IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 17, 2019, accepted January 24, 2019, date of publication January 30, 2019, date of current version March 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2896241

Downscaling Land Surface Temperatures Using a Random Forest Regression Model With Multitype Predictor Variables

HUA WU[1,2,3] AND WAN LI[1,2]

[1]State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

Corresponding author: Hua Wu (wuhua@igsnrr.ac.cn)

This work was supported in part by the Chinese Academy of Sciences through the Strategic Priority Research Program under Grant XDA20030302, in part by the National Key R&D Program of China under Grant 2018YFB0504800, and in part by the National Natural Science Foundation of China under Grant 41871267 and Grant 41571352.

ABSTRACT In this paper, a random forest regression model with multitype predictor variables (MTVRF) was utilized with four kinds of input variables, including surface reflectance, spectral indices, terrain factors, and land cover types, to establish the nonlinear relationship between land surface temperature (LSTs) and other land surface parameters. The main objective of this paper is to analyze the superiority of MTVRF model in multivariable regression wherever on the simple or complex underlying surface and further to demonstrate the robustness of random forest (RF) regression downscaling model trained in one study area while being applied to another area. The spatial resolution of the Moderate Resolution Imaging Spectroradiometer LST product was downscaled by MTVRF from 990 to 90 m. A comparison with two other downscaling methods, such as the basic RF model and the thermal sharpening algorithm, was also made. By computing the mean error, the determination coefficient (R^2), and the root mean square error (RMSE) between the downscaled and referenced LSTs, the MTVRF model achieved a satisfied performance. Further satisfactory results were also obtained for the MTVRF to downscale LSTs for different land covers and evaluate the training model in various regions. The RMSE of the MTVRF model trained on study area B and evaluated on study area A was 3.13k, while the RMSE trained on study area A and evaluated on study area B was 2.11k; this shows the MTVRF model trained in a specific region is thought to be robust enough to downscale LSTs under other various surface conditions.

INDEX TERMS Land surface temperature, downscaling, random forest, thermal remote sensing, thermal sharpening, robustness.

I. INTRODUCTION

Land surface temperature (LST), which is described as one of the most important parameters for atmospheric and land surface interactions, material cycles and energy exchanges in the terrestrial ecosystem at regional and global scales, plays an essential role in the energy balance of land surfaces and the solution of biophysical parameters [1]–[6]. LSTs with a finer resolution have been widely used in hydrological equilibrium assessments, global warming studies, urban heat island effect assessments and surface evapotranspiration calculations [7]–[12].

The associate editor coordinating the review of this manuscript and approving it for publication was Tomasz Trzcinski.

Thermal infrared remote sensing data, which usually derive from thermal infrared sensors such as the Thermal Infrared Sensor (TIRS), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Moderate Resolution Imaging Spectroradiometer (MODIS), have been continuously used to retrieve land surface temperatures. However, remotely sensed LST products yield a tradeoff between high temporal and high spatial resolutions [13]. For example, the resolution of Landsat ETM plus at band 6 is 60 m but with a 16-day cycle, which cannot meet continuous time series observations for the same area. High temporal resolution sensors, such as MODIS, obtain observations twice every day, but only 1-km LSTs can be provided. Thermal infrared images with low spatial resolutions are affected by spatial heterogeneity due to the subcomponents within each pixel. This issue impedes the application of LSTs. Therefore, an effective downscaling method is desired to improve the resolution of LSTs [14]–[16].

During the past few years, various downscaling algorithms have been proposed to downscale LSTs from coarser to finer resolutions [13], [17]–[25], where downscaling methods based on scale factors are recognized as one kind of popular algorithm [15]–[16], [26]. Downscaling methods based on scale factors are commonly divided into three categories: statistical regression-based downscaling methods [13], [19], [27], [28], modulation distribution-based downscaling methods and linear spectral mixture model-based downscaling methods [29]–[35]. Statistical regression methods, such as the disaggregation procedure for radiometric surface temperature (DisTrad) method [13] and the thermal sharpening (TsHARP) algorithm [19], have been commonly accepted as easily manipulated and satisfactorily accurate among these methods. Early regression algorithms tend to focus on the statistical relationship between vegetation indices and LST, such as the normalized difference vegetation index (NDVI) [13], fractional vegetation cover [17]–[19], and soil-adjusted vegetation index (SAVI) [24], based on the belief that the spatial variance of land surface temperature is mostly controlled by vegetation coverage [36], [37]. However, the main limitation of these algorithms is that the correlation between a vegetation index and LST appears to be insufficient for regions with impervious surfaces, barren lands and wetlands [29], [38], [39]. Thus, many other spectral indices have been used to fit the nonlinear relationships between LST and land parameters, including the normalized difference building index (NDBI), modified normalized water index (MNDWI) [28], and normalized difference drought index (NDDI) [22]. Other scale factors, such as surface albedo, digital evaluation models and land cover types, have continuously been introduced to improve the statistical regression algorithm [27], [41].

In addition to this method, recent studies also addressed the issues of linear or nonlinear regression algorithms between LST and the mentioned parameters [42]–[45], for example, Duan and Li proposed an original geographically weighted regression downscaling model and achieved good results [45]. Machine learning algorithms, such as the Bayesian-based model [42], artificial neural network (ANN) [46]–[48], support vector machine (SVM) [49], combined global window and moving window regression trees and random forest (RF) [25], [28], [50], [51], have obtained great accuracies when fitting the nonlinear relationship between LST and other variables.

Recently, the random forest model has been utilized in the LST downscaling of vegetation and arid areas [28], [39]. In addition, a great contribution to introduce land surface reflectance, terrain factors (DEM, slope) and land use map into RF downscaling model and achieved good results in simulative Landsat LST as well as MODIS LST product [51]. However, something detailed still worthy to be discussed such as the application of RF downscaling model in different kind of regions, as well, the selection of predictors should be considered more comprehensively and systematically. Therefore, the main objective of this study is to verify that the random forest regression model with multitype predictor variables (MTVRF) has a satisfied performance on LST downscaling over various underlying surface and the robustness of the MTVRF model is also comprehensively evaluated. In this study, the MTVRF model is first applied to downscale the resampled LST product of MODIS (990 m) to 90 m by using four kinds of predictor variables, including surface reflectance, spectral indices, terrain factors and a land classification map; then, the results are evaluated by the LST product of ASTER (90 m). Two distinguished areas with different land cover and terrain factors are selected to further compare the downscaling performance of the MTVRF model. The rest of this paper is organized as follows: Section 2 presents the study areas, data and methodology. Section 3 gives a detailed analysis of the downscaling results. Section 4 derives further discussion, and Section 5 draws a conclusion.

II. MATERIALS AND METHODS

A. STUDY AREA

Two typical areas were selected in this study with different terrains and land covers to fully understand the underlying surface characteristics in these two areas. Figure 1 shows the false color image of these two areas derived from Landsat-8 OLI reflectance data.

Study area A comprises the Peñarora mountain region, which is located in Segovia, Spain. The latitude and longitude of this area are 40°N to 41°N and 4°W to 3°W, respectively. The variation in the elevation range is from 687 m to 2414 m. More than 80% of this area is covered by vegetation, which is mainly composed of forests and croplands. The climate is Mediterranean subtropical, with temperatures ranging from −17.0°C to 39.7°C. The annual mean temperature of this area is 11.5°C, and the annual precipitation is approximately 464 mm.

Study area B is located northwest of Beijing, China and includes the districts of Haidian and Changping. The coordinate range for this area is from 39°N to 40°N and 115°E to 116°E. Most of this area has a flat terrain, except for the western mountainous districts. The climate is typically a semihumid continental monsoon climate, with an annual mean temperature ranging from 10°C to 12°C and mean precipitation ranging from 450 mm to 550 mm. This study area contains four kinds of land cover: vegetation, croplands, impervious surfaces (including buildings and roads) and water; the area is mainly characterized by vegetation in the western regions and impervious surfaces in the eastern regions.

For a more detailed discussion on LST downscaling, three subareas were also selected: subarea 1 in study area A, which is the Santillana reservoir and is marked as a water region with a maximum surface cover of 1052 ha; subarea 2 in study area B, which is dominated by forests and is regarded as a

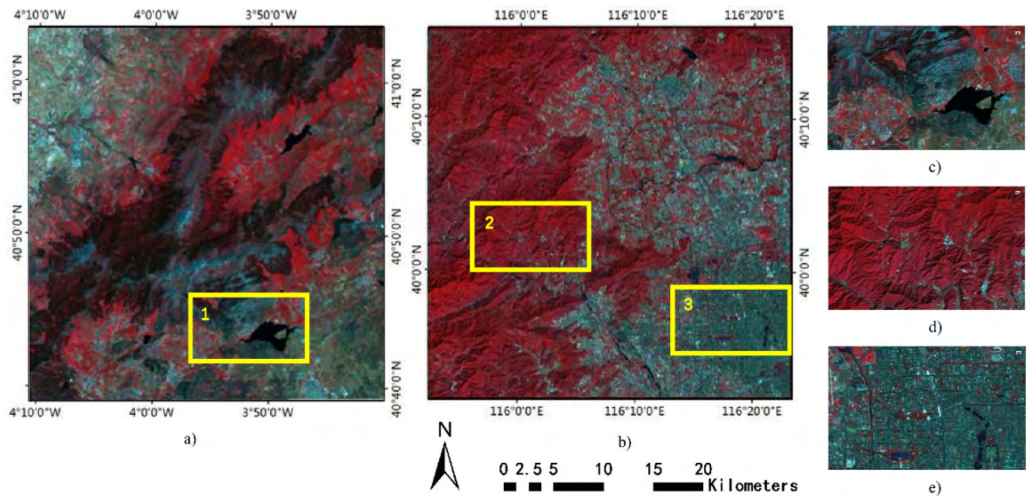

FIGURE 1. Geolocation of the study area with false color images generated from Landsat 8 data (R: band 5; G: band 4; B: band 3). (a) Study area A in Segovia, Spain, (b) study area B in Beijing, China, (c) subarea 1 in the water region, (d) subarea 2 in the vegetation region, and (e) subarea 3 in the impervious surface region.

vegetation region; and subarea 3 in study area B (i.e., Xicheng District), which has high-rise buildings and crisscross roads and can be described as an impervious surface region.

B. DATA PREPARATION

1) LANDSAT DATA

The Landsat 8 Operational Lad Imager (OLI) and TIRS image were acquired at the USGS Earth Resources Observation and Science (EROS) Center Science Processing Architecture and were retrieved with USGS Earth Explorer (http://earthexplorer.usgs.gov/) with the resolution of 30 m and 100 m, respectively. In this study, two Landsat tiles (WRS-2 Path 201 / Row 32 and Path 123 / Row 32) are applied to cover the selected study areas in Segovia and Beijing. The Landsat land surface reflectance products with bands 2-7 and the processed spectral indices are served as one of the input variables for downscaling MODIS LST product from 990 m to 90 m.

2) ASTER DATA

The LST at finer resolution was derived from the ASTER, which is a sensor aboard the Terra satellite launched on December 18, 1999. The ASTER LST products (AST08), available from the NASA Earthdata Search (https://search.earthdata.nasa.gov/) with a spatial resolution of 90 m, were generated from the Temperature/emissivity Separation (TES) algorithm with the accuracy of about 1.5 K [52]. The ASTER LST will be served as the reference data to validate the performance of the MTVRF model at the finer scale.

3) MODIS DATA

The MODIS products, which were acquired on 22 June 2016 for study area A and 24 July 2014 for study area B, derive from another sensor aboard the Terra satellite. Because the same satellite platform is used for both the ASTER and MODIS sensors, discrepancies caused by geometric observation deviations in the LST products tend to be neglected. The concurrent data also eliminate observed time differences between the ASTER and MODIS LST products. Although the LSTs of MODIS can be generated by a generalized split-window (GSW) algorithm [53] and TES algorithm [54], we have no alternative but to use GSW-based sensors LSTs because of the unavailability of TES-based LSTs due to science data quality issues. The collection 6 MOD11A1 LSTs are tile-based global products that provide per-pixel temperature and emissivity values, with a resolution of 1 km and an accuracy of approximately 1 K [55], [56]. These GSW-based LST products were registered into WGS 84/UTM Zone 30 N for study area A and WGS 84/UTM Zone 50 N for study area B, with a resampling interval of 990 m. It is noteworthy that there is a deviation between ASTER TES-based LSTs and MODIS GSW-based LSTs because different algorithms are used to generate the LST products. Fortunately, a simple linear regression has been used to remove any systematic LST discrepancies between different sensors [57]. This linear regression model was first established for ASTER and MODIS LST products at coarser resolutions. Subsequently, the MODIS LST products at coarser resolution were converted using this simple linear regression and then were applied to MTVRF model training instead of original MODIS LST products. Figure 2 shows a scatter plot and linear regression relationship between ASTER and MODIS LSTs at a 990-m resolution. The determination coefficient (R^2) is 0.93 in study area A and 0.90 in study area B within the 95% confidence interval, which shows a significant correlation between the ASTER and MODIS LST products. The RMSE is 1.82 K in study area A and 1.07 K in study area B. There are about 60-70% of points whose RMSEs range from -1 to 1 K, which indicates that this

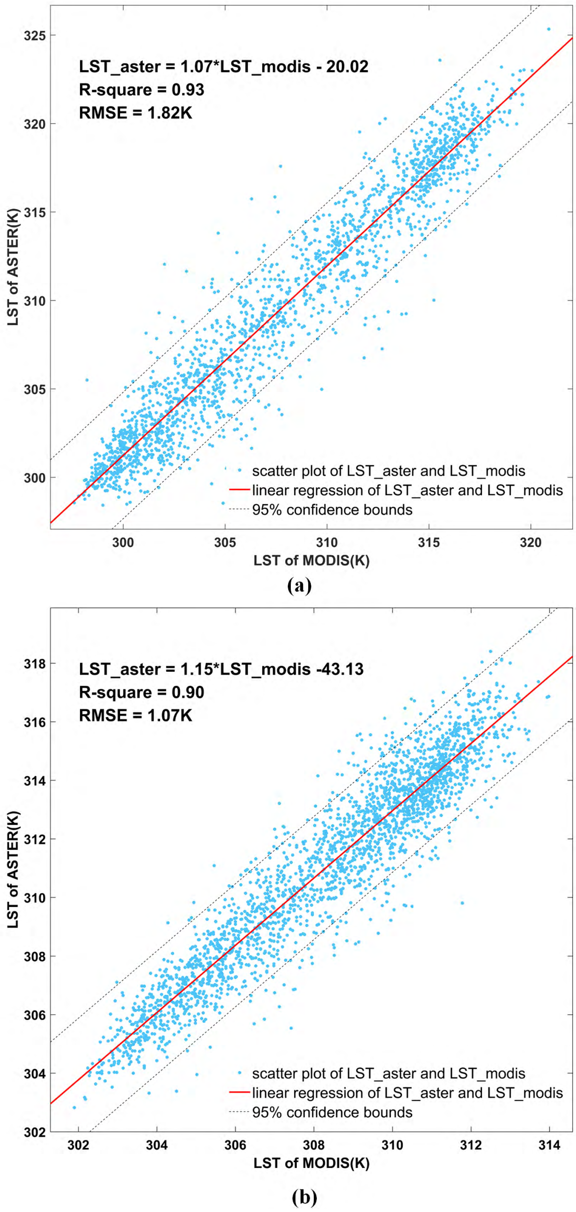

FIGURE 2. **The scatter plot and linear regression relationship between LST products from MODIS and ASTER in study areas A and B.**

conversion may be suitable to evaluate the accuracies of the downscaled LSTs.

4) SRTM DATA

The digital elevation model (DEM) data are derived from The NASA's Shuttle Radar Topography Mission (SRTM), which is a joint effort between NASA, the National Bureau of Defense (NIMA), Germany and the Italian space agency. The SRTM DEM data, with a spatial resolution of 90 m, were also registered to WGS 84/UTM Zone 30 N and WGS 84/UTM Zone 50 N. Subsequently, the aspect, slope and hill-shade were calculated by the spatial analysis module of ARCGIS 10.2. Those terrain factors were spatially aggregated to a resolution of 990 m by spatial averaging to the resolution of the MODIS LST for the RF model training.

C. METHODOLOGY

1) MTVRF

The RF is an integrated machine learning algorithm that evolved from the bagging algorithm [58]–[60]. As a nonlinear statistical ensemble regression method, the RF is constructed by a set of uncorrelated classification and decision regression trees. Every bootstrap sample is selected from the training set, and the features used are extracted randomly from all features in a certain proportion of the set while training each mode of the tree [61]. The results of RF training turn out to be the voting output for all decision trees. The RF promises to obtain almost all desired results with high adaptability to the data and the parameters used, and there is no need to adjust the parameters tediously, as is done in the SVM. The RF is regarded as one of the best machine learning algorithms today and has been widely used in various fields involving remote sensing image processing, such as remote sensing image classification, feature recognition and spatial downscaling [62]–[65].

The predictor variable datasets in the RF downscaling algorithm are intended to reflect the spatial variation in LSTs efficiently over different regions. In TsHARP and the basic RF model, only minority variables of the highest correlation with LST like vegetation coverage are qualified to be introduced into the model. Nevertheless, these variables may have weak performances due to the complexity of land cover and further impede the effects while being tested on other regions. In this paper, MTVRF will make a trade-off between algorithm complexity and the number of input variables on the premise of the insensitivity to multicollinearity, which pledges the robustness of the result for missing and nonequilibrium data and has a satisfactory prediction for thousands of inputs. In addition, MTVRF also could avoid over-fitting, which improves the generalization of the model, thus it could depict the complicated surface status and applied to other regions with acceptable results.

According to the statistical correlation between LST and biophysical parameters, four main kinds of predictor variables were introduced into the MTVRF downscaling model to improve the generalization ability. The input variables are listed as follows:

a). Surface reflectance of visible, near infrared and shortwave infrared bands, which contains plenty of vegetation cover and soil moisture conditions.

b). Typical spectral indices, which may be sensitive to specific land cover types. For example, the NDVI, vegetation fraction, and SAVI satisfactorily indicate the vegetation density and biomasses, while the NDBI is trained to recognize impervious surface more precisely, other spectral indices like NDDI, NMDI and MNDWI are used to fit the relationship between land surface temperatures and soil moisture which has a great impact on LST variation.

c). Terrain factors, including the DEM and its derivatives, such as aspect, slope and hill-shade of the study area, are assumed to have a significant correlation with LST in mountainous areas.

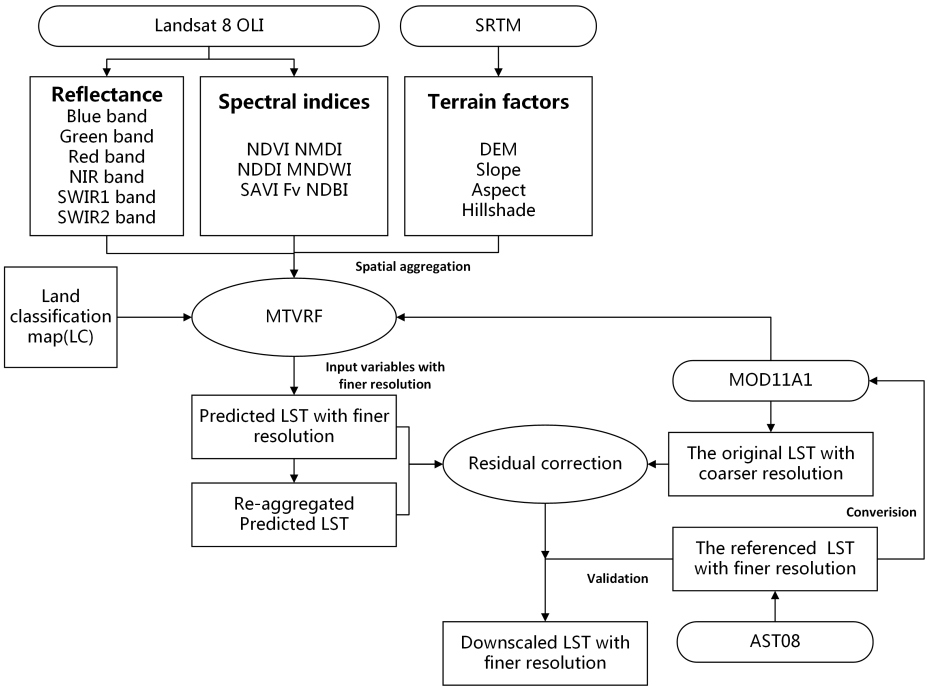

FIGURE 3. Schematic of the MTVRF land surface temperature downscaling procedure.

d). Land classification map, which is regarded as the predictor to promote the recognition of the land cover's influence on LSTs in different regions.

2) STEPS OF THE DOWNSCALING MODEL

The specific steps for the developed MTVRF LST downscaling model are shown in Figure 3 and summarized as shown below:

(1) The input variables should be aggregated to match the coarse resolution of the product to be downscaled. In this study, image data form Landsat 8 OLI and SRTM, with resolutions of approximately 30 m and 90 m, respectively, were first aggregated to 990 m (i.e., the same as that of the LST data from MODIS) using a spatial averaging method, where the statistical relationship between the explanatory variables and LST can be established at a coarse level.

(2) At the 990-m level, the MTVRF regression model between LST and explanatory variables can be expressed as:

$$LST_c = F\left((\rho_i)_c, (S_i)_c, (TF_i)_c, (LC)_c\right) \qquad (1)$$

where the subscript c represents the variable with a coarser resolution, subscript i represent the i-th variable, ρ represents the reflectance, S represents the spectral index, TF represents the terrain factor and LC represents the land cover type. The function F $(\cdot)$ indicates a nonlinear relationship between converted MODIS LST and these variables.

(3) It is worth noting that the residual temperature calculated from the difference between the original LST with a coarser resolution and the simulated LST (LST_c) is intended to correct the prediction ME, as the RF regression cannot explain all spatiotemporal variations in the LST distribution [51].

This is called the model error:

$$\Delta LST_c = LST_o - LST_c \qquad (2)$$

where the subscript o indicates the original MODIS LST converted by ASTER.

(4) The trained model is subsequently used on a finer scale (90 m) given the scale invariance in the relationship between LST and other variables. The model error after resampling was allocated for each pixel at a finer resolution. Thus, the final downscaled LST at a 90-m resolution can be worked out as follows:

$$LST_f = F\left((\rho_i)_f, (S_i)_f, (TF_i)_f, (LC)_f\right) + \Delta LST_c \qquad (3)$$

where the subscript f refers to the variable with a finer resolution.

One of the most important parameters of RF model is out-of-bag error, of which out-of-bag (OOB) samples means about one-third of samples that were not participated in model training. The OOB samples in each tree served as a testing dataset to ensure an unbiased estimation of error. Therefore, there is no need to execute cross-validation or use a single testing dataset to obtain the unbiased estimation of error, since the unbiased estimation could be established during the generation process of model. Averagely, the size of training dataset is about 10000 of each study areas.

The key parameters of RF model which brought out the best regression result, the number of decision trees (n_estimators) and the maximum number of features to be split (max_features), are traversed according to OOB error

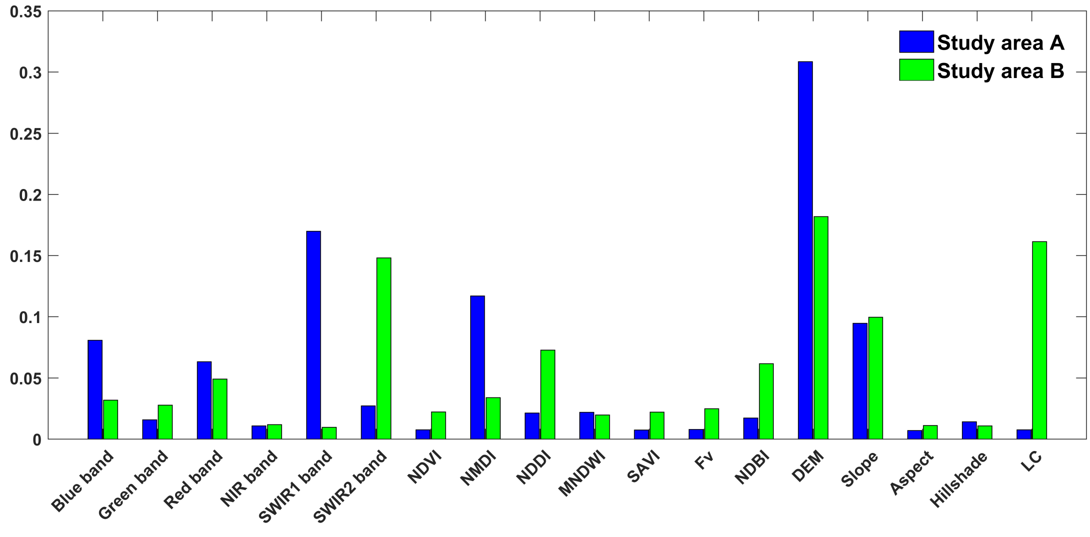

FIGURE 4. Random forest variable importance scores averaged across two study areas.

estimates, and the best results turn out that n estimators is 200 and max_features is 9.

The performance of the MTVRF downscaling model is compared with (1) the basic RF model proposed by Hutengs and Vohland [51], where the reflectance values of the NIR and red bands, DEM and land use map are selected as input variables, and (2) the TsHARP method, which is based on the linear relationship between LST and NDVI.

D. ERROR EVALUATION

The downscaling results were evaluated by R^2, ME and RMSE, which are commonly used as measurable indicators for fitting problems. The expressions are given as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (LST_\mathrm{e} - LST_r)^2}{\sum_{i=1}^{n} \left(LST_e - \overline{LST_r}\right)^2} \tag{4}$$

$$ME = \frac{\sum_{i=1}^{\mathrm{n}} (LST_e - LST_r)}{n} \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (LST_e - LST_r)^2} \tag{6}$$

where LST_e refers to the estimated LST from MTVRF, LST_r represents the reference ASTER LST, and n represents the total number of samples involved in the estimation.

III. RESULTS AND ANALYSIS

A. IMPORTANCE OF THE SELECTED VARIABLES

Figure 4 shows the importance ranking of the input variables calculated by the MTVRF in two study areas. It is worth mentioning that this importance score gives a relative ranking regarding the contribution of the input variables, but it is not equivalent to the correlation coefficient. The contributions of the terrain factors, especially for the DEM and slope, are shown to be greater than those of any other variable in both areas. This implies that there are large topographic effects on solar incident radiation and longwave surface cooling along mountainous surfaces. In addition, study area A, which is terrain-dominated, shows a significant inconsistency between the DEM and other kind of variables, while the importance scores in study area B show homogeneous variations because of the flat landscape.

TABLE 1. Downscaling statistics for the MTVRF model, basic RF model and TsHARP method.

Method	Study area A			Study area B		
	RMSE	R^2	ME	RMSE	R^2	ME
Basic RF	3.32	0.66	-0.10	2.14	0.63	-0.25
TsHARP	3.62	0.69	-0.12	2.16	0.69	0
MTVRF	2.67	0.85	-0.05	2.10	0.71	-0.03

For study area A, which is largely covered by hills with elevations ranging from ~616-2409 m, the spatial distribution of LSTs is almost entirely controlled by these terrain factors and vegetation fraction correlated factors, such as the red band and NMDI. For study area B, the scene acquired over the downtown area, with a complicated landscape and relatively flat terrain, has a DEM that becomes less important than that in study area A and a classification map that has a higher weight. The NDBI, which is regarded as the indicator of buildings and roads in cities, also shows higher importance.

Because these importance scores vary with the number of input variables, which means that these importance scores would change when input variables are added or removed, they may give evidence for the selection of input variables with a higher correlation. For further analysis, quantified analyses should be performed to evaluate the downscaling effect of the MTVRF.

B. DOWNSCALING PERFORMANCE UNDER DIFFERENT TERRAIN CONDITIONS

The LST downscaling performance of the RF model from 990m to 90m is shown in Table 1, all three models improve

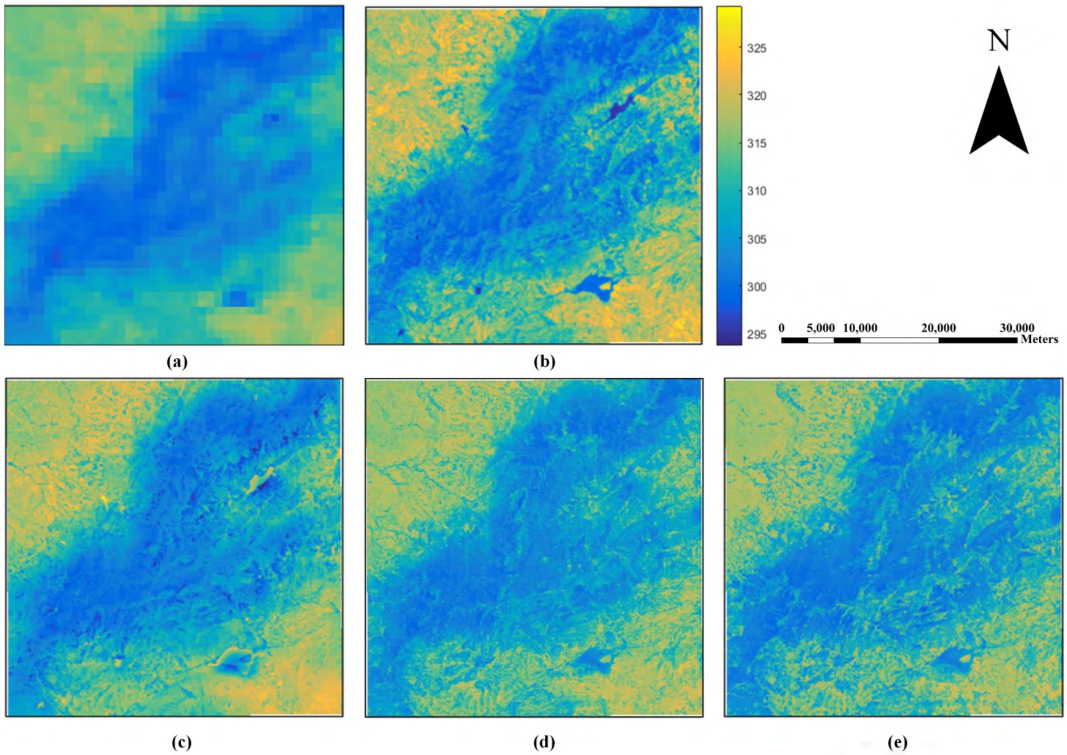

FIGURE 5. Spatial distribution of LSTs for study area A. (a) 990-m MODIS LST, (b) 90-m ASTER LST, (c) 90-m downscaled TsHARP LST, (d) 90-m downscaled basic RF LST, and (e) 90-m downscaled MTVRF LST.

the resolution of LSTs from 990 m to 90 m with satisfactory accuracies. On average, the best downscaling results are acquired from the MTVRF model, followed by the basic RF model and the TsHARP method.

For study area A, as shown in Figure 5, the ME and RMSE of the downscaled LSTs for the MTVRF model are -0.05 and 2.67 K, respectively. The accuracy is improved by approximately 20% compared to the basic RF model, with a RMSE of 3.32 K, and by approximately 26% compared to the TsHARP method, with a RMSE of 3.62 K. The satisfactory results of the MTVRF model may account for the introduction of multiple spectral indices and terrain factors, especially for the DEM, which is related to the large topographic influence on land surface patterns. As shown in Figure 5, the detailed texture described by the MTVRF model (Figure 5e) cannot be found in the same spot by using the basic RF model (Figure 5d) and TsHARP method (Figure 5c) because of the impact of terrain factors. Furthermore, the TsHARP method, which is reported as performing well on the full cover of natural vegetation, has a better downscaling result than the basic RF model in the study area.

For study area B, as shown in Figure 6, the best downscaling results are also obtained by the MTVRF model, with a ME and RMSE of -0.03 and 2.10 K, respectively. The accuracy is improved by approximately 2% compared to the basic RF model, with a RMSE of 2.14 K, and by approximately 3% compared to the TsHARP model, with a RMSE of 2.16 K. With the introduction of NDBI, which identify impervious surfaces well, the spatial pattern in the MTVRF model is more robust and detailed. However, the complexity of the land cover types may result in a problem of mixed pixels, and flat plains may neglect the importance score for terrain factors, especially for DEM.

As shown in Figures 7 and 8, most errors in these three methods are distributed from -5 to 5 K. The downscaling accuracy for LSTs is highest when using the MTVRF model, followed by the TsHARP method and basic RF model. Although there are some underestimations (overestimations) in the maximum (minimum) LST because of the deviation between the downscaled LST and referenced ASTER LST, we can still draw the conclusion that the MTVRF model outperforms the TsHARP and basic RF models when successfully downscaling LSTs under different terrain conditions.

C. DOWNSCALING PERFORMANCE FOR DIFFERENT LAND COVER TYPES

To fully evaluate the downscaling performances for different land cover types, three kinds of subareas (i.e., vegetation, water and imperious surfaces) are extracted from study areas A and B.

As shown in Table 2, the downscaled LSTs in the vegetation regions show the most satisfactory results, with RMSEs of 2.80 K, 1.91 K and 1.92 K for the three downscaling models, followed by the impervious surfaces and water regions.

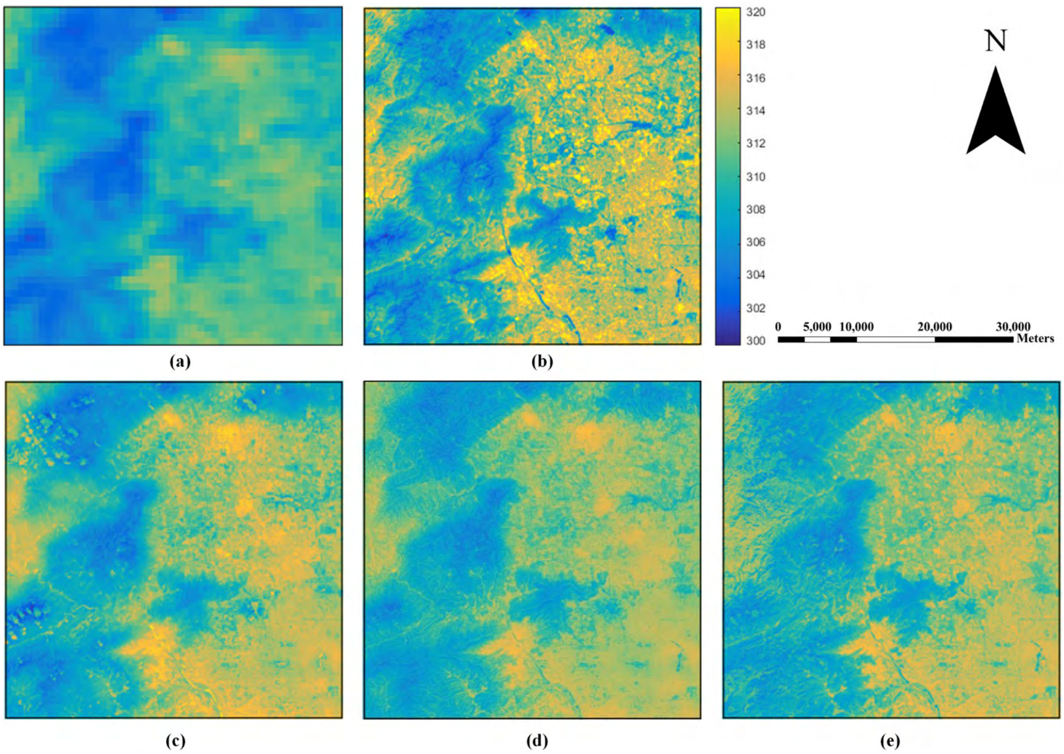

FIGURE 6. Spatial distribution of LSTs for study area B. (a) 990-m MODIS LST, (b) 90-m ASTER LST, (c) 90-m downscaled TsHARP LST, (d) 90-m downscaled basic RF LST, and (e) 90-m downscaled MTVRF LST.

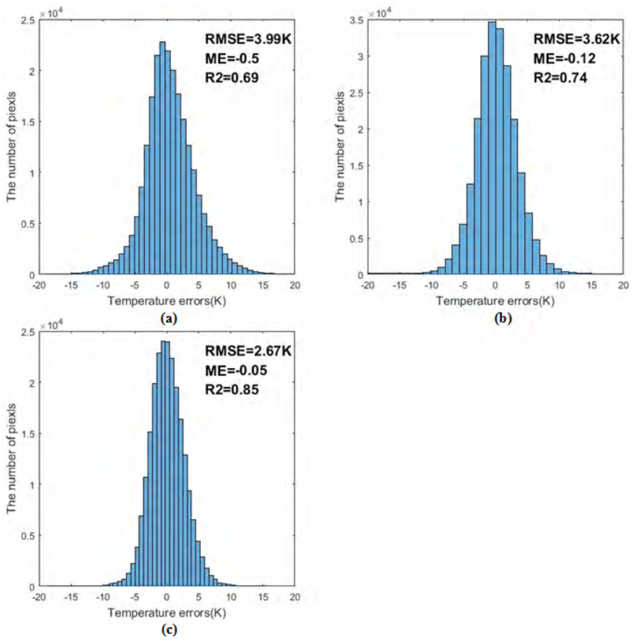

FIGURE 7. Distribution of the LST errors between the estimated and reference LSTs for study area A. (a) 90-m downscaled basic RF LST, (b) 90-m downscaled TsHARP LST, and (c) 90-m downscaled MTVRF LST.

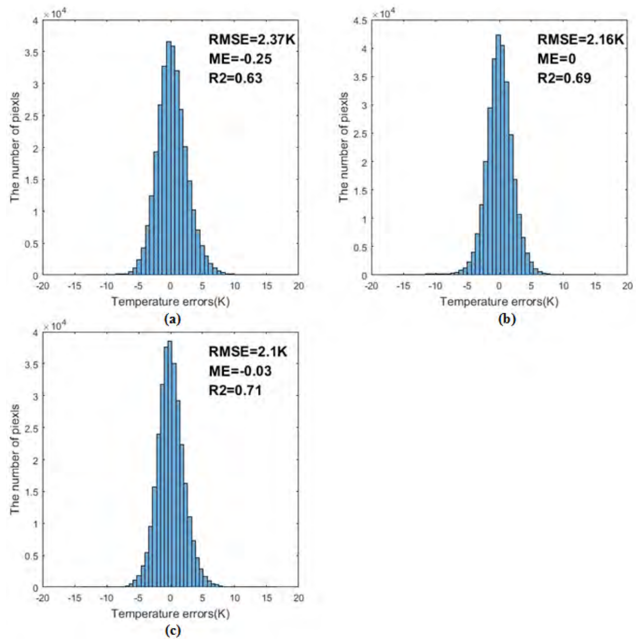

FIGURE 8. Distribution of the LST errors between the estimated and reference LSTs for study area B. (a) 90-m downscaled basic RF LST, (b) 90-m downscaled TsHARP LST, and (c) 90-m downscaled MTVRF LST.

In Figure 9, large improvements in the MTVRF model over the water region can be seen from the spatial distribution of LST errors in Figures 9a-9c, where obvious underestimations of the water surface and overestimations around the water region appeared when using the TsHARP method, which can be attributed to the insensitivity of water to the NDVI or other vegetation factors. In addition, the accuracy of the basic RF model is also inferior to the MTVRF model

TABLE 2. Downscaling statistics for vegetation, water and impervious surfaces using three downscaling methods.

	Water		Vegetation		Impervious Surface	
	RMSE	ME	RMSE	ME	RMSE	ME
Basic RF	4.92	-1.10	2.21	-0.33	2.73	-0.53
TsHARP	8.53	0.20	3.11	-1.81	2.72	1.38
MTVRF	2.80	-0.01	1.91	0.4	1.92	0.29

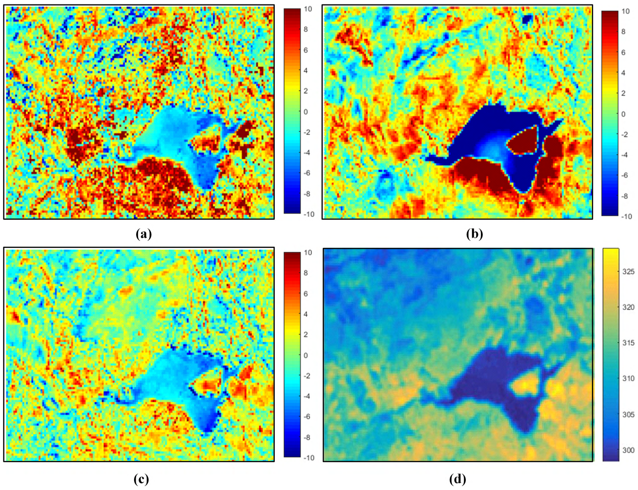

FIGURE 9. Spatial distributions of LST differences over water regions by using the (a) basic RF model, (b) TsHARP model, (c) MTVRF model and (d) the referenced ASTER LST.

due to the simple depiction of the surrounding environment. As shown in Figures 5 and 6, the MEs are close to zero when the regression models trained for the total area are used to predict the same location. However, MEs may appear in subareas with single land cover types because of the different underlying attributes. The MEs in Table 2 show that larger MEs appear over the water region when using the TsHARP method and basic RF model, while the ME is close to zero when using the MTVRF model, which proves its adequate robustness over water regions. Additionally, the MTVRF model retains the homogeneity of water surface temperatures because of the MNDWI, which plays a crucial role in water recognition.

From Figure 10, which shows the predictions over the vegetation region, good accuracies remain, and the ME and RMSE of the three models do not change too much because of the general similarity of the underlying situation between the total study area and subarea covered with vegetation. In addition, the loss of spatial details for the MODIS LST at a coarser resolution are restored clearly by the MTVRF model, which can be seen from the spatial distribution of LST differences in Figure 10c, where more detailed LST variations derived from terrain factors are reserved. It can also be concluded that the TsHARP method performs better than the basic RF model, as vegetation is controlled.

In Figures 11a-11c, there are few differences in the spatial distribution of LST differences among the three models over

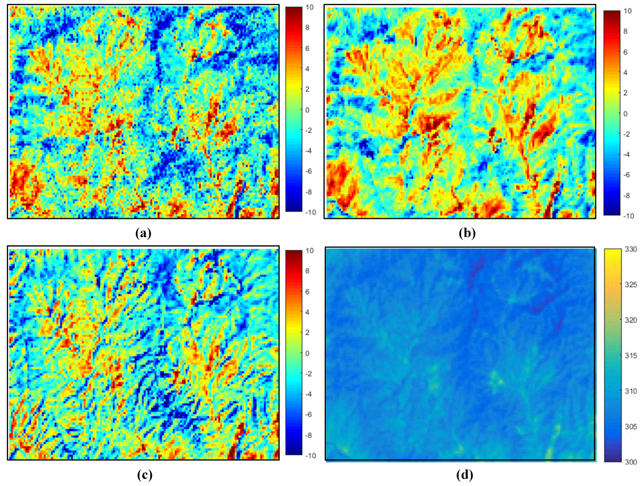

FIGURE 10. Spatial distributions of LST differences over vegetation region by using (a) basic RF model, (b) TsHARP model, (c) MTVRF model and (d) the referenced ASTER LST.

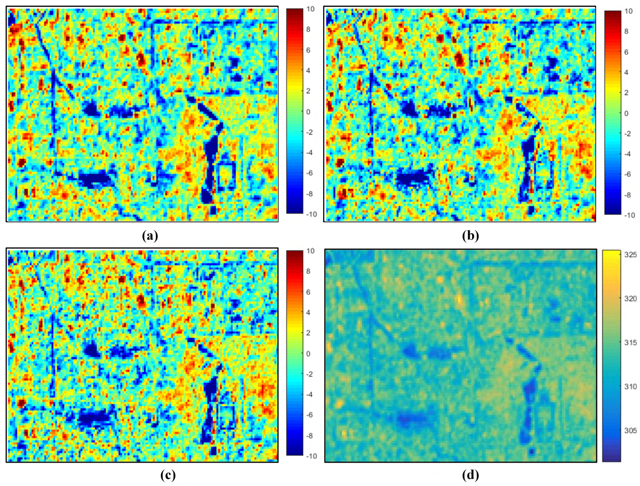

FIGURE 11. Spatial distributions of LST differences over impervious surfaces by using (a) basic RF model, (b) TsHARP model, (c) MTVRF model and (d) the referenced ASTER LST.

urban areas with complex land cover types and mixed pixel problems. The difficulty in constructing the stable relationship between LST and land parameters weakens the performance of the downscaling models. The RMSEs of the three models all range from 2 to 3 K. However, together with the results in Table 2, supplemental proof shows the robustness of the MTVRF model when being trained in impervious surfaces, with the lowest ME of approximately 0.29 K.

D. ROBUSTNESS EVALUATION FOR THE MTVRF MODEL

The RF downscaling model and other nonlinear regressions are trained separately for each region, which may lead to concerns regarding overfitting problems, as they are not generalized enough to downscale LSTs for other regions when the same area is utilized for training and prediction. For this reason, the MTVRF model trained over study area A

TABLE 3. Downscaling statistics for various regions.

	Study area A with model trained on study area B		Study area B with model trained on study area A	
	RMSE	ME	RMSE	ME
Basic RF	4.53	-0.36	3.07	0.42
TsHARP	6.68	0.26	3.54	-2.28
MTVRF	3.13	-0.30	2.11	-0.05

is intended to downscale the LSTs in study area B, and vice versa. The results from the training model tested over different regions are summarized in Table 3; compared with the results in Table 1, the RMSE increases by ~0.46 K for the downscaled LSTs in study area A, and the RMSE increases by ~0.01 K in study area B. The same experiment was tested on the other two downscaling models, and in general, the results were not better than those of the MTVRF model, which means that the basic RF model and TsHARP method are more regionally restricted. The major decrease in the prediction accuracies of the basic RF model and TsHARP method compared with that of the MTVRF can be attributed to the differences in underlying surface characteristics between the two areas, as a simple regression relationship, such as the LST-NDVI, cannot be directly relocated from one area (such as a downtown area) to another (such as a mountainous area).

By training separately and evaluating mutually in both study areas, the slight decrease in the prediction accuracies seems to be acceptable to support our conclusions, which further proves that the MTVRF model is robust enough to downscale LSTs in various regions with few overfitting problems. Therefore, the MTVRF model may lead to better results not only in a single region but also in various regions with satisfactory accuracies.

IV. DISCUSSION

The MTVRF model is shown to have a nonlinear relationship between LSTs and other land surface parameters. Downscaling results for study area A and study area B show evidence that the MTVRF model works better than the basic RF model and TsHARP algorithm when downscaling LSTs in any region with simple or complex land surfaces. The detailed experiments on vegetation, water and impervious surfaces give further proof that the MTVRF model yields the best downscaling results for these three kinds of regions with comprehensive input variables. Indeed, this comparison seems to be unfair between the MTVRF model with more variables and TsHARP as well as the basic RF model with less variables input, but it drives the promising method that with adequate input variables introduced, the downscaling model could keep region insensitivity and accuracy could also be ensured.

However, there are still some limitations in our research. First, the discrepancy between the ASTER and MODIS LST products has a negative effect on the evaluation of the downscaling results. The ASTER TES-based LST and MODIS GSW-based LST are generated from different algorithms, which may lead to concerns regarding the accuracy of the reference dataset to validate the downscaled results. The ideal solution to overcome this problem seems to use the same algorithm, such as the TES algorithm for both MODIS and ASTER. Fortunately, Hulley and Hook (2011) have tried to generate consistent LST products between ASTER and MODIS [66]. The new product (MOD21) generated by the TES algorithm has been reported to be released in MODIS Collection 6 [56], [67]. However, this new product is unavailable now due to data quality issues according to the announcement of the Land Processes Distributed Active Archive Center (LP DAAC). Thus, the compromise is that a simple linear regression is established between the ASTER TES-based LSTs and MODIS GSW-based LSTs to avoid algorithm differences. The determination coefficients of 0.93 for study area A and 0.90 for study area B within the 95% confidence interval prove good correlations between the LSTs from ASTER and MODIS; therefore, the comparison between these two products appears to be meaningful. For future studies, the cross-validation with ASTER LSTs will not be necessary. It is recommended to acquire in situ LSTs on the ground at the time of the satellite overpass to give a more accurate validation of the downscaling performance.

Second, the primary superiority of the RF model compared to other nonlinear regressions and its insensitivity to multicollinearity allows for as many predictor variables as need to be added into the RF model without selection. The importance score gives a relative ranking among the input variables to show the near or far relationships between the input variables and LST. As shown in Figure 4, the vegetation indices and terrain factors are more dominant than the other kind of variables, with almost coverage over the vegetation and mountainous regions in study area A; in contrast, the impervious surface indices seem to be more correlative, as the artificial materials comprising the downtown region cover study area B. This may give the indication regarding the selection of the input variables. However, the introduction of too many variables causes complexity in the MTVRF model. The number of input variables should be selected rationally rather than as much as possible.

In addition, limited by the specific area or data scene availability, the MTVRF model has not been evaluated for long periods of time, and only several subareas in the study area dominated by different land cover types were tested to support the robustness of the model. To supplement this method, mutual validations have been conducted for the training models in the two study areas, and acceptable results have proven the robustness of the MTVRF to a certain extent.

V. CONCLUSIONS

In this paper, the performance of random forest regression model based on the relationships between multitype predictor variables, including surface reflectance, spectral indices, terrain factors and land classification maps, and land surface temperature has been comprehensively discussed to downscale MODIS LSTs from 990 m to 90 m and evaluated by ASTER LST products (90 m) with RMSE and ME. Both the visual comparison and statistical measures prove that the MTVRF model achieves a desired downscaling result either on regions with single types like vegetation, water as well as impervious surface, or on the complicated underlying surfaces, which can be validated by the comparison with the basic RF model and TsHARP method.

The appropriate nonlinear model should be chosen to fit the correlation between LST and other biophysical variables. Random forest regressions, given their great tolerance to multicollinearity, can address high-dimension datasets without overfitting. This meets the requirement of the regression downscaling model to train all related variables and determine the desired result. The relative correlations, given by the importance scores calculated by the RF model, guide the variable selection. The TsHARP method, which uses the vegetation fraction as the input variable, usually has an ideal result in regions with high vegetation; the single indicator always results in a substantial amount of loss in the spatial pattern. In this paper, the MTVRF model and two other downscaling methods were applied to downscale LSTs under various surface conditions. Satisfactory results were obtained for the MTVRF model, with a RMSE of 2.67 K in Segovia and 2.10 K in Beijing, which outperformed the other two methods, with a RMSE of 3.32 K in Segovia and 2.14 K in Beijing for the basic RF model and a RMSE of 3.62 K in Segovia and 2.16 K in Beijing when using TsHARP. The RMSEs of the MTVRF model are 2.80 K, 1.91 K and 1.92 K for water, vegetation and impervious surfaces, respectively, which are still lower than those of the basic RF model and TsHARP method. In addition, the robustness evaluation for the downscaling model gives further evidence about the superiority of our MTVRF, with a reduced RMSE while simultaneously predicting LSTs by using the downscaling model trained in other regions. That is, the MTVRF is less likely to be limited in the study site where it is trained. From the above results, it can be deduced that the MTVRF model outperforms the TsHARP method and basic RF model when downscaling LST under different terrain conditions and for land cover types.

In addition, the MTVRF model provided us with an indication for the selection of predictor variables in different kinds of regions. While being trained in mountainous areas, terrain factors, such as DEMs, would play a significant role in the description of LST variations, while areas with single land cover types that are being trained and targeted spectral indices with higher importance scores should be considered, such as the NDVI for vegetation regions and NDBI for impervious surfaces. Future studies will concentrate on the introduction of geographical information (e.g., latitude and longitude), temporal information (e.g., Julian days), and other surface status information (e.g., net surface shortwave radiation, soil moisture and wind velocity) to further improve the generalization ability of our MTVRF model.

APPENDIX

The brief introduction of spectral indices in the MTVRF model is listed as following:

Variables	Formula
NDVI	$NDVI = \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + \rho_R}$
Fv	$Fv = \frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}$
NMDI	$NMDI = \frac{\rho_{NIR} - (\rho_{SWIR1} - \rho_{SWIR2})}{\rho_{NIR} + (\rho_{SWIR1} - \rho_{SWIR2})}$
NDDI	$NDDI = \frac{\rho_{SWIR2} - \rho_B}{\rho_{SWIR2} + \rho_B}$
MNDWI	$MNDWI = \frac{\rho_G - \rho_{SWIR1}}{\rho_G + \rho_{SWIR1}}$
SAVI	$SAVI = \frac{(\rho_{NIR} - \rho_R)(1+L)}{\rho_{NIR} + \rho_R + L}$
NDBI	$NDBI = \frac{\rho_{SWIR1} - \rho_{NIR}}{\rho_{SWIR1} + \rho_{NIR}}$

Remarks : ρ refers to the reflectance in VNIR/SWIR bands, the subscripts like R, G, B, NIR as well as SWIR represent the red, green, blue, near infrared red and short-wave infrared bands of the image. L means the soil adjustment coefficient determined by vegetation density, with the value range from 0 to 1. In study area A, the value of L is 0.8 with much vegetation covered while L in study area B is 0.5 with almost half vegetation covered.

REFERENCES

[1] L. V. Alexander, S. K. Allen, and L. Nathaniel, "Climate change 2013: The physical science basis summary for policymakers," Intergovernmental Panel Climate Change, Geneva, Switzerland, Tech. Rep. 12, 2013.
[2] N. A. Brunsell and R. R. Gillies, "Length scale analysis of surface energy fluxes derived from remote sensing," *J. Hydrometeorol.*, vol. 4, no. 6, pp. 1212–1219, Dec. 2003.
[3] A. Karnieli *et al.*, "Use of NDVI and land surface temperature for drought assessment: Merits and limitations," *J. Climate*, vol. 23, pp. 618–633, Feb. 2010.
[4] W. Kustas and M. Anderson, "Advances in thermal infrared remote sensing for land surface modeling," *Agricult. Forest Meteorol.*, vol. 149, pp. 2071–2081, Dec. 2009.
[5] Z.-L. Li *et al.*, "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013.
[6] R. Zhang, J. Tian, H. Su, X. Sun, S. Chen, and J. Xia, "Two improvements of an operational two-layer model for terrestrial surface heat flux retrieval," *Sensors*, vol. 8, no. 10, pp. 6165–6187, 2008.
[7] A. J. Arnfield, "Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island," *Int. J. Climatol.*, vol. 23, no. 1, pp. 1–26, 2003.
[8] W. G. M. Bastiaanssen *et al.*, "A remote sensing surface energy balance algorithm for land (SEBAL).: Part 2: Validation," *J. Hydrometeorol.*, vol. 212, nos. 1–4, pp. 213–229, 1998.
[9] J. Hansen, R. Ruedy, M. Sato, and K. Lo, "Global surface temperature change," *Rev. Geophys.*, vol. 48, no. 4, 2010, Art. no. RG4004.
[10] J. D. Kalma, T. R. McVicar, and M. F. McCabe, "Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data," *Surv. Geophys.*, vol. 29, no. 4, pp. 421–469, Oct. 2008.
[11] F. N. Kogan, "Operational space technology for global vegetation assessment," *Bull. Amer. Meteorol. Soc.*, vol. 82, pp. 1949–1964, Sep. 2001.
[12] J. C. Price, "Using spatial context in satellite data to infer regional scale evapotranspiration," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 5, pp. 940–948, Sep. 1990.

[13] W. P. Kustas, J. M. Norman, M. C. Anderson, and A. N. French, "Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship," *Remote Sens. Environ.*, vol. 85, no. 4, pp. 429–440, 2003.

[14] P. M. Atkinson, "Downscaling in remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 22, pp. 106–114, Jun. 2013.

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

[16] M. Bisquert, J. M. Sánchez, and V. Caselles, "Evaluation of disaggregation methods for downscaling MODIS land surface temperature to Landsat spatial resolution in Barrax test site," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 4, pp. 1430–1438, Apr. 2016.

[17] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and P. D. Colaizzi, "Utility of thermal sharpening over Texas high plains irrigated agricultural fields," *J. Geophys. Res.-Atmos.*, vol. 112, no. D19, pp. 1–10, 2007.

[18] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and P. D. Colaizzi, "Utility of thermal image sharpening for monitoring field-scale evapotranspiration over rainfed and irrigated agricultural regions," *Geophys. Res. Lett.*, vol. 35, no. 2, pp. 196–199, 2008.

[19] 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, 2007.

[20] L. Chen, G. Yan, H. Ren, and A. Li, "A modified vegetation index based algorithm for thermal imagery sharpening," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 2444–2447.

[21] W. Essa, B. Verbeiren, J. van der Kwast, T. Van de Voorde, and O. Batelaan, "Evaluation of the DisTrad thermal sharpening methodology for urban areas," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 19, no. 1, pp. 163–172, 2012.

[22] J. J. Qu, X. Hao, M. Kafatos, and L. Wang, "Asian dust storm monitoring combining terra and aqua MODIS SRB measurements," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 4, pp. 484–486, Oct. 2006.

[23] I. Sandholt, C. Nielsen, and S. Stisen, "A simple downscaling algorithm for remotely sensed Land Surface Temperature," in *Proc. Amer. Geophys. Union Spring Meeting Abstr.*, Washington, DC, USA, 2014.

[24] H. Xu, "A new index for delineating built-up land features in satellite imagery," *Int. J. Remote Sens.*, vol. 29, no. 14, pp. 4269–4276, 2008.

[25] K. Zakšek and K. Oštir, "Downscaling land surface temperature for urban heat island diurnal cycle analysis," *Remote Sens. Environ.*, vol. 117, no. 1, pp. 114–124, 2012.

[26] Y. Chen, W. Zhan, J. Quan, J. Zhou, X. Zhu, and H. Sun, "Disaggregation of remotely sensed land surface temperature: A generalized paradigm," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5952–5965, Sep. 2014.

[27] A. Dominguez, J. Kleissl, J. C. Luvall, and D. L. Rickman, "High-resolution urban thermal sharpener (HUTS)," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1772–1780, 2011.

[28] 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, 2017.

[29] X. Chen, W. Li, J. Chen, Y. Rao, and Y. Yamaguchi, "A combination of TsHARP and thin plate spline interpolation for spatial sharpening of thermal imagery," *Remote Sens.*, vol. 6, no. 4, pp. 2845–2863, 2014.

[30] M. Cubero-Castan, J. Chanussot, V. Achard, X. Briottet, and M. Shimoni, "A physics-based unmixing method for thermal hyperspectral images," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2015, pp. 5082–5086.

[31] C. Deng and C. Wu, "Examining the impacts of urban biophysical compositions on surface urban heat island: A spectral unmixing and thermal mixing approach," *Remote Sens. Environ.*, vol. 131, pp. 262–274, Apr. 2013.

[32] L. J. Guo and J. M. Moore, "Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery," *Int. J. Remote Sens.*, vol. 19, no. 13, pp. 2477–2491, 1998.

[33] J. Nichol, "An emissivity modulation method for spatial enhancement of thermal satellite images in urban heat island analysis," *Photogramm. Eng. Remote Sens.*, vol. 75, no. 5, pp. 547–556, 2009.

[34] M. Stathopoulou and C. Cartalis, "Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation," *Remote Sens. Environ.*, vol. 113, no. 12, pp. 2592–2605, 2009.

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

[36] T. N. Carlson and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol. 62, no. 3, pp. 241–252, 1997.

[37] G. Petropoulos, T. N. Carlson, M. J. Wooster, and S. Islam, "A review of Ts/VI remote sensing based methods for the retrieval of land surface energy fluxes and soil surface moisture," *Prog. Phys. Geogr., Earth Environ.*, vol. 33, no. 2, pp. 224–250, 2009.

[38] C. Jeganathan, N. A. S. Hamm, S. Mukherjee, P. M. Atkinson, P. L. N. Raju, and V. K. Dadhwal, "Evaluating a thermal image sharpening model over a mixed agricultural landscape in India," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 2, pp. 178–191, 2011.

[39] Y. Yang, X. L. Li, and C. Cao, "Downscaling urban Land Surface Temperature based on multi-scale factor," *Sci. Surveying Mapping*, vol. 42, no. 10, pp. 73–79, 2017.

[40] G. Yang, R. Pu, C. Zhao, W. Huang, and J. Wang, "Estimation of subpixel land surface temperature using an endmember index based technique: A case examination on ASTER and MODIS temperature products over a heterogeneous area," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1202–1219, 2011.

[41] B. Aiazzi, L. Alparone, S. Baronti, L. Santurri, and M. Selva, "Spatial resolution enhancement of ASTER thermal bands," *Proc. SPIE*, vol. 5982, Oct. 2005, Art. no. 59821G.

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

[43] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006.

[44] E. Pardo-Igúzquiza, M. Chica-Olmo, and P. M. Atkinson, "Downscaling cokriging for image sharpening," *Remote Sens. Environ.*, vol. 102, no. 1, pp. 86–98, 2006.

[45] S.-B. Duan and Z.-L. Li, "Spatial Downscaling of MODIS land surface temperatures using geographically weighted regression: Case study in northern china," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6458–6469, Nov. 2016.

[46] G. P. Lemeshewsky, "Neural network-based sharpening of Landsat thermal-band images," *Proc. SPIE*, vol. 3387, pp. 378–386, Jul. 1998.

[47] 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, 2001.

[48] G. Yang, R. Pu, W. Huang, J. Wang, and C. Zhao, "A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2170–2178, Apr. 2010.

[49] F. Gao, W. P. Kustas, and M. C. Anderson, "A data mining approach for sharpening thermal satellite imagery over land," *Remote Sens.*, vol. 4, no. 11, pp. 3287–3319, 2012.

[50] J. A. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 2, Jul. 2000, pp. 813–815.

[51] 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.

[52] A. Gillespie *et al.*, "A temperature and emissivity separation algorithm for advanced Spaceborne thermal emission and reflection radiometer (ASTER) images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1113–1126, Apr. 1998.

[53] 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.

[54] G. Hulley, S. Veraverbeke, and S. Hook, "Thermal-based techniques for land cover change detection using a new dynamic MODIS multispectral emissivity product (MOD21)," *Remote Sens. Environ.*, vol. 140, no. 1, pp. 755–765, 2014.

[55] 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.

[56] S.-B. Duan, Z.-L. Li, J. Cheng, and P. Leng, "Cross-satellite comparison of operational land surface temperature products derived from MODIS and ASTER data over bare soil surfaces," *ISPRS J. Photogramm. Remote Sens.*, vol. 126, pp. 1–10, Apr. 2017.

[57] S. L. Ermida, C. C. Dacamara, I. F. Trigo, A. C. Pires, D. Ghent, and J. Remedios, "Modelling directional effects on remotely sensed land surface temperature," *Remote Sens. Environ.*, vol. 190, pp. 56–69, Mar. 2017.

[58] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.

[59] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[60] T. G. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," *Mach. Learn.*, vol. 40, no. 2, pp. 139–157, 2000.

[61] Y. Amit and D. Geman, "Shape quantization and recognition with randomized trees," in *Neural Computation*. Cambridge, MA, USA: MIT Press, 1997, pp. 1545–1588.

[62] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, "Random Forests for land cover classification," *Pattern Recognit. Lett.*, vol. 27, no. 4, pp. 294–300, 2006.

[63] L. R. Iverson, A. M. Prasad, S. N. Matthews, and M. Peters, "Estimating potential habitat for 134 eastern US tree species under six climate scenarios," *Forest Ecol. Manage.*, vol. 254, no. 3, pp. 390–406, 2008.

[64] M. Pal, "Random forests for land cover classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 6, Jul. 2004, pp. 3510–3512.

[65] G. James, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. New York, NY, USA: Springer, 2013.

[66] G. C. Hulley and S. J. Hook, "Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for earth science research," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1304–1315, Apr. 2011.

[67] G. C. Hulley, S. J. Hook, and C. Hughes, "MODIS MOD21 Land Surface Temperature and emissivity algorithm theoretical basis document," Jet Propuls. Lab., Nat. Aeronaut. Space Admin., Pasadena, CA, USA, Tech. Rep. 12-17, 2012.

HUA WU received the B.S. degree in photogrammetric engineering and remote sensing from Wuhan University, Wuhan, China, in 2003, the M.S. degree in cartography and geographical information system from Beijing Normal University, Beijing, China, in 2006, and the Ph.D. degree in cartography and geographical information system from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, in 2010.

He is currently an Associate Research Fellow with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. His research interests include the retrieval, validation, and scaling of remotely sensed products.

WAN LI received the B.S. degree in geographical condition monitoring from Wuhan University, Wuhan, China, in 2012. She is currently pursuing the M.S. degree in cartography and geographical information system with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2018. Her research interest includes the scale transformation of land surface temperature.