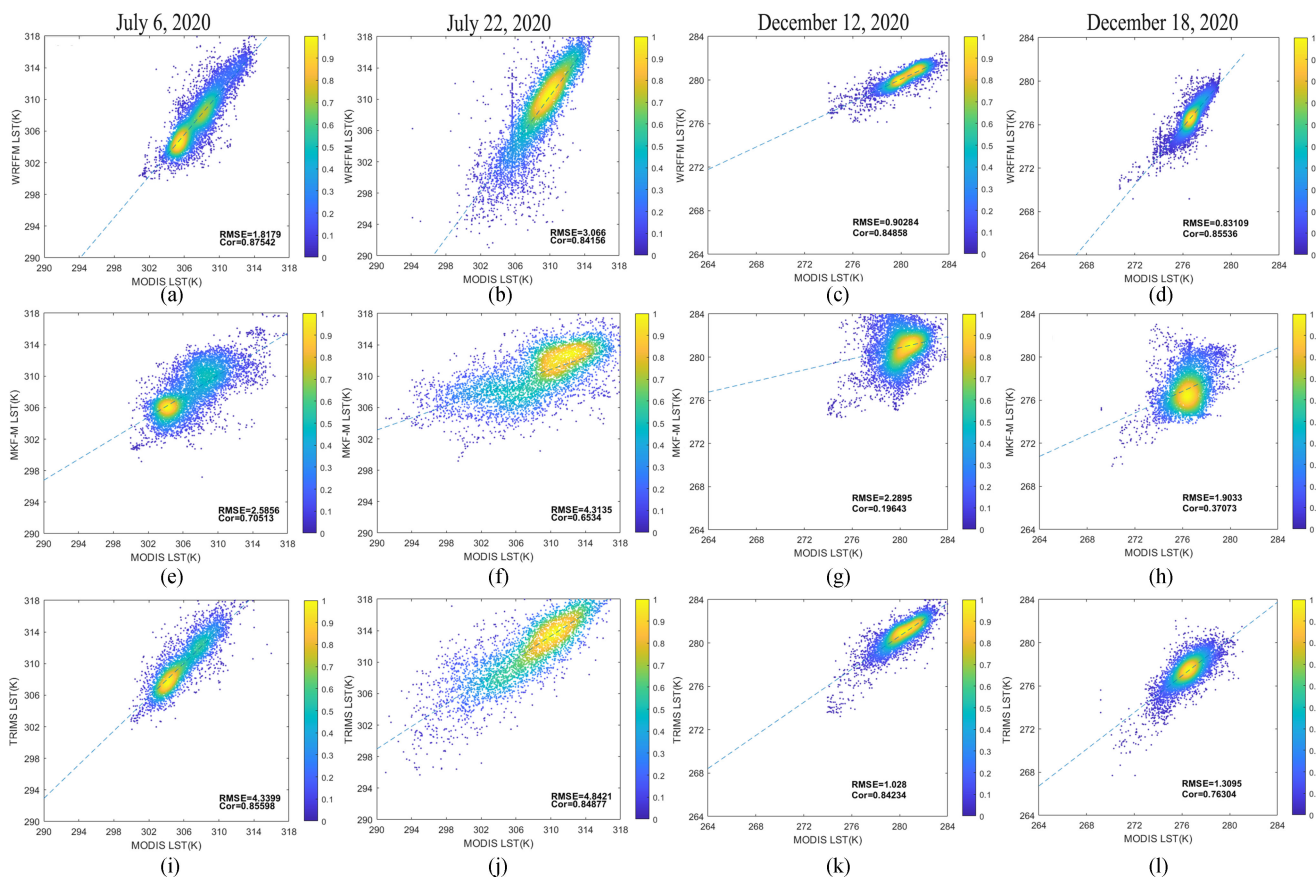
Fig. 10.   Scatter density map of WRFFM LST and MODIS LST (a)–(d). Scatter density map of TRIMS LST and MODIS LST (e)–(h). Scatter density map of MKF-M LST and MODIS LST (i)–(l).
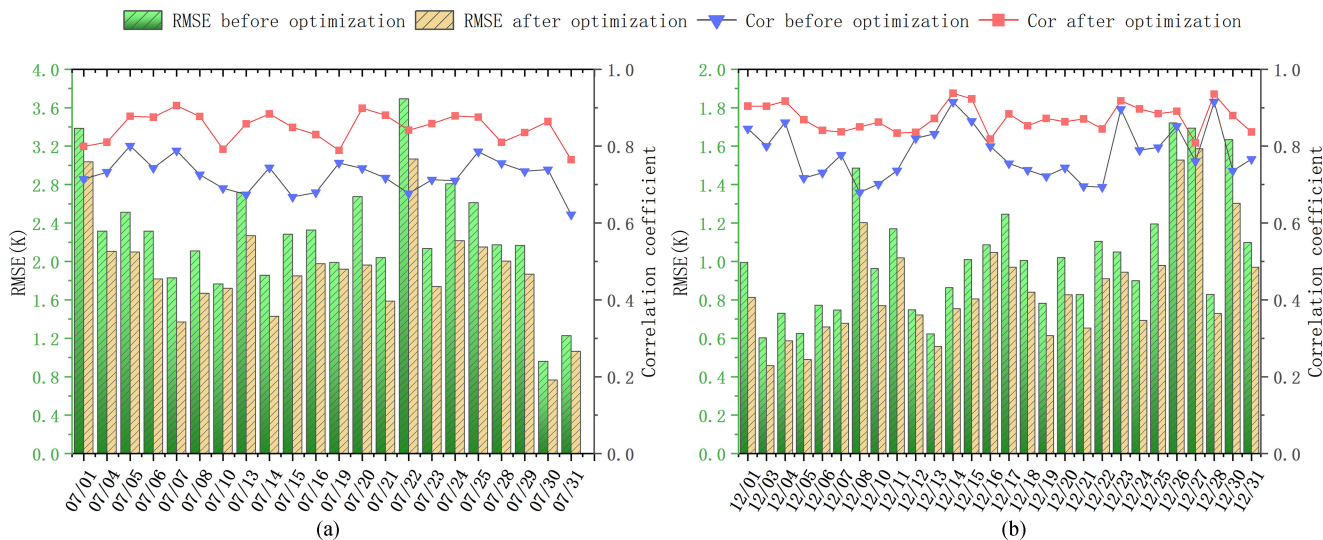

Fig. 11.   Comparison of RMSE errors and Pearson correlation coefficients derived before and after the image optimization process.

shown in Fig. 12. According to the grades of the SUHI effect on surface cities listed in Table II, on most days in July, SUHIs were more intense during the day than at night, and half of the days exhibited "weak SUHIs." From July 28 to 31, the daytime heat island effect was weaker than the nighttime effect. Even on July 19 and 30, "weak SUCI" phenomena appeared, and among

nighttime observations, the number of days with "no SUHI" effect was highest, followed by the number of days with "weak SUHI" effects. On the whole, in December, nighttime UHIs were more intense than daytime SUHIs, and most identified SUHIs were nighttime "weak SUHIs"; almost no heat islands were observed in the daytime, with only one daytime "weak
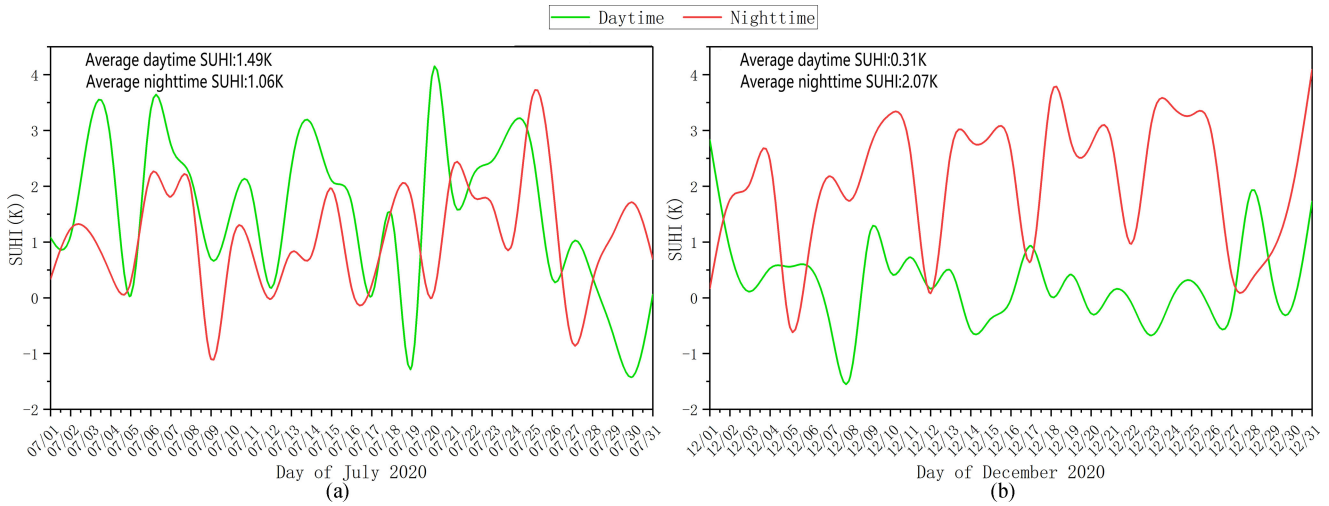
Fig. 12. UHI effects (a) and (b) during the daytime and nighttime in July and December 2020.
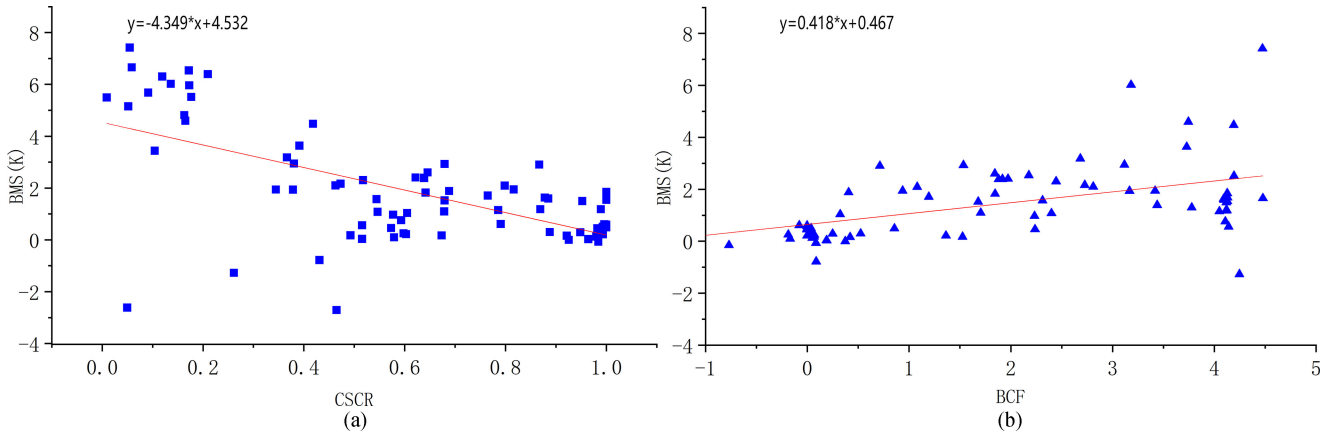

Fig. 13. Relationship between the urban and suburban clear-sky pixel coverage ratio (CSCR) and the deviations in SUHI intensity between MODIS LSTs and WRFFM LSTs under different weather conditions (BMS) (a); biases of the ratio of urban clear-sky pixels to suburban clear-sky pixels under cloud-contamination conditions and the same ratio under cloud-free conditions (BCF) and BMS (b).

SUHI" appearing, on December 8. In July, the average heat island temperature in daytime was 0.43 K higher than that at night, and in December, the average heat island temperature at night was 1.76 K higher than that in the daytime, indicating that the SUHI effect in Tianjin undergoes great seasonal differences.

To the best of our knowledge, current SUHI research that is based on LSTs obtained by MODIS is limited to clear-sky conditions, so the results cannot accurately reflect the actual SUHI intensities on certain days. To quantify the biases caused by performing MODIS-guided SUHI research only under clear-sky conditions, we investigated the relationship between the CSCR and SUHI intensity bias (BMS); the results are shown in Fig. 13(a). As the number of filled clear-sky pixels increases, the overall SUHI intensity bias decreases, when the proportion of clear-sky pixels is below 30%, the deviation is above 3 K, and when the proportion of clear-sky pixels is above 80%, the deviation is below 0.6 K. When the pixel percentage is between 45% and 65%, the BMS value does not vary extensively.

According to the SUHI calculation index utilized in this article, if no cloud contamination is present in a city or suburbs, the ratio of the number of pixels representing cities and suburbs must be approximately 1:19; that is, each urban pixel needs to correspond to 19 suburban pixels. Under these conditions, the bias of the derived SUHI will not be excessively large. However, because the locations of cloud-contaminated MODIS LSTs are random, this ratio cannot be fixed, so researching the bias between the urban-to-suburban-pixel ratio under cloud-free and cloudy (BCF) conditions and the corresponding SUHI intensity bias (BMS) is necessary. Fig. 13(b) shows that as the BCF increases, according to formula (8), the proportion of clear-sky pixels in suburbs becomes smaller than that in urban areas; that is, when the ratio of urban pixels to suburban pixels exceeds 1:19, the BMS increases accordingly, with a maximum difference of 7.41 K. In summary, the BMS is greater under cloudy conditions than under cloud-free conditions, indicating that the MODIS SUHI has a large deviation from the actual SUHI under cloudy

conditions, and that the number of clear-sky pixels in a city is an important factor influencing the bias relative to the actual SUHI.

V. DISCUSSION

A. Comparisons With Other Studies on LST Reconstruction Methods

When MODIS data were used to study the SUHI effect, spatial discontinuities caused by cloud contamination arose in the image data when the sensor obtained surface information, and blank areas appeared when pixels were missing. Such conditions can greatly affect research on the SUHI effect. Several studies [75], [76], [77] have addressed this issue by quantifying only nighttime heat islands, as daytime solar radiation increases the instability of the boundary layer and makes cloud formation more likely [13]. In addition, in some studies, only cloud-free LST images [78], [79], [80] have been selected for short-term case studies, or the mean has been used for long-term series studies; however, these approaches are not suitable for monitoring large areas. In this article, we combined remote sensing and meteorological methods to give full play to the advantages of both. By introducing WRF LSTs and MODIS LSTs into the RF model-fitting and subsequent optimization processes, clear-sky MODIS-like LST images with high spatial continuity and high accuracy were obtained. We then researched the nighttime and daytime SUHIs in Tianjin city. Like the WRFF model established by Fu et al. [43], the LSTs under cloudy conditions were reconstructed based on the WRF model and the RF model, but in the case of complete cloud contamination, LSTs were retrieved by using temporally adjacent images, which may not hold when the search for temporally adjacent images of the target date is extended over a longer period of time. In this article, the WRFFM model integrated the information of the two satellite MODIS images of the day to ensure the validity of the independent variables entering the RF model. Second, in the case of complete cloud contamination, we used the time-adjacent model for weighted calculation, while instead of temporally adjacent images, the results are more plausible. Compared with other filling methods, the model in this article has an advantage in filling pixels contaminated by clouds in that it considers land surface and atmospheric conditions when forming a coupled WRF/UCM simulation framework [81]. Furthermore, Terra and Aqua satellite-derived data were used to ensure that a sufficient number of effective pixels were available to input into the RF model for fitting, thus greatly improving the accuracy of the model retrieval results, and the WRFFM model can obtain LST under cloudless and cloudy conditions at the same time with high accuracy. In model verification (Fig. 10), the optimized RMSEs in July 2020 were found to be in the range of 0.5–4 K, while the December RMSEs ranged from 0.4 to 1.8 K. Compared with LST reconstruction algorithms of previous articles, such as PMW (yielding RMSEs between 2 and 5 K), and SEB model to infer LST under cloudy conditions (yielding RMSEs between 3 and 5 K), and the gap-filling method (yielding RMSEs of approximately 3.35 K) [82], [83], [84], [85], which mostly rely on spatiotemporal information [82], [86], the WRFFM model

has strong overall performance, and the RMSEs are within an acceptable range.

B. Applications in SUHI Researches

The WRFFM model simulation results show that the daytime and nighttime SUHIs derived in Tianjin in July and December of 2020 exhibit obvious seasonal characteristics. Wang et al. [87] pointed out that urban–suburban surface thermal differences, which result in daytime UHIs, are driven by the seasonal solar heat radiation driver. In July, which is in summertime, the daytime solar thermal radiation of the Sun is intense, and heat is emitted from buildings and via transportation and the daily activities of residents throughout the city; in contrast, in suburban areas, the vegetation is denser and has a larger surface heat capacity, which can reduce the LSTs [88]. In winter, sunset occurs early, sunrise occurs late, the days are short, and the nights are long; these conditions are conducive to the formation of inversion layers [89]. In addition, in wintertime, there is no solar radiation at night, the suburban vegetation coverage is low, and the heat capacity of the land surface is relatively small [90], [91]; thus, the nighttime heat islands are more intense than the daytime heat islands in wintertime. A comparison of the WRFFM SUHIs with the cloud-contaminated MODIS SUHIs revealed that, the number of urban/suburban clear-sky pixels was not significantly negatively correlated with the SUHI deviation bias; this result was observed because cloud layers have a mitigating effect on SUHI intensity [92] and because the strength of this effect differs between day and night. Due to the randomness of the spatial distribution of clouds, the ratio of the number of urban pixels to the number of suburban pixels is greater than 1:19, so high numbers of urban and suburban clear-sky pixels can be observed. However, the calculated SUHI BMS deviations were large. Romanov [93] studied the influence of cloud cover on UHIs and found that cloud cover was more expansive in summer than in winter, which is consistent with our findings (Fig. 10). The height of the boundary layer is generally higher in urban regions than in suburban areas due to the higher surface temperatures of urban areas; thus, low-level coupling on the urban surface can affect the formation and location of clouds [94]. Cloudy urban environments affect SUHI research [95], and remote sensing can play an optimal role only under clear and cloudless conditions; thus, the results obtained with the model developed in this article can provide a reference for long-term SUHI monitoring.

C. Limitations and Future Perspectives

Because MODIS does not yield the correct LST under clouds, our model predicted the LST under clouds based on clear-sky pixels. However, as neither the validation set nor the pixels under clouds participated in the model training, the validation set can be regarded as the pixels under clouds. The accuracy attained with the two sets is the same, and the accuracy of the validation set is included in the overall accuracy, so the LST filled by this model can be regarded as the LST under clouds. This article has some limitations. The LSTs simulated by the WRF model developed in this article were more sensitive to the utilized physical scheme than the results obtained using

traditional models [96], although a complete analysis of the WRF parameter changes was computationally expensive [97]. However, as cloud computing resources become available, the proposed WRFFM model has high application potential. In future article, data assimilation techniques can be introduced to use multisource remote sensing data to drive the WRFFM model and thereby estimate LST under cloudy conditions with greater accuracy. In addition, in this article, we used the results of the WRFFM model to quantify SUHIs in July and December 2020. In our future article, we will continue to investigate in detail the changes in derived long-term SUHI series and the driving factors.

VI. CONCLUSION

In this article, we took Tianjin as the research area, and by combining remote sensing and meteorological methods for application under cloudy conditions, merged the WRF model and RF method to construct a highly practical model (WRFFM) for generating high-quality MODIS-like LST data to research SUHIs. This article shows that the WRF simulation results are consistent with the hourly temperature data monitored by local automatic weather stations, and the extracted four-day verification results show that the correlation coefficients were above 0.8 and the RMSEs were between 0.8 and 1.7 K. To use Terra/MODIS-derived surface temperatures to fill gaps in Aqua/MODIS data caused by cloud contamination, we adjusted for the different overpass times of the two satellite platforms at different overpass times. The effective pixel coverage of the Terra-Aqua/MODIS data was compared with that of the Aqua/MODIS data, and the coverage of the Terra-Aqua/MODIS results during the daytime and nighttime were increased by 4.27% and 3.71%, respectively, in July and by 3.59% and 3.13%, respectively, in December. The WRFFM model results show that the WRFFM LSTs and MODIS LSTs exhibit very similar spatial patterns and can capture the UHI effect well; the correlation coefficients describing the relationship between these two datasets were above 0.8. The RMSE ranges were 0.5–3.8 K and 0.4–1.8 K in July and December, respectively; thus, the model accuracy was high. The SUHIs identified in Tianjin in July and December based on the WRFFM model results suggested that July was characterized as "high in the day and low in the night" SUHIs, and the average daytime heat island was 0.73 K warmer than the average nighttime heat island. December was characterized "low in the day and high in the night" SUHIs, and the average nighttime heat island was 1.81 K warmer than the average daytime heat island, revealing obvious seasonal differences. When the clear-sky pixel percentage was below 30%, the bias was above 3 K; when the clear-sky pixel percentage was above 80%, the bias was below 0.6 K. Thus, filled city pixels are an important factor affecting the measured SUHI intensity.

REFERENCES

[1] C. Lahariya, "The state of the world population 2007: Unleashing the potential of urban growth," *Indian Pediatr.*, vol. 45, no. 6, Jun. 2008, Art. no. 476.

[2] X. Huang, J. Y. Huang, D. W. Wen, and J. Y. Li, "An updated MODIS global urban extent product (MGUP) from 2001 to 2018 based on an automated mapping approach," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 95, Mar. 2021, Art. no. 102255.

[3] K. C. Wang, J. K. Wang, P. C. Wang, M. Sparrow, J. Yang, and H. B. Chen, "Influences of urbanization on surface characteristics as derived from the moderate-resolution imaging spectroradiometer: A case study for the Beijing metropolitan area," *J. Geophys. Res.-Atmos.*, vol. 112, no. D22, Aug. 2007, Art. no. D22s06.

[4] U. Eichelmann, *Boundary Layer Climates*. London, U.K.: Routledge, 1987, pp. 265–265.

[5] G. Manley, "On the frequency of snowfall in metropolitan England," *Quart. J. Roy. Meteorol. Soc.*, vol. 84, no. 359, pp. 70–72, 1958.

[6] G. Ren et al., "An integrated procedure to determine a reference station network for evaluating and adjusting urban bias in surface air temperature data," *J. Appl. Meteorol. Climatol.*, vol. 54, no. 6, pp. 1248–1266, Jun. 2015.

[7] S. H. Mahmoud and T. Y. Gan, "Long-term impact of rapid urbanization on urban climate and human thermal comfort in hot-arid environment," *Building Environ.*, vol. 142, pp. 83–100, Sep. 2018.

[8] X. Y. Fang et al., "The multi-scale numerical modeling system for research on the relationship between urban planning and meteorological environment," *Adv. Atmos. Sci.*, vol. 21, no. 1, pp. 103–112, Jan. 2004.

[9] M. Santamouris, C. Cartalis, A. Synnefa, and D. Kolokotsa, "On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—A review," *Energy Buildings*, vol. 98, pp. 119–124, Jul. 2015.

[10] N. B. Grimm et al., "Global change and the ecology of cities," *Science*, vol. 319, pp. 756–760, 2008.

[11] W. Chen, F. Zhang, Y. Zhu, L. Yang, and P. Bi, "Analysis of the impact of multiscale green landscape on urban PM2.5," *Air Qual. Atmos. Health*, vol. 15, pp. 1319–1332, 2022.

[12] Q. Weng, U. Rajasekar, and X. Hu, "Modeling urban heat islands and their relationship with impervious surface and vegetation abundance by using ASTER images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 4080–4089, Oct. 2011.

[13] J. A. Voogt and T. R. Oke, "Thermal remote sensing of urban climates," *Remote Sens. Environ.*, vol. 86, no. 3, pp. 370–384, Aug. 2003.

[14] Y. Zhang, P. Jiang, H. Zhang, and P. Cheng, "Study on urban heat island intensity level identification based on an improved restricted Boltzmann machine," *Int. J. Environ. Res. Public Health*, vol. 15, no. 2, Feb. 2018, Art. no. 186.

[15] R. Yao, L. Wang, X. Huang, W. Zhang, J. Li, and Z. Niu, "Interannual variations in surface urban heat island intensity and associated drivers in China," *J. Environ. Manage.*, vol. 222, pp. 86–94, Sep. 2018.

[16] J. Wang, B. Huang, D. Fu, P. M. Atkinson, and X. Zhang, "Response of urban heat island to future urban expansion over the Beijing-Tianjin-Hebei metropolitan area," *Appl. Geogr.*, vol. 70, pp. 26–36, May 2016.

[17] J. S. Silva, R. M. da Silva, and C. A. Guimaraes Santos, "Spatiotemporal impact of land use/land cover changes on urban heat islands: A case study of Paco do Lumiar, Brazil," *Building Environ.*, vol. 136, pp. 279–292, May 2018.

[18] Y. Lin, C. Y. Jim, J. Deng, and Z. Wang, "Urbanization effect on spatiotemporal thermal patterns and changes in Hangzhou (China)," *Building Environ.*, vol. 145, pp. 166–176, Nov. 2018.

[19] L. Prihodko and S. N. Goward, "Estimation of air temperature from remotely sensed surface observations," *Remote Sens. Environ.*, vol. 60, no. 3, pp. 335–346, 1997.

[20] G. V. Mostovoy, R. L. King, K. R. Reddy, V. G. Kakani, and M. G. Filippova, "Statistical estimation of daily maximum and minimum air temperatures from MODIS LST data over the state of Mississippi," *Gisci. Remote Sens.*, vol. 43, no. 1, pp. 78–110, Jan.–Mar. 2006.

[21] 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.

[22] J. P. A. Martins et al., "An all-weather land surface temperature product based on MSG/SEVIRI observations," *Remote Sens.*, vol. 11, no. 24, Dec. 2019, Art. no. 3044.

[23] 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.

[24] L. Hu and N. A. Brunsell, "The impact of temporal aggregation of land surface temperature data for surface urban heat island (SUHI) monitoring," *Remote Sens. Environ.*, vol. 134, pp. 162–174, Jul. 2013.

[25] Z. Wan, "New refinements and validation of the MODIS land-surface temperature/emissivity products," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 59–74, Jan. 2008.

[26] Y. P. Mo, Y. M. Xu, H. J. Chen, and S. Y. Zhu, "A review of reconstructing remotely sensed land surface temperature under cloudy conditions," *Remote Sens.*, vol. 13, no. 14, Jul. 2021, Art. no. 2838.

[27] 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, Mar. 2014.

[28] L. Ke, X. Ding, and C. Song, "Reconstruction of time-series MODIS LST in central Qinghai-Tibet plateau using geostatistical approach," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1602–1606, Nov. 2013.

[29] J. Kang, J. Tan, R. Jin, X. Li, and Y. Zhang, "Reconstruction of MODIS land surface temperature products based on multi-temporal information," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 1112.

[30] 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.

[31] R. Van De Kerchove, S. Lhermitte, S. Veraverbeke, and R. Goossens, "Spatio-temporal variability in remotely sensed land surface temperature, and its relationship with physiographic variables in the Russian Altay mountains," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 20, pp. 4–19, Feb. 2013.

[32] Z. Liu et al., "Balancing prediction accuracy and generalization ability: A hybrid framework for modelling the annual dynamics of satellite-derived land surface temperatures," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 189–206, May 2019.

[33] Y. Julien, J. A. Sobrino, and W. Verhoef, "Changes in land surface temperatures and NDVI values over Europe between 1982 and 1999," *Remote Sens. Environ.*, vol. 103, no. 1, pp. 43–55, Jul. 2006.

[34] 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, Sep. 2015.

[35] Z. Yin et al., "Spatiotemporal fusion of land surface temperature based on a convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1808–1822, Feb. 2021.

[36] 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, Sep. 2020, Art. no. 111863.

[37] P. Wu et al., "Spatially continuous and high-resolution land surface temperature product generation: A review of reconstruction and spatiotemporal fusion techniques," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 3, pp. 112–137, Sep. 2021.

[38] 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, May 2016.

[39] P. H. Wu et al., "A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data," *Remote Sens. Environ.*, vol. 277, Aug. 2022, Art. no. 113070.

[40] C. R. Hain and M. C. Anderson, "Estimating morning change in land surface temperature from MODIS day/night observations: Applications for surface energy balance modeling," *Geophys. Res. Lett.*, vol. 44, no. 19, pp. 9723–9733, Oct. 2017.

[41] M. L. 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, Feb. 2000.

[42] 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, Feb. 2019, Art. no. 336.

[43] P. Fu, Y. Xie, Q. Weng, S. Myint, K. Meacham-Hensold, and C. Bernacchi, "A physical model-based method for retrieving urban land surface temperatures under cloudy conditions," *Remote Sens. Environ.*, vol. 230, Sep. 2019, Art. no. 111191.

[44] L. R. Diaz et al., "Land surface temperature retrieval using high-resolution vertical profiles simulated by WRF model," *Atmosphere*, vol. 12, no. 11, Nov. 2021, Art. no. 1436.

[45] Y. Wang, R. Amundson, and S. Trumbore, "The impact of land use change on C turnover in soils," *Glob. Biogeochem. Cycles*, vol. 13, no. 1, pp. 47–57, 1999.

[46] F. O. Becker and Z. L. Li, "Surface temperature and emissivity at various scales: Definition, measurement and related problems," *Remote Sens. Rev.*, vol. 12, no. 3–4, pp. 225–253, 1995.

[47] W. L. Crosson, M. Z. Al-Hamdan, S. N. J. Hemmings, and G. M. 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.

[48] M. Chen, Y. Zhou, M. Hu, and Y. Zhou, "Influence of urban scale and urban expansion on the urban heat island effect in metropolitan areas: Case study of Beijing–Tianjin–Hebei urban agglomeration," *Remote Sens.*, vol. 12, no. 21, 2020, Art. no. 3491.

[49] Y.-Z. Zhang, X.-G. Jiang, and H. Wu, "A generalized split-window algorithm for retrieving land surface temperature from GF-5 thermal infrared data," in *Proc. Prog. Electromagn. Res. Symp. Spring*, 2017, pp. 2766–2771.

[50] Z. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sens. Environ.*, vol. 140, pp. 36–45, Jan. 2014.

[51] R. Qu, X. Cui, H. Yan, E. Ma, and J. Zhan, "Impacts of land cover change on the near-surface temperature in the north China plain," *Adv. Meteorol.*, vol. 2013, 2013, Art. no. 409302.

[52] W. Chen, X. Zhang, and Y. Huang, "Spatial and temporal changes in ecosystem service values in karst areas in southwestern China based on land use changes," *Environ. Sci. Pollut. Res.*, vol. 28, pp. 45724–45738, 2021.

[53] C. B. Wei Wang, M. Duda, and J. Dudhia, WRF ARW V3:User's Guide, 2012. [Online]. Available: http://www.mmm.ucar.edu/wrf/users

[54] D. Li, E. Bou-Zeid, M. Barlage, F. Chen, and J. A. Smith, "Development and evaluation of a mosaic approach in the WRF-Noah framework," *J. Geophys. Res.-Atmos.*, vol. 118, no. 21, pp. 11918–11935, Nov. 2013.

[55] S. Leroyer, S. Belair, and J. Mailhot, "Microscale numerical prediction over Montreal with the Canadian external urban modeling system," *J. Appl. Meteorol. Climatol.*, vol. 50, no. 12, pp. 2410–2428, Dec. 2011.

[56] J. P. Evans and M. F. McCabe, "Regional climate simulation over Australia's Murray-Darling basin: A multitemporal assessment," *J. Geophys. Res.-Atmos.*, vol. 115, Jul. 2010, Art. no. D14114.

[57] A. Martilli, "Numerical study of urban impact on boundary layer structure: Sensitivity to wind speed, urban morphology, and rural soil moisture," *J. Appl. Meteorol.*, vol. 41, no. 12, pp. 1247–1266, Dec. 2002.

[58] H. Kusaka, H. Kondo, Y. Kikegawa, and F. Kimura, "A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models," *Boundary-Layer Meteorol.*, vol. 101, no. 3, pp. 329–358, 2001.

[59] F. Chen, H. Kusaka, R. Bornstein, J. Ching, and S. Miao, "The integrated WRF/urban modelling system: Development, evaluation, and applications to urban environmental problems," *Int. J. Climatol.*, vol. 31, pp. 273–288, 2011.

[60] J. Wei et al., "Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111221.

[61] G. Zhang and Y. Lu, "Bias-corrected random forests in regression," *J. Appl. Statist.*, vol. 39, no. 1, pp. 151–160, 2012.

[62] D. Lu and Q. Weng, "Use of impervious surface in urban land-use classification," *Remote Sens. Environ.*, vol. 102, no. 1–2, pp. 146–160, May 2006.

[63] 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.

[64] P. Zhang, M. L. Imhoff, R. E. Wolfe, and L. Bounoua, "Urban heat island effect across biomes in the continental USA," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2010, pp. 1920–1923.

[65] S. Chen and T. Wang, "A comparative study on the definition of urban heat island by equal interval method and mean standard deviation method," *J. Geo-Inf. Sci.*, vol. 11, no. 02, pp. 145–150, 2009.

[66] Q. Weng and P. Fu, "Modeling annual parameters of clear-sky land surface temperature variations and evaluating the impact of cloud cover using time series of landsat TIR data," *Remote Sens. Environ.*, vol. 140, pp. 267–278, Jan. 2014.

[67] J. Zhou, X. Zhang, W. Tang, L. Ding, J. Ma, and X. Zhang, "Daily 1-km all-weather land surface temperature dataset for the Chinese landmass and its surrounding areas (TRIMS LST; 2000-2021)," National Tibetan Plateau Data Center, 2021.

[68] S. Xu and J. Cheng, "A new land surface temperature fusion strategy based on cumulative distribution function matching and multiresolution Kalman filtering," *Remote Sens. Environ.*, vol. 254, 2021, Art. no. 112256.

[69] 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, Aug. 2017.

[70] X. Zhang, J. Zhou, S. Liang, and D. Wang, "A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature," *Remote Sens. Environ.*, vol. 260, 2021, Art. no. 112437.

[71] X. Zhang, J. Zhou, F. Goettsche, 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, Jul. 2019.

[72] Q. Zhang, N. Wang, J. Cheng, and S. Xu, "A stepwise downscaling method for generating high-resolution land surface temperature from AMSR-E data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5669–5681, 2020.

[73] Q. Zhang and J. Cheng, "An empirical algorithm for retrieving land surface temperature from AMSR-E data considering the comprehensive effects of environmental variables," *Earth Space Sci.*, vol. 7, no. 4, Apr. 2020, Art. no. UNSP e2019EA001006.

[74] T. M. Giannaros, D. Melas, I. A. Daglis, I. Keramitsoglou, and K. Kourtidis, "Numerical study of the urban heat island over Athens (Greece) with the WRF model," *Atmos. Environ.*, vol. 73, pp. 103–111, Jul. 2013.

[75] C. J. Tomlinson, L. Chapman, J. E. Thornes, and C. J. Baker, "Derivation of Birmingham's summer surface urban heat island from MODIS satellite images," *Int. J. Climatol.*, vol. 32, no. 2, pp. 214–224, Feb. 2012.

[76] M. Tiangco, A. M. F. Lagmay, and J. Argete, "ASTER-based study of the night-time urban heat island effect in metro manila," *Int. J. Remote Sens.*, vol. 29, no. 10, pp. 2799–2818, 2008.

[77] D. R. Streutker, "Satellite-measured growth of the urban heat island of Houston, Texas," *Remote Sens. Environ.*, vol. 85, no. 3, pp. 282–289, May 2003.

[78] B. Tan et al., "The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 98–114, Nov. 2006.

[79] S. Kato and Y. Yamaguchi, "Analysis of urban heat-island effect using ASTER and ETM+ data: Separation of anthropogenic heat discharge and natural heat radiation from sensible heat flux," *Remote Sens. Environ.*, vol. 99, no. 1–2, pp. 44–54, Nov. 2005.

[80] X.-L. Chen, H.-M. Zhao, P.-X. Li, and Z.-Y. Yin, "Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes," *Remote Sens. Environ.*, vol. 104, no. 2, pp. 133–146, Sep. 2006.

[81] W. C. Skamarock, J. B. Klemp, J. Dudhia, D. O. Gill, and J. G. Powers, "A description of the advanced research WRF version 2," *NCAR Tech.*, vol. 113, pp. 7–25, 2018.

[82] R. Yao et al., "A robust method for filling the gaps in MODIS and VIIRS land surface temperature data," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 12, pp. 10738–10752, Dec. 2021.

[83] W. Tan, C. Wei, Y. Lu, and D. Xue, "Reconstruction of all-weather daytime and nighttime MODIS aqua-terra land surface temperature products using an XGBoost approach," *Remote Sens.*, vol. 13, no. 22, Nov. 2021, Art. no. 4723.

[84] H. T. 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, Oct. 2019, Art. no. 101879.

[85] 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, Apr. 2011.

[86] P. Fu and Q. Weng, "Consistent land surface temperature data generation from irregularly spaced landsat imagery," *Remote Sens. Environ.*, vol. 184, pp. 175–187, Oct. 2016.

[87] K. Wang, J. Wang, P. Wang, M. Sparrow, J. Yang, and H. Chen, "Influences of urbanization on surface characteristics as derived from the moderate-resolution imaging spectroradiometer: A case study for the Beijing metropolitan area," *J. Geophys. Res.-Atmos.*, vol. 112, no. D22, Aug. 2007, Art. no. D22s06.

[88] T. Shi, Y. Yang, J. Ma, L. Zhang, and S. Luo, "Spatiotemporal characteristics of representative urban heat island effect in Anhui province based on MODIS," *J. Appl. Meteorol. Sci.*, vol. 24, no. 04, pp. 484–494, 2013.

[89] S. Grossman-Clarke, J. A. Zehnder, T. Loridan, and C. S. B. Grimmond, "Contribution of land use changes to near-surface air temperatures during recent summer extreme heat events in the phoenix metropolitan area," *J. Appl. Meteorol. Climatol.*, vol. 49, no. 8, pp. 1649–1664, Aug. 2010.

[90] X. Liu, F. Hu, L. Li, and Y. Wang, "Analysis and research on urban climate trends and environmental effects in Beijing area in summer," *Chin. J. Geophys.*, vol. 3, pp. 689–697, 2006.

[91] J. Zhang, Y. Hou, G. Li, H. Yan, and L. Yang, "Satellite remote sensing research and influence factor analysis of diurnal variation and seasonal characteristics of heat island in Beijing and its surrounding areas," *Sci. China Earth Sci.*, vol. 35, no. S1, pp. 187–194, 2005.

[92] Y. Liao et al., "Surface urban heat island detected by all-weather satellite land surface temperature," *Sci. Total Environ.*, vol. 811, pp. 151405–151405, Mar. 2022.

[93] P. Romanov, "Urban influence on cloud cover estimated from satellite data," *Atmos. Environ.*, vol. 33, no. 24–25, pp. 4163–4172, Oct. 1999.

[94] W. M. Angevine, A. B. White, C. J. Senff, M. Trainer, R. M. Banta, and M. A. Ayoub, "Urban-rural contrasts in mixing height and cloudiness over Nashville in 1999," *J. Geophys. Res.-Atmos.*, vol. 108, no. D3, Feb. 2003, Art. no. 4092.

[95] C. J. G. Morris, I. Simmonds, and N. Plummer, "Quantification of the influences of wind and cloud on the nocturnal urban heat island of a large city," *J. Appl. Meteorol.*, vol. 40, no. 2, pp. 169–182, Feb. 2001.

[96] J. Fallmann, P. Suppan, and S. Emeis, "Modeling of the urban heat island (UHI) using WRF—Assessment of adaptation and mitigation strategies for the city of Stuttgart," EGU General Assembly, 2013.

[97] Z.-H. Wang, E. Bou-Zeid, S. K. Au, and J. A. Smith, "Analyzing the sensitivity of WRF's single-layer urban canopy model to parameter uncertainty using advanced Monte Carlo simulation," *J. Appl. Meteorol. Climatol.*, vol. 50, no. 9, pp. 1795–1814, Sep. 2011.

Fengjiao Zhang is currently working toward the master's degree in geodesy and survey engineering from the College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China.

Her research interests include satellite remote sensing radiometric calibration, land surface processes and urban climate, and air quality.

Xuepeng Zhang received the master's degree in geodesy and surveying engineering from the China University of Mining & Technology, Beijing, China, in 2022.

He is currently a Research Associate with the Research Center of Big Data Technology, Nanhu Laboratory, Jiaxing, China. His research interests include valuation of ecosystem service values based on multisource remote sensing data and cloud removal of remote sensing satellite images.

Wei Chen received the B.S. degree in geographical information system from the Beijing Normal University, Beijing, China, in 2008, and the Ph.D. degree in photogrammetry and remote sensing from the Peking University, Beijing, in 2013.

He is currently an Associate Professor with the College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing, China. His research interests include remote sensing retrieval and remote sensing applications.
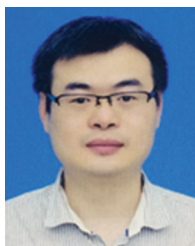
Bin Yang received the B.S. degree in science and technology of remote sensing from the Beihang University, Beijing, China, in 2012, and the Ph.D. degree in photogrammetry and remote sensing from the Peking University, Beijing, in 2017.

He is currently an Associate Professor with the College of Electrical and Information Engineering, Hunan University, Changsha, China, and with the Key Laboratory of Visual Perception and Artificial Intelligence of Hunan Province, Changsha, China. His research interests include remote sensing and artificial intelligence.

Zhenting Chen received the B.S. degree in applied mathematics from the Kunming University of Science and Technology, Kunming, China, in 2003, the M.S. degree in GIS from the Kunming University of Science and Technology, Kunming, China, in 2012, and the Ph.D. degree in optics from the University of Science and Technology of China, Hefei, China, in 2018.

From 2008 to 2014, he mainly engaged in the research of geographic information system. From 2015 to 2018, he participated in the development of DPC sensor of GF-5 satellite with the Hefei Institute of Material Science, Chinese Academy of Sciences, Hefei, China. His research interests include quantitative remote sensing, satellite technology, big data, AI, etc.

Hongzhao Tang received the B.S. degree in surveying and mapping engineering from the Wuhan University, in 2007, Wuhan, China, the M.S. degree in photogrammetry and remote sensing from the Peking University, Beijing, China, in 2010, the M.Sc. degree in ecology from the Colorado State University, Fort Collins, CO, USA, in 2013, and the Ph.D. degree in geographic information system from the Peking University, in 2022.

He is currently an Associate Research Fellow with the Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, China. His research interests include radiometric calibration and validation of satellite sensor and hyperspectral remote sensing application.

Zhe Wang is currently working toward the master's degree in geodesy and survey engineering with the College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing, China.

His research interests include remote sensing of the ecological environment, computer vision, and deep learning.

Pengshuai Bi is currently working toward the master's degree in surveying and mapping engineering with the College of Geoscience and Surveying Engineering, China University of Mining and Technology, Beijing, China.

His research interests include polarization remote sensing, image dehazing, and application of deep learning in remote sensing.

Lan Yang is currently working toward the master's degree with the College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing, China.

Her research interests include polarization remote sensing, extraction of vegetables under shadow conditions, and application of deep learning in remote sensing.

Guangchao Li is currently working toward the Ph.D. degree in photogrammetry and remote sensing with the College of Geoscience and Surveying Engineering, China University of Mining & Technology, Beijing, China.

His research interests include atmospheric aerosols, and ecological environment monitoring and evaluation.

Zhe Jia received the B.S. degree in environmental science from Northwest Normal University, Lanzhou, China, in 2002 and the M.S. degree in environmental engineering from the Xi'an University of Science and Technology, Xi'an, China.

He is currently working with the College of Geomatics, Xi'an University of Science and Technology.

His research interests include mine geological ecological environment protection, environmental damage, and remediation.