A physical method for downscaling land surface temperatures using surface energy balance theory

Yongxin Hu, Ronglin Tang, Xiaoguang Jiang, Zhao-Liang Li, Yazhen Jiang, Meng Liu, Caixia Gao, Xiaoming Zhou

A Physical Method for Downscaling Land Surface Temperatures Using Surface Energy Balance Theory

Yongxin Hu[1,2,3], Ronglin Tang[2,1,*], Xiaoguang Jiang[1,4,*], Zhao-Liang Li[2,1,3], Yazhen Jiang[2,1], Meng Liu[5], Caixia Gao[4] and Xiaoming Zhou[6]

[1]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China;

[2]State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China;

[3]ICube Laboratory, (UMR 7357), CNRS, University of Strasbourg, 300 bd Sébastien Brant, CS 10413, F-67412 Illkirch, France;

[4]Key Laboratory of Quantitative Remote Sensing Information Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China;

[5]Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing, China;

[6]School of Civil Engineering, Lanzhou University of Technology, Lanzhou, Gansu, China.

* Corresponding author:

Ronglin Tang (tangrl@lreis.ac.cn), Xiaoguang Jiang (xgjiang@ucas.ac.cn)

Abstract

Fine-resolution land surface temperature (LST) derived from thermal infrared remote sensing images is a good indicator of surface water status and plays an essential role in the exchange of energy and water between land and atmosphere. A physical surface energy balance (SEB)-based LST downscaling method (DTsEB) is developed to downscale coarse remotely sensed thermal infrared LST products with fine-resolution visible and near-infrared data. The DTsEB method is advantageous for its ability to mechanically interrelate surface variables contributing to the spatial variation of LST, to quantitatively weigh the contributions of each related variable within a physical framework, and to efficaciously avoid the subjective selection of scaling factors and the establishment of statistical regression relationships. The applicability of the DTsEB method was tested by downscaling 12 scenes of 990 m Moderate Resolution Imaging Spectroradiometer (MODIS) and aggregated Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) LST products to 90 m resolution at six overpass times between 2005 and 2015 over three 9.9 km by 9.9 km cropland (mixed by grass, tree, and built-up land) study areas. Three typical LST downscaling methods, namely the widely applied TsHARP, the later developed least median square regression downscaling (LMS) and the geographically weighted regression (GWR), were introduced for intercomparison. The results showed that the DTsEB method could more effectively reconstruct the subpixel spatial variations in LST within the coarse-resolution pixels and achieve a better downscaling accuracy than the TsHARP, LMS and GWR methods. The DTsEB method yielded, on average, root mean square errors (RMSEs) of 2.01 K and 1.42 K when applied to the MODIS datasets and aggregated ASTER datasets, respectively, which were lower than those obtained with the TsHARP method, with average RMSEs of 2.41 K and 1.71 K, the LMS method, with average RMSEs of 2.35 K and 1.63 K, and the GWR method, with average RMSEs of 2.38 K and 1.64 K, respectively. The contributions of the related surface variables to the subpixel spatial variation in the LST varied both spatially and temporally and were different from each other. In summary, the DTsEB method was demonstrated to outperform the TsHARP, LMS, and GWR methods and could be used as a good alternative for downscaling LST products from coarse to fine resolution with high robustness and accuracy.

Keywords: Land surface temperature; Thermal infrared remote sensing; Surface energy balance; Downscaling; DTsEB

1. Introduction

As a key parameter in the characterization of energy and water exchange between land and atmosphere [*Li et al., 2013; Duan & Li, 2016; Anderson et al., 2008*], land surface temperature (LST) has been widely applied in a variety of disciplines and related studies on evapotranspiration estimation [*Tang et al., 2010; Tang & Li, 2017a*], urban heat island monitoring [*Weng et al., 2004; Quan et al., 2014*], forest fire detection [*Eckmann et al., 2008*] and biogeochemical process modeling [*Zhan et al., 2016*]. Satellite-based thermal infrared (TIR) data are directly linked to LSTs through the radiative transfer equation and are recognized as the most reliable source for deriving regional LSTs in a globally consistent and economically feasible manner. However, because of technical constraints, satellite-derived LST datasets always reflect a tradeoff between temporal and spatial resolutions [*Bindhu et al., 2013; Liu et al., 2020*]. Even within the same satellite, thermal sensors have much lower spatial resolution than visible and near-infrared (VNIR) sensors due to the relatively lower levels of thermal radiation that are emitted by land surface. The relatively low spatial resolution of satellite-derived TIR LSTs often leads to a thermal mixture effect (i.e., blending of multiple thermal elements within a single coarse spatial resolution pixel) [*Yang et al., 2011; Zhan et al., 2013*] and provides a weak representation of the heterogeneity of the land surface water status and energy exchange. Therefore, techniques to enhance the spatial resolution of satellite-derived LST datasets are highly desirable [*Atkinson, 2013; Hutengs & Vohland, 2016*].

Over the past decades, a large number of LST downscaling (also called *LST sharpening* or *LST disaggregating*) methods have been proposed to transform coarse-resolution LSTs to fine resolution with thermal radiances at the two resolutions invariantly maintained [*Kustas et al., 2003; Zhan et al., 2012, 2013*]. Most of these LST downscaling methods assume that the relationships between surface parameters (e.g., independent variables) and LSTs (e.g., dependent variable) are scale invariant, and they first construct a mathematical relationship at coarse resolution and subsequently apply the relationship to surface parameters at fine resolution to obtain fine-resolution LSTs. The surface parameters used in LST downscaling have been collectively referred to as scaling factors in some studies (also called *kernels* or *modulation factors* in some other studies); they represent indicators that connect the LSTs and should be achieved at both coarse and fine resolutions [*Zhan et al., 2011, 2013; Liu & Moore, 1998; Stathopoulou & Cartalis, 2009; Chen et al., 2014*]. Based on the fact that the regional LST variability is most affected by vegetation

coverage, vegetation indices (VIs) are the most widely used scaling factors in LST downscaling algorithms. Since the pioneering LST downscaling methods developed by *Kustas et al. [2003]* and *Agam et al. [2007]* (named *DisTrad* and *TsHARP*) were proposed, successful tests and applications of various VIs, such as the normalized difference vegetation index (NDVI) or its derivatives – fractional vegetation cover (f_c), soil-adjusted vegetation index (SAVI) [*Yang et al., 2011*], enhanced vegetation index (EVI) [*Zakšek & Oštir, 2012*], and green ratio vegetation index (GRVI) [*Bonafoni, 2016*] – over different land cover types (especially agricultural areas) have been reported in the LST downscaling literature [*Jeganathan et al., 2011; Essa et al., 2012; Bisquert et al., 2016a; Bisquert et al., 2016b; Olivera-Guerra et al., 2017*]. However, because the LST is influenced by multiple factors, the relationships between VIs and LSTs display great limitations for heterogeneous underlying surfaces [*Inamdar & French, 2009; Nichol, 2009; Stathopoulou & Cartalis, 2009; Zakšek & Oštir, 2012*]. To compensate for the deficiency of using only simple vegetation parameters in LST downscaling, spectral indices representing different types of land surfaces, such as the normalized difference water index (NDWI), the normalized difference built-up index (NDBI), the enhanced built-up and bareness index (EBBI), the bare soil index (BI), and the temperature vegetation dryness index (TVDI), have also been introduced [*Yang et al., 2011; Zakšek & Oštir, 2012; Merlin et al., 2012; Bonafoni, 2016; Li et al., 2019; Agathangelidis & Cartalis, 2019; Liu et al., 2020*]. In addition, to modulate the land surface energy distribution, topographic variables, namely, digital elevation models (DEMs) and slope angle, surface emissivity and broadband albedo, were also suggested to be good scaling factors [*Duan & Li, 2016; Li et al., 2019; Agathangelidis and Cartalis, 2019; Inamdar & French, 2009; Nichol, 2009; Stathopoulou & Cartalis, 2009; Zakšek & Oštir, 2012; Merlin et al., 2010, 2012; Dominguez et al., 2011*]. However, although many studies have reported that the accuracy of LST downscaling results increases with the introduction of new or multiple scaling factors, the selection of scaling factors (usually without explicit physical implications) remains mostly subjective, uncertain, and application-specific.

For the relationships between scaling factors and LSTs, statistical regression has been widely used in LST downscaling. Both simple linear or quadratic regression relationships and multivariate nonlinear regression relationships between single/multiple scaling factors (e.g., NDVI, DEM, broadband albedo, and surface emissivity) and LSTs were explored in early years [*Kustas et al., 2003; Agam et al., 2007; Dominguez et al, 2011; Yang et al., 2011; Merlin et al., 2010, Merlin et*

al., 2012; Bonafoni, 2016; Bisquert et al., 2016a]. In recent years, more advanced regression methods, including tree-based regression techniques (such as the gradient boosting machine), kernel-based regression techniques (such as the support vector machine), artificial neural networks, and random forest techniques, which have great advantages in establishing stable relationships between multiple scaling factors and LST, have attracted extensive attention [*Yang et al., 2010; Ghosh et al., 2014; Hutengs and Vohland, 2016; Li et al., 2019; Wu and Li, 2019; Agathangelidis and Cartalis, 2019*]. Although great success has been reported for LST downscaling, the lack of a clear physical mechanism in these empirical regression relationship-based methods has resulted in difficulties in understanding the interactions of various scaling factors with LSTs and has further hindered the extension of LST downscaling methods to other study areas. Moreover, the selection of different scaling factors and the establishment of corresponding empirical relationship between scaling factors and LST in these different statistics-based methods can even result in different LSTs for the same pixel under different relationship expression formulas and/or different sizes of areas of interest (such as the different LST downscaling results for Changping area in the researches of *Yang et al.* [*2011*] and *Wu & Li* [*2019*]). To date, compared to the statistics-based LST downscaling methods, (semi)physical methods for achieving LST downscaling have seldom been proposed [*Bechtel et al., 2012*]. While tests have been developed to enhance the thermal details, the apparently unreasonable assumptions of invariable emitted energy (with surface emissivity as the sole factor influencing temperature variation) among the subpixels [*Nichol, 2009*] or isothermal pixels [*Liu & Pu, 2008; Liu & Zhu, 2012*] have greatly hindered the wide application of these (semi)physical methods. The dual band method proposed by Dozier [*Dozier, 1981*] to estimate the percentage of thermally anomalous coverage and its temperature also faces technical limitations because of the requirements of a priori knowledge (such as the predefined specific end members) [*Zakšek & Oštir, 2012; Bechtel et al., 2012*]. In fact, in the energy and water exchange between land and atmosphere, the spatial variation of LST is more physically and fundamentally linked to the surface energy budget (i.e., the surface net radiation, soil heat flux, sensible heat flux and latent heat flux) than the land surface vegetation information or land cover details [*Zhan et al., 2013*]. From this point of view, the surface energy balance is very likely to provide a robust and physical LST downscaling framework to take into account all parameters that drive the spatial variation of LSTs [*Zhan et al., 2013; Merlin et al., 2010*].

The objective of this study is to develop a physical surface energy balance (SEB)-based LST downscaling method (DTsEB) to avoid a subjective selection of scaling factors and the establishment of statistical regression relationships. To this end, the theoretical relationship between scaling factors and LSTs is at first deduced from the SEB equation and Penman-Monteith equation. Subsequently, the DTsEB method is developed by calculating the total differential of the LST, and then, the fine-resolution LSTs can be obtained by converting the differences between the LSTs at fine and coarse resolutions into differences between the VNIR/SWIR scaling factors at the two resolutions. Finally, this new approach is tested on both 990 m resolution Moderate Resolution Imaging Spectroradiometer (MODIS) LST and aggregated 990 m resolution Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) LST data at six overpass times between 2005 and 2015 over three 9.9 km by 9.9 km cropland (mixed by grass, tree, and built-up land) study areas. As a reference, three typical LST downscaling methods, namely the well-known and most widely applied TsHARP method, the later developed least median square regression downscaling (LMS) method, and the geographically weighted regression (GWR) method, are intercompared with the DTsEB method. Section 2 describes the study areas, the remotely sensed MODIS and ASTER data, and the auxiliary ground-based meteorological data involved in this study. Section 3 presents the methodology of how LST downscaling is performed with the DTsEB method. Section 4 provides the LST downscaling results and discusses the uncertainty, superiority, and weakness of the DTsEB method. Finally, a summary and conclusions are presented in Section 5.

2. Study Area and Data

2.1. Study area

Three study areas (Figure 1) with spatial dimensions of 9.9 km by 9.9 km and characterized by different climates and soil types were selected to evaluate the adaptability and accuracy of the DTsEB method.

Study area A, which surrounds the Yucheng site (36.8291 °N/116.5703 °E, indicated by the solid black triangle in Figure 1), is located in the southwestern part of Yucheng County, Shandong Province, North China. This area is characterized by a subhumid monsoon climate with a mean annual temperature and precipitation of 13.1 ℃ and 528 mm, respectively. The soil type is classified as sandy loam and the land cover types primarily consist of croplands, bare soil and

built-up lands (including roads and buildings). Winter wheat (seeded in mid-October and harvested in mid-June) and summer corn (seeded in late June and harvested in early October) are rotated in this study area. More details of regarding the characteristics of this area can be found in the work by *Tang & Li* [*2017b*].

Study area B, which surrounds the US_Bo1 site (40.0062 °N/88.2904 °W, marked by the solid black star in Figure 1), is located in the midwestern part of the United States near Champaign, Illinois. Study area B has a deep silty clay loam soil type and is characterized by a humid continental climate with a mean annual temperature of 11 ℃ and mean annual precipitation of 991 mm. The land cover types mainly include crops, trees, and built-up lands. Corn and soybeans are rotated annually in this study area. The detailed information on study area B was provided by *Meyers & Hollinger* [*2004*].

Study area C, which surrounds the US_Ne1 site (41.1651 °N/96.4766 °W, indicated by the solid black circle in Figure 1), is located at the University of Nebraska Agricultural Research and Development Center near Mead, Nebraska, United States. Study area C has a humid continental climate with a mean annual temperature of 10 ℃ and mean annual precipitation of 790 mm. The soil type is classified as deep silty clay loam and the land cover/land use types of this study area are mainly crops, built-up areas, trees, and grass. Maize and soybeans are seeded under no-till and are harvested in early November and October, respectively. The detailed information of study area C was provided by *Verma et al.* [*2005*].

Surface four-component radiation (downwelling and upwelling shortwave and longwave radiation) and meteorological variables, including precipitation, wind speed, relative humidity, air temperature, and atmospheric pressure were regularly measured at half-hourly intervals at the three ground-based sites. The auxiliary meteorological data required for evaluating the LST downscaling methods were obtained from the three sites over the three study areas (Yucheng site in study area A, US_Bo1 site in study area B and US_Ne1 site in study area C).

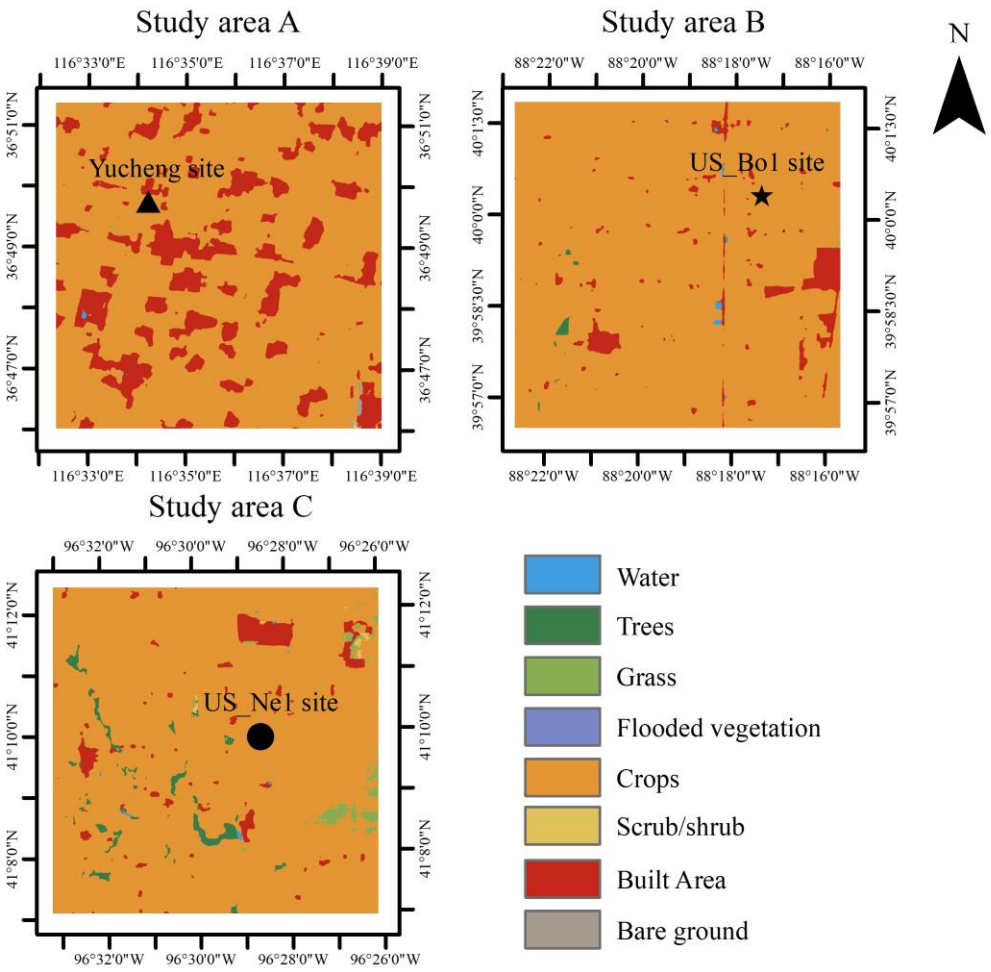

Figure 1. Site geolocations (Yucheng site in study area A, US_Bo1 site in study area B and US_Ne1 site in study area C) and land use/land cover (LULC) types of the three study areas (retrieved from ESA Sentinel-2 imagery at 10 m resolution), including water, trees, grass, flooded vegetation, crops, built-up area, and bare ground.

2.2 Data

The data used in this study include remotely sensed MODIS and ASTER datasets onboard the same satellite platform and ground-based meteorological datasets. In each study area, datasets from two different dates were used to test the LST downscaling algorithm.

2.2.1 MODIS datasets

MODIS data collected from the MODIS/Terra platform, including the MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1 km SIN Grid product (MOD11A1, Collection 6) and MODIS/Terra Surface Reflectance Daily L2G Global 1 km and 500 m SIN Grid product (MOD09GA, Collection 6), were used as the original coarse-resolution data in this study. These

data were downloaded from the National Aeronautics and Space Administration's (NASA's) Level-1 and Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC) (https://ladsweb.modaps.eosdis.nasa.gov/search/). The two MODIS products were acquired on April 24, 2006, and October 3, 2010, for study area A, on April 15, 2005 and July 29, 2005, for study area B, and on September 2, 2014, and August 20, 2015, for study area C in HDF-EOS format and were reprojected into UTM WGS 1984 50 N, UTM WGS 1984 16 N, and UTM WGS 1984 14 N, respectively, with a resampling interval of 990 m by using the MODIS Reprojection Tool (MRT).

Coarse-resolution parameters, such as the normalized difference vegetation index (NDVI) and fractional vegetation cover (f_c) (used to calculate the soil heat flux), broadband albedo and surface emissivity (used to calculate the surface net radiation), and surface resistance and aerodynamic resistance (used to solve the Penman-Monteith equation) at 990 m can be obtained with the MOD11A1 and MOD09GA products. For example, the NDVI is calculated using the reflectance values in the red and near-infrared bands; the surface broadband albedo, r, is estimated using the method of *Liang* [*2001*] (the method of *Mokhtari et al. [2013]* is applied for ASTER data after April 2008 due to the malfunction of the shortwave infrared detectors), and the surface emissivity, ε_s, is calculated using the algorithm proposed by *Qin et al.* [*2004*] (see the Appendix D).

2.2.2 ASTER datasets

Concurrent remote sensing data from the ASTER sensor onboard the Terra satellite, including the ASTER L2 Surface Kinetic Temperature product (AST_08) and ASTER L2 Surface Reflectance Visible and Near Infrared (VNIR) and Shortwave Infrared (SWIR) product (AST_07), were used as the fine-resolution data in this study. These data were collected from NASA's Earthdata Search web application (https://search.earthdata.nasa.gov/search/).

The ASTER surface reflectance products (AST_07), with a pixel size of 15 m in the VNIR region and 30 m in the SWIR region were spatially aggregated to 90 m to match the original spatial resolution of ASTER LST (AST_08) in this study. The surface parameters, including NDVI, f_c, r, ε_s, surface resistance, and aerodynamic resistance at 90 m resolution can be calculated with the AST_07 data (see the Appendix D).

2.2.3 Ground-based meteorological data

Half-hourly surface meteorological variables from the three ground sites (e.g., the Yucheng site in study area A, the US-Bo1 site in study area B and the US_Ne1 site in study area C), including downward solar radiation, wind speed, relative humidity, air temperature and vapor pressure, were collected at the Terra satellite overpass times as the auxiliary data in this study. Given the limited spatial dimensions of the study areas, the spatial variation of near-surface meteorological data was low and might contribute less to the downscaled LST than the subpixel heterogeneity of surface parameters. The meteorological data from the ground sites were thus regarded as spatially representative over the entire study area.

2.2.4 Digital elevation data

The digital elevation data (DEM) was used as the auxiliary data in the GWR method. 30 m DEM data of the three study areas (N36E116 for study area A, N39W089 and N40W089 for study area B, and N41W097 for study area C) collected from the ASTGTM product (Version 3, https://search.earthdata.nasa.gov/search/) were spatially aggregated to 90 m and 990 m to match the fine and coarse resolution ASTER and MODIS datasets, respectively.

3. Methodology

3.1 DTsEB method

LST is a direct indicator in the exchange of long-wave radiation and turbulent heat fluxes at the land–atmosphere interface and can effectively characterize the physical processes of surface energy and water balance at local to global scales [*Li et al., 2013*]. Considering the physical interconnections between LSTs and land surface energy, we propose the DTsEB method (the flow chart is shown in Figure 2) by introducing the surface energy balance equation and Penman-Monteith equation, as follows:

$$R_n = G + H + LE \tag{1}$$

$$LE = \frac{\Delta(\mathrm{R_n} - \mathrm{G}) + \rho \mathrm{C}_p \, VPD / r_a}{\Delta + \gamma\left(1 + r_s / r_a\right)} \tag{2}$$

where R_n is the surface net radiation, W/m^2; G is the soil heat flux, W/m^2; H is the sensible heat flux, W/m^2; LE is the latent heat flux, W/m^2; Δ is the slope of the saturated vapor pressure versus air temperature curve, kPa/℃; ρ is the air density, kg/m^3; C_p is the specific heat of air, J/(kg℃);

VPD is the vapor pressure deficit of air, kPa; r_a is the aerodynamic resistance, s/m; r_s is the surface resistance, s/m; and γ is the psychrometric constant, kPa/ ℃.

The sensible heat flux (H) can be approximated by combining the difference between LST and air temperature (T_a) with the aerodynamic resistance (r_a), as follows:

$$H = \rho C_p \frac{LST - T_a}{r_a} \tag{3}$$

The soil heat flux (G) is often expressed as a fraction of the surface net radiation (R_n) from the remote sensing perspective and it is estimated following the work of *Su* [*2002*] in this study, as follows:

$$G = R_n \left[\Gamma_c + \left(1 - f_c\right)\left(\Gamma_s - \Gamma_c\right) \right] \tag{4}$$

where f_c is the fractional vegetation cover, and Γ_c and Γ_s are the fractions of G to R_n for fully covered vegetation and dry bare soil, respectively.

By combining and rearranging equations (1), (2), (3) and (4), we can mathematically express the LST with the following equations:

$$LST = \frac{r_a \left[1 - \Gamma_s + \left(\Gamma_s - \Gamma_c\right) f_c \right] R_n}{\rho C_p} \frac{\gamma\left(1 + r_s / r_a\right)}{\Delta + \gamma\left(1 + r_s / r_a\right)} - \frac{VPD}{\Delta + \gamma\left(1 + r_s / r_a\right)} + T_a \tag{5}$$

with

$$R_n = \left(1 - r\right) R_g + \varepsilon_s \varepsilon_a \sigma T_a^4 - \varepsilon_s \sigma LST^4 \tag{6}$$

$$f_c = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \right)^2 \tag{7}$$

$$r_a = \frac{\left[\ln\left(\frac{Z_u - d}{Z_{om}} \right) - \Psi_m \right] \left[\ln\left(\frac{Z_t - d}{Z_{oh}} \right) - \Psi_h \right]}{k^2 u} \tag{8}$$

$$r_s = \frac{1}{C_L m\left(T_{min}\right) m\left(VPD\right) LAI} \tag{9}$$

where R_g is the global solar radiation, W/m^2; r is the surface albedo; ε_s is the surface emissivity; ε_a is the atmospheric emissivity; σ is the Stefan-Boltzmann constant; and $NDVI_{min}$ and $NDVI_{max}$ are the minimum NDVI corresponding to bare soil and the maximum NDVI corresponding to fully vegetated surfaces, respectively; Z_u and Z_t are the heights at which the wind speed and air temperature are observed, respectively, m; k is the von Karman constant; u is the wind speed, m/s; d is the zero plane displacement height, m; z_{om} is the surface momentum roughness height, m; z_{oh} is the roughness height for surface heat transfer, m; Ψ_m and Ψ_h are the stability correction functions for momentum and heat transfer, respectively [*Paulson*, 1970]; *LAI* is the leaf area index; C_L is the mean potential stomatal conductance per unit leaf area; $m(T_{min})$ is a multiplier that limits the potential stomatal conductance by the minimum air temperature; and $m(VPD)$ is a multiplier used to reduce the potential stomatal conductance when the VPD is sufficient to reduce the canopy conductance [*Mu et al., 2007, 2011*]. In Eq. (5), all the surface environmental and ecophysiological parameters including the fractional vegetation cover, surface emissivity, broadband albedo, aerodynamic resistance, surface resistance, and surface net radiation that are involved in the surface energy budget and drive the spatial variation of the LSTs are mechanically interrelated.

In this work, $\Gamma_c = 0.05$ and $\Gamma_s = 0.4$ are assumed in the soil heat flux calculations [*Daughtry et al., 1990; Li & Lyons, 1999; Tang et al., 2010*]; $z_{oh} = 0.1z_{om}$ where z_{om} is 0.125 times the vegetation height, h ($z_{om} = 0.125h$), and in this study, following the work of *Tang et al.* [*2013*] and *Teixeira et al.* [*2009*], h is estimated as a function of the surface albedo and NDVI. A summary of how these intermediate variables/parameters are estimated is provided in the Appendix D.

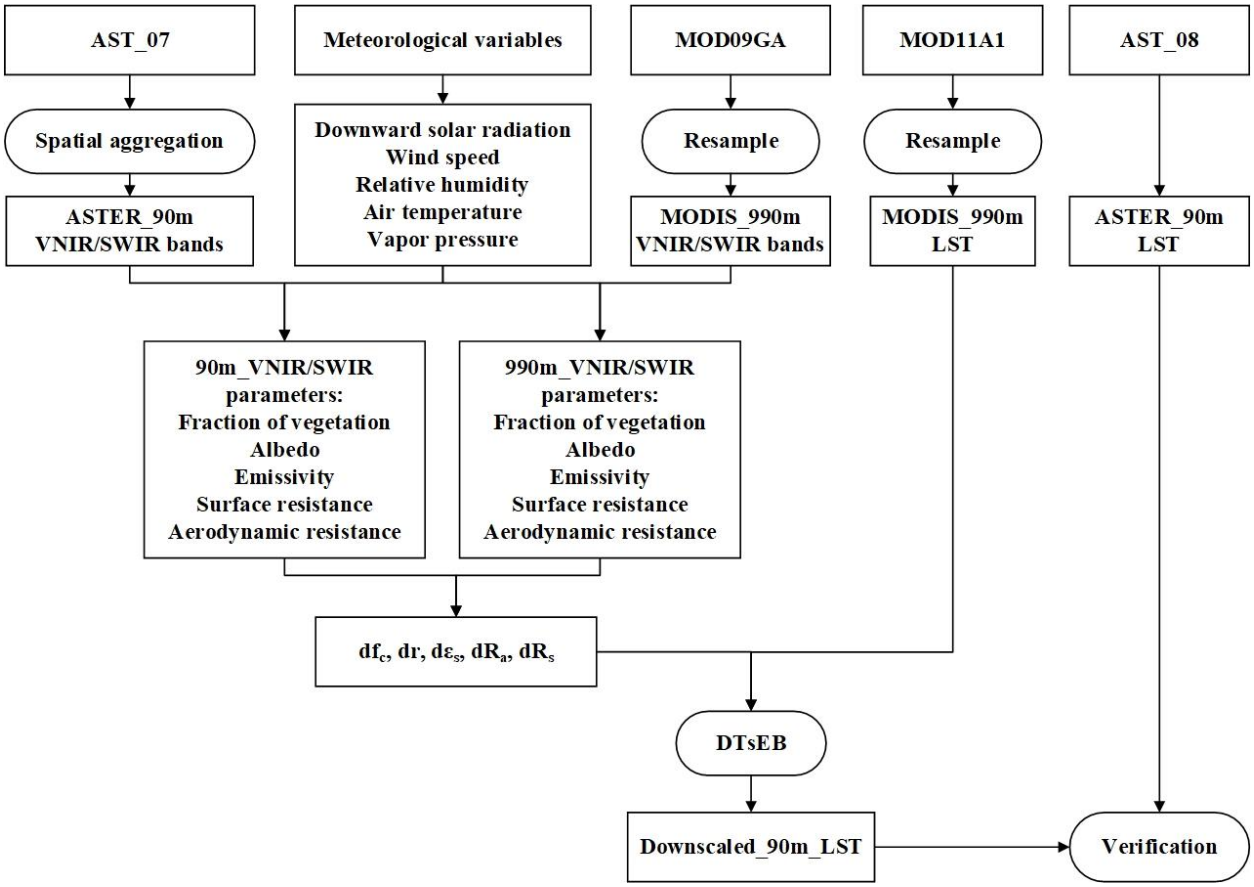

Figure 2. Flow chart of LST downscaling based on the DTsEB algorithm

Given the negligible spatial variations in the atmospheric parameters over the subpixels within a coarse pixel, the difference between fine-resolution LST and coarse-resolution LST primarily comes from the heterogeneity of the surface parameters (e.g., albedo, emissivity, fractional vegetation cover, resistance) within the coarse pixel and can be obtained by calculating the total differential of Eq. (5), as follows:

$$dLST = \frac{\partial LST}{\partial R_n} dR_n + \frac{\partial LST}{\partial f_c} df_c + \frac{\partial LST}{\partial r_a} dr_a + \frac{\partial LST}{\partial r_s} dr_s \tag{10}$$

with

$$dR_n = -R_g dr + \left(\varepsilon_a \sigma T_a^4 - \sigma LST_{CR}^4 \right) d\varepsilon_s - 4\varepsilon_s \sigma LST_{CR}^3 dLST \tag{11}$$

We combine equation (11) into equation (10), as follows:

$$dLST = \frac{\frac{\partial LST}{\partial R_n}\left[-R_g dr + \left(\varepsilon_a \sigma T_a^4 - \sigma LST_{CR}^4\right) d\varepsilon_s\right] + \frac{\partial LST}{\partial f_c} df_c + \frac{\partial LST}{\partial r_a} dr_a + \frac{\partial LST}{\partial r_s} dr_s}{1 + \frac{\partial LST}{\partial R_n} 4\varepsilon_s \sigma LST_{CR}^3} \quad (12)$$

in which,

$$\frac{\partial LST}{\partial R_n} = \frac{r_a\left[1-\Gamma_s+\left(\Gamma_s-\Gamma_c\right)f_c\right]}{\rho C_p} \frac{\gamma\left(1+r_s/r_a\right)}{\Delta+\gamma\left(1+r_s/r_a\right)} \quad (13)$$

$$\frac{\partial LST}{\partial f_c} = \frac{r_a\left(\Gamma_s-\Gamma_c\right)R_n}{\rho C_p} \frac{\gamma\left(1+r_s/r_a\right)}{\Delta+\gamma\left(1+r_s/r_a\right)} \quad (14)$$

$$\frac{\partial LST}{\partial r_a} = \frac{\left[1-\Gamma_s+\left(\Gamma_s-\Gamma_c\right)f_c\right]R_n}{\rho C_p} \frac{\gamma\Delta+\gamma^2\left(1+r_s/r_a\right)^2}{\left[\Delta+\gamma\left(1+r_s/r_a\right)\right]^2} - \frac{\gamma VPD\, r_s/r_a{}^2}{\left[\Delta+\gamma\left(1+r_s/r_a\right)\right]^2} \quad (15)$$

$$\frac{\partial LST}{\partial r_s} = \frac{\left[1-\Gamma_s+\left(\Gamma_s-\Gamma_c\right)f_c\right]R_n}{\rho C_p} \frac{\gamma\Delta}{\left[\Delta+\gamma\left(1+r_s/r_a\right)\right]^2} + \frac{\gamma VPD/r_a}{\left[\Delta+\gamma\left(1+r_s/r_a\right)\right]^2} \quad (16)$$

where the $\frac{\partial LST}{\partial R_n}$, $\frac{\partial LST}{\partial f_c}$, $\frac{\partial LST}{\partial r_a}$ and $\frac{\partial LST}{\partial r_s}$ can be calculated with the coarse resolution VNIR/SWIR and ground-based meteorological data, the dR_n, df_c, dr_a, dr_s, dr, and $d\varepsilon_s$ are the variations of R_n, f_c, r_a, r_s, r, and ε_s between fine and coarse resolutions, respectively, and the $\frac{\partial LST}{\partial R_n} dR_n$, $\frac{\partial LST}{\partial f_c} df_c$, $\frac{\partial LST}{\partial r_a} dr_a$, and $\frac{\partial LST}{\partial r_s} dr_s$ represent the contribution of surface net radiation, the contribution of fraction of the vegetation, the contribution of aerodynamic resistance, and the contribution of surface resistance to *dLST*, respectively. The subscript CR stands for the variable at coarse resolution.

Overall, combining the surface energy balance equation and Penman-Monteith equation, we can express the differences between LSTs at fine and coarse resolutions as a function of the differences between the surface parameters (i.e., $\frac{\partial LST}{\partial R_n} dR_n$, $\frac{\partial LST}{\partial f_c} df_c$, $\frac{\partial LST}{\partial r_a} dr_a$ and $\frac{\partial LST}{\partial r_s} dr_s$, which can be obtained with VNIR/SWIR data) at the two resolutions. The final downscaled fine-resolution LST (LST_{FR}) can be expressed as the sum of the coarse-resolution LST (LST_{CR}) and the LST difference, as follows:

$$LST_{FR} = LST_{CR} + dLST \quad (17)$$

3.2 Model application

To better understand error propagation in LST downscaling and how the reference LST truth affects the model performance, the proposed physical-based DTsEB method is evaluated for two different cases. As a reference, the downscaling results from the widely applied TsHARP method [*Agam et al., 2007; Kustas et al., 2003*], the LMS method [*Mukherjee et al., 2014; Bisquert et al., 2016*], and the GWR method [Duan *et al., 2016*] (see the Appendix A, B, and C for a description of these three methods) are also intercompared.

Case 1: downscaling of the 990 m MODIS LST product to 90 m and validation against the concurrent 90 m ASTER LST product. In this case, six scenes of original coarse-resolution MODIS LST data (990 m) for different growing dates of crops over three study areas were first downscaled to 90 m. Then, the fine-resolution ASTER LST products at 90 m (AST_08) were used as the reference LST data to validate the downscaled results. Because both the MODIS and ASTER sensors are onboard the same satellite platform (Terra), errors caused by altitude and time differences between different satellites can be eliminated.

Case 2: downscaling of 990 m aggregated ASTER LST data to 90 m and validation against the 90 m ASTER LST product that was used for aggregation. In this case, the original 90 m ASTER datasets were first spatially aggregated to a resolution of 990 m by assuming the conservation of surface emitted energy for AST_08 LST products and by the arithmetic mean for the AST_07 reflectance products. The LST downscaling methods were then performed on these aggregated datasets. The aggregated 990 m LSTs were used here with an implicit assumption that the coarse-resolution LSTs are highly consistent with the reference fine-resolution LSTs. The uncertainties introduced by the differences between the coarse-resolution LSTs and reference fine-resolution LSTs can therefore be excluded to some extent.

During model application, essential quality control procedures were conducted to remove pixels (~ 0.3%) characterized by low vegetation cover (NDVI < 0.1) with extremely high r_a and r_s (> 1,000 s/m) values at 90 m resolution, to reduce the abnormities in the downscaled LSTs in the DTsEB method. Note that fine-resolution data in this study were from 90 m ASTER NVIR and SWIR reflectance products. More generally, for a targeted coarse-resolution LST to be downscaled

in practical applications, concurrent (or adjacent) reflectance measurements with the sensor onboard the same (or other) satellite platform can be used as the fine-resolution data.

3.3 Statistical Analyses

Once the LST downscaling results are obtained, they were compared with the 90 m reference fine-resolution ASTER LST. The following statistical metrics, namely the root mean square error (RMSE), mean absolute error (MAE), mean bias (BIAS), normalized root mean square error (NRMSE), and correlation coefficient (R), were calculated to measure the model performance, as shown in Equations (18) - (22):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(LST_{dsl,i} - LST_{ref,i})^2}{n}} \tag{18}$$

$$MAE = \frac{\sum_{i=1}^{n}\left|LST_{dsl,i} - LST_{ref,i}\right|}{n} \tag{19}$$

$$BIAS = \frac{\sum_{i=1}^{n}(LST_{dsl,i} - LST_{ref,i})}{n} \tag{20}$$

$$NRMSE = \frac{RMSE}{SD} \tag{21}$$

$$R = \frac{\sum_{i=1}^{n}(LST_{dsl,i} - \overline{LST_{dsl,i}})(LST_{ref,i} - \overline{LST_{ref,i}})}{\sqrt{\sum_{i=1}^{n}(LST_{dsl,i} - \overline{LST_{dsl,i}})^2}\sqrt{\sum_{i=1}^{n}(LST_{ref,i} - \overline{LST_{ref,i}})^2}} \tag{22}$$

where, LST_{dsl} is the downscaled LST, LST_{ref} is the reference LST, $\overline{LST_{dsl}}$ and $\overline{LST_{ref}}$ are the average values of LST_{dsl} and LST_{ref}, respectively, SD is the standard deviation of reference LST.

4. Results and Discussion

4.1 Application to MODIS datasets

Before using the ASTER LST data to evaluate the downscaling results of the coarse-resolution MODIS LST data, the original 990 m MODIS LST data, the aggregated 990 m ASTER

LST data and the reference 90 m ASTER LST data from study areas A, B and C were compared, as shown in Table 1 and Figure 3. A deviation was found between the ASTER LST products (AST08) and the MODIS LST products (MOD11A1) because the former were generated from the Temperature/Emissivity Separation (TES) algorithm while the latter were generated from the generalized spilt-window (GSW) algorithm. From a visual comparison, clear spatial pattern differences between the MODIS LST (Figure 3a) and ASTER LST (Figure 3b) data were observed. The spatial distribution of the MODIS LSTs was smooth, whereas the ASTER LSTs could more effectively reflect sharp variations and spatial heterogeneity over the three study areas. In addition, compared with the ASTER LSTs, the MODIS LSTs were lower overall (approximately 2 K lower on average, see Table 1), especially in the high value range of LSTs. For example, the differences between the highest ASTER and MODIS LSTs on April 24, 2006 for study area A, on July 29, 2005 for study area B and on August 20, 2015 for study area C were as great as ~10 K, ~14 K and ~12 K, respectively, while in the low value range of LSTs, the MODIS LSTs were observed to be somewhat higher than the ASTER LSTs (Figure 3 and Table 1). These differences imply that the MODIS LSTs had a narrower value distribution than the ASTER LSTs, and this narrower value distribution was much remarkable than that of aggregated ASTER LSTs.

Table 1. The minimum, maximum, and mean of the coarse-resolution MODIS LSTs and the aggregated ASTER LSTs for model applications and the reference fine-resolution ASTER LSTs for validation.

Study Area & Date		MODIS LST at 990 m			Aggregated ASTER LST at 990 m			ASTER LST at 90 m		
		Min. (K)	Max. (K)	Mean (K)	Min. (K)	Max. (K)	Mean (K)	Min. (K)	Max. (K)	Mean (K)
A	April 24, 2006	294.36	297.80	295.94	295.85	303.43	298.30	294.30	307.34	298.25
A	October 3, 2010	297.64	300.16	298.73	298.63	302.06	300.00	295.20	306.16	299.97
B	April 15, 2005	298.10	303.74	301.28	299.60	307.19	303.89	293.78	310.47	303.86
B	July 29, 2005	299.56	304.18	300.79	300.36	309.91	302.32	298.70	318.20	302.35
C	September 2, 2014	298.0	301.48	299.19	298.66	304.60	300.84	296.50	312.90	300.83
C	August 20, 2015	298.64	302.04	299.64	300.72	305.93	302.42	296.50	314.30	302.41

Figures 3 and 4 show the spatial patterns of the LSTs downscaled from MODIS products using the DTsEB, TsHARP, LMS and GWR methods and scatter plots of the comparisons between the downscaled LSTs and ASTER LSTs, respectively. In general, the spatial distributions of the downscaling results based on the DTsEB (Figure 3c), TsHARP (Figure 3d), LMS (Figure 3(e)) and GWR (Figure 3(f)) were basically consistent with the spatial distribution of the ASTER LSTs (Figure 3b). All the four LST downscaling methods were able to enhance the spatial details of the original coarse pixels to some extent. However, in reference to the ASTER LSTs, the DTsEB method could more successfully reconstruct the subpixel spatial variations within a coarse MODIS LST pixel and, in particular, display the texture features better than the TsHARP, LMS, and GWR methods for all three study areas. For example, from circled region 1 in study area A on April 24, 2006 and circled region 3 in study area C on September 2, 2014, the high-value subpixels within the low-value coarse-resolution pixels were better reproduced in the DTsEB downscaled LSTs, whereas smooth subpixel variation was represented in the TsHARP, LMS, and GWR downscaled LSTs. For the generally high LST values in study area B on April 15, 2005 (see circled region 2 in Figure 3), the spatial details of the low LST subpixels that were mixed with high-value pixels could also be better displayed by the DTsEB method. In summary, the DTsEB method outperformed the TsHARP, LMS, and GWR methods in all the study areas, i.e., A, B and C. Especially for high value ranges of LSTs, the DTsEB method better reproduced the spatial details of ASTER LSTs (but possibly with some scatters over study areas B and C, as shown in Figure 4) whereas the TsHARP, LMS, and GWR methods significantly underestimated the reference LSTs. In addition, the LST downscaling results of study areas A and C were better than those of study area B regardless of which method was used. Overall, for the six scenes of LST downscaling, the DTsEB method, with a lower root mean square error (RMSE) of 1.46~3.02 K, mean absolute error (MAE) of 1.20~2.53 K, mean bias of -2.06~-0.23 K, normalized RMSE (by standard deviation of 90 m reference ASTER data) of 0.70~1.29 and correlation coefficient of 0.38~0.79, achieved a higher accuracy than the TsHARP, LMS, and GWR methods. By contrast, the TsHARP method, with RMSE values of 1.57~3.21 K, MAE values of 1.31~2.71 K, mean bias values of -2.43~-1.19 K, normalized RMSE values of 0.90~1.39 and correlation coefficient values of 0.46~0.76, performed slightly worse than the LMS method, and the GWR method. With an intermediate performance, the LMS method had RMSE values of 1.56~3.38 K, MAE values of 1.30~2.88 K, mean bias values of -2.59~-1.04 K, normalized RMSE values of 0.89~1.44 and correlation

coefficient of values 0.39~0.77 and the GWR method had RMSE values of 1.57~3.25 K, MAE values of 1.30~2.74 K, mean bias values of -2.41~-1.15 K, normalized RMSE values of 0.90~1.33 and correlation coefficient of values 0.40~0.83 (Table 2 and 4). In other words, the RMSE decreased by 0.11~0.70 K (~17% on average), 0.04~0.50 K (~14% on average), and 0.08~0.58 K (~16% on average) and the MAE decreased by 0.11~0.71 K (~16% on average), 0.09~0.51 K (~14% on average) and 0.10~0.67 K (~15%) when the DTsEB method was applied for LST downscaling instead of the TsHARP method, the LMS method, and the GWR method, respectively.

As mentioned in the Methodology section, the sum of the surface net radiation (R_n), fractional vegetation cover (f_c), aerodynamic resistance (r_a), and surface resistance (r_s) contributions constitutes the final *dLST* (the difference in LST between 990 m and 90 m resolutions) in the DTsEB method. Figure 5 displays the spatial distribution of the contributions of the above four surface parameters in study areas A, B and C. From a visual comparison, the contributions of each of the four parameters varied both spatially and temporally and were different from each other. Compared with those of the other three parameters, the spatial variation of the contribution of r_s (Figure 5(d)) was more similar to that of *dLST* (Figure 5e). The contribution of r_s (1.26 ± 1.39 K on average) had the greatest impact on the final *dLST* results, especially for the high-value ranges of *dLST*, whereas the lowest impacts and narrower variation ranges were found for the contributions of R_n (0.33 ± 0.21 K on average, Figure 5a). In addition, the contributions of f_c (Figure 5b) were negatively correlated with *dLST*, while the contributions of r_s were positively correlated with *dLST*, which means that in the process of LST downscaling from a coarse resolution to a fine resolution, an increase in f_c or decrease in r_s could lead to a decrease in LST and vice versa. However, no simple correlation relationship was found between the contribution of R_n or r_a and *dLST*.

Table 2. Statistical metrics of the validation of the downscaled 90 m LST by the DTsEB, TsHARP, LMS, and GWR methods from the 6 scenes of 990 m MODIS LST over three study areas with the corresponding 90 m ASTER LST. RMSE is the root mean square error, MAE is the mean absolute error, and NRMSE is the RMSE normalized by the standard deviation (SD) of the reference 90 m ASTER LST.

Study area & Date		Methods															
		DTsEB				TsHARP				LMS				GWR			
		RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE
A	Apr. 24, 2006	1.84	1.37	-0.89	0.70	2.37	1.72	-1.51	0.90	2.34	1.71	-1.51	0.89	2.36	1.69	-1.56	0.90
A	Oct. 3, 2010	1.46	1.20	-1.04	1.22	1.57	1.31	-1.24	1.31	1.56	1.30	-1.22	1.30	1.57	1.30	-1.23	1.32
B	Apr. 15, 2005	3.02	2.53	-2.06	1.29	3.21	2.71	-2.43	1.37	3.38	2.88	-2.59	1.44	3.25	2.74	-2.41	1.38
B	Jul. 29, 2005	2.03	1.26	-0.74	0.79	2.46	1.54	-1.38	0.96	2.32	1.46	-1.27	0.91	2.44	1.53	-1.39	0.95
C	Sept. 2, 2014	1.92	1.33	-0.23	0.91	2.10	1.52	-1.19	1.00	1.96	1.42	-1.04	0.93	2.00	1.45	-1.15	0.95
C	Aug. 20, 2015	2.06	1.54	-1.05	1.04	2.76	2.25	-2.22	1.39	2.52	2.05	-2.00	1.27	2.64	2.22	-2.20	1.33
Overall		2.01	1.54	-1.00	0.99	2.41	1.84	-1.66	1.16	2.35	1.80	-1.61	1.12	2.38	1.82	-1.66	1.14

Table 3. Same as Table 2, but for the 990 m aggregated ASTER data.

Study area & Date		Methods															
		DTsEB				TsHARP				LMS				GWR			
		RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE	RMSE (K)	MAE (K)	BIAS (K)	NRMSE
A	Apr. 24, 2006	1.54	1.25	0.30	0.59	1.74	1.35	-0.06	0.66	1.69	1.28	0.05	0.65	1.76	1.34	0.05	0.67
A	Oct. 3, 2010	0.95	0.72	0.12	0.79	1.17	0.90	0.41	0.98	1.10	0.86	0.03	0.92	1.12	0.87	0.03	0.94
B	Apr. 15, 2005	1.85	1.45	0.14	0.79	2.32	1.73	-0.03	0.99	2.03	1.58	0.03	0.86	2.36	1.79	0.02	1.00
B	Jul. 29, 2005	1.46	0.99	0.06	0.57	1.97	1.35	-0.24	0.77	2.12	1.30	-0.03	0.83	1.67	1.11	-0.08	0.65
C	Sept. 2, 2014	1.45	1.03	0.25	0.69	1.53	1.13	-0.09	0.73	1.52	1.10	0.01	0.72	1.59	1.16	-0.01	0.75
C	Aug. 20, 2015	1.29	0.92	0.24	0.65	1.51	1.11	-0.08	0.76	1.31	0.89	0.01	0.66	1.32	0.90	0.01	0.66
Overall		1.42	1.06	0.19	0.68	1.71	1.26	-0.02	0.82	1.63	1.17	0.03	0.77	1.64	1.20	0.003	0.78

Table 4. The correlation coefficient (R) between 90 m downscaled LST by the DTsEB, TsHARP, LMS, and GWR methods and the reference 90 m ASTER LST over the three study areas.

Study area & Date		Methods							
		DTsEB		TsHARP		LMS		GWR	
		MODIS LST	Aggregated ASTER LST	MODIS LST	Aggregated ASTER LST	MODIS LST	Aggregated ASTER LST	MODIS LST	Aggregated ASTER LST
A	Apr. 24, 2006	0.79	0.83	0.76	0.84	0.77	0.85	0.83	0.85
A	Oct. 3, 2010	0.55	0.63	0.59	0.61	0.58	0.61	0.58	0.57
B	Apr. 15, 2005	0.38	0.63	0.46	0.56	0.39	0.60	0.40	0.47
B	Jul. 29, 2005	0.66	0.82	0.66	0.80	0.67	0.80	0.69	0.80
C	Sept. 2, 2014	0.61	0.75	0.59	0.76	0.63	0.76	0.68	0.74
C	Aug. 20, 2015	0.59	0.78	0.58	0.72	0.64	0.80	0.73	0.79
Overall		0.60	0.74	0.61	0.72	0.61	0.74	0.65	0.70

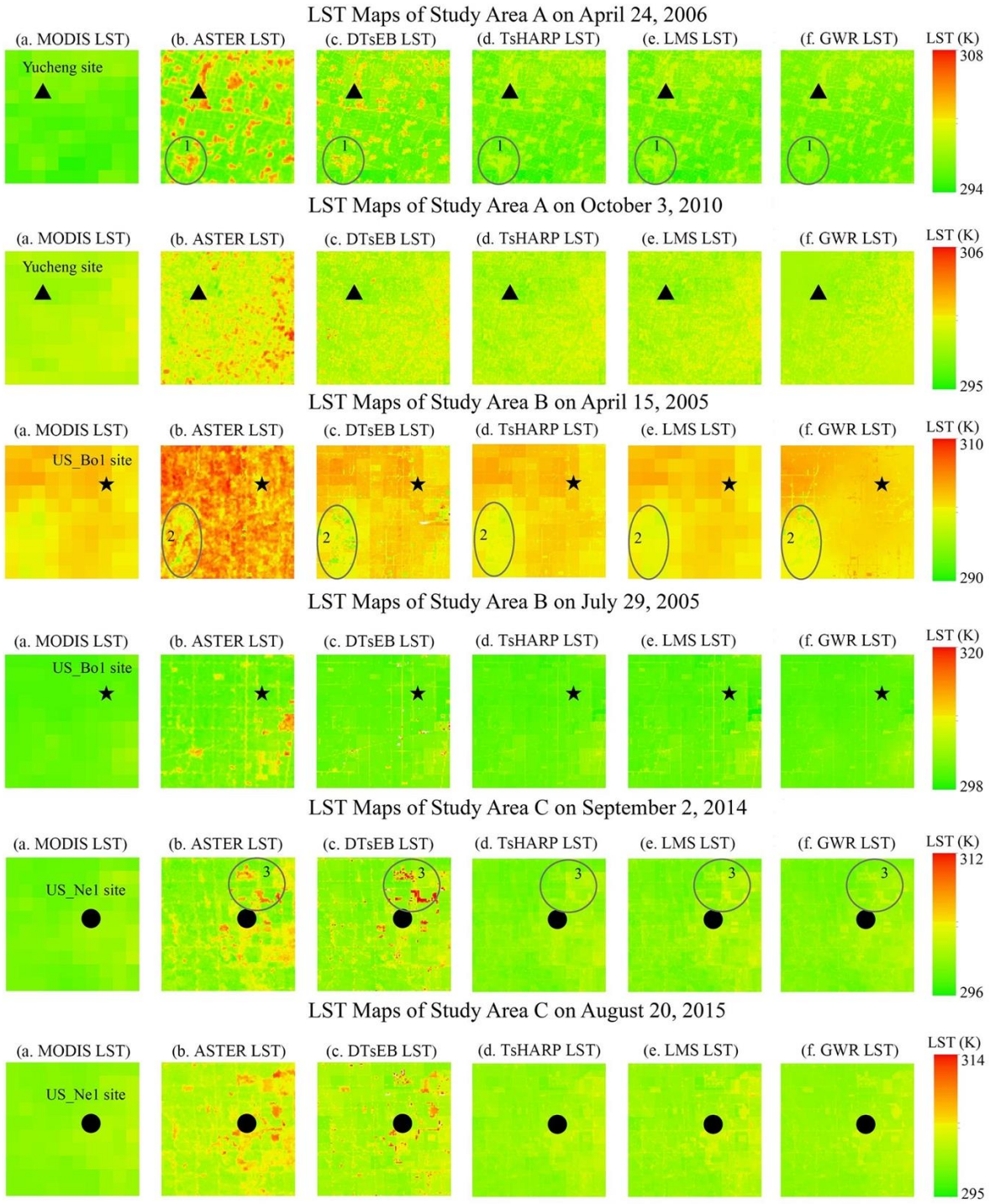

Figure 3. Spatial distribution of the (a) 990 m MODIS LST, (b) 90 m ASTER LST, (c) 90 m LST downscaled by the DTsEB method, (d) 90 m LST downscaled by the TsHARP method, (e) 90 m LST downscaled by the LMS method, and (f) 90 m LST downscaled by the GWR method for study areas A, B and C. Circles 1, 2 and 3 (in study area A on April 24, 2006, in study area B on April 15, 2005, and in study area C on September 2, 2014, respectively) are typical areas highlighted for comparison.

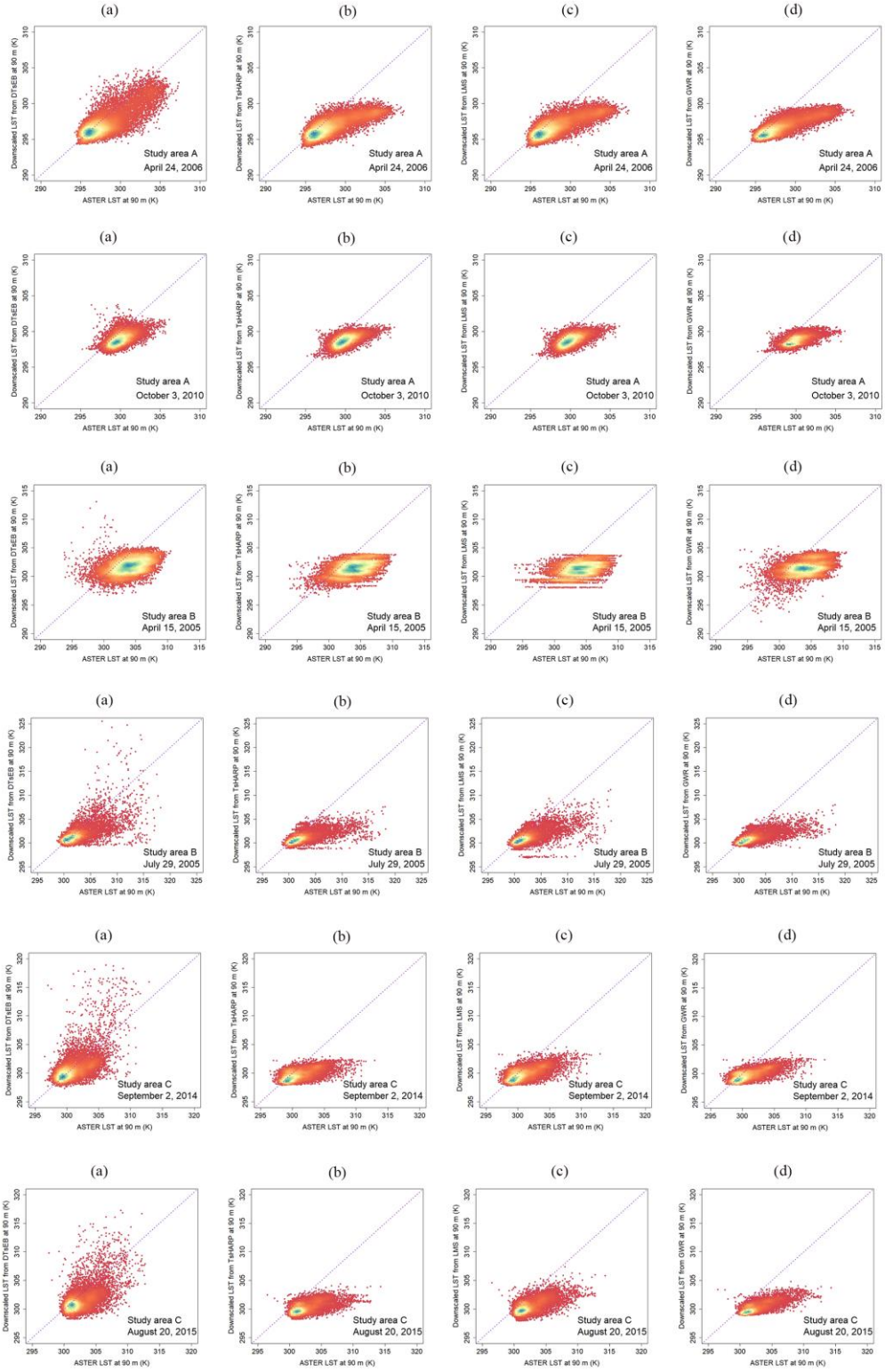

Figure 4. Comparisons of the 90 m LST downscaled from 990 m MODIS products using the (a) DTsEB (left panel), (b) TsHARP, (c) LMS, and (d) GWR methods with the ASTER LST for the three study areas.

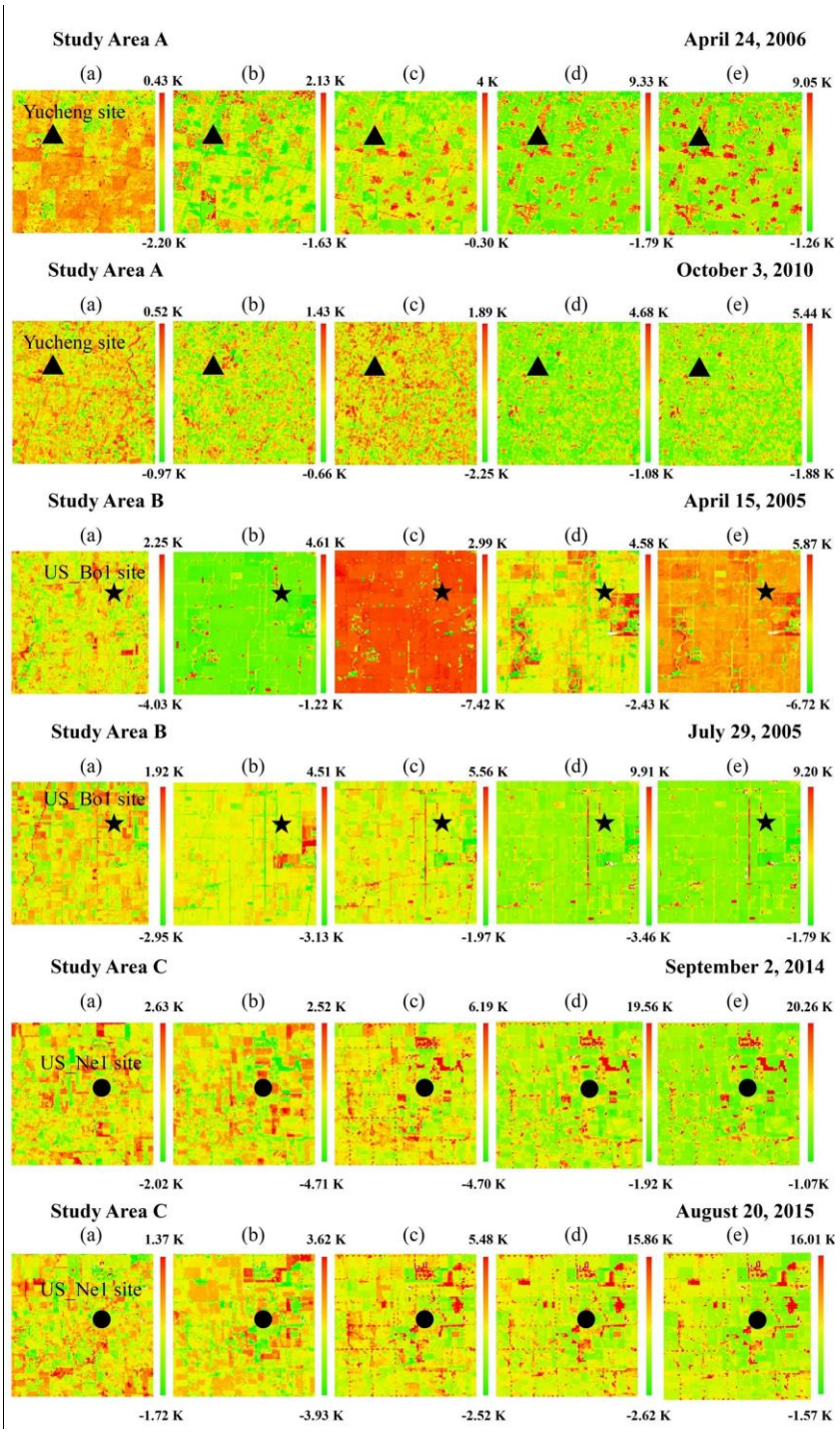

Figure 5. Spatial distribution of the contributions of the four surface parameters to the final *dLST* results in study areas A, B and C by using the MODIS datasets: (a) contribution of the surface net radiation, $\frac{\partial LST}{\partial R_n} dR_n$; (b) contribution of fraction of the vegetation, $\frac{\partial LST}{\partial f_c} df_c$; (c) contribution of the aerodynamic resistance, $\frac{\partial LST}{\partial r_a} dr_a$; (d) contribution of the surface resistance, $\frac{\partial LST}{\partial r_s} dr_s$; and (e) the estimated *dLST*.

4.2 Application to aggregated ASTER datasets

Similar to the application to the MODIS datasets, the DTsEB, TsHARP, LMS, and GWR methods were also applied to downscale the 990 m resolution aggregated ASTER LSTs to 90 m resolution. Overall, the aggregated ASTER LSTs (Figure 6a) were higher than the MODIS LSTs (Figure 3a) in all the study areas, i.e., A, B and C. Especially in the high-value LST ranges, the aggregated coarse-resolution ASTER LSTs presented a much broader value distribution. The maximum value of the aggregated ASTER LSTs (Table 1) was approximately 4 K higher than that of the MODIS LSTs on average. Furthermore, from a visual comparison, the pixel-to-pixel LST variations were also observed to be larger in the aggregated coarse-resolution LSTs than in the MODIS LSTs. Compared with the reference ASTER LSTs without aggregation, the aggregated ASTER LSTs roughly exhibited the expected similar spatial distributions in all the study areas, i.e., A, B and C. The differences between the mean aggregated ASTER LSTs and the mean fine-resolution ASTER LSTs over the three study areas were less than 0.1 K. The notable differences between the coarse-resolution LSTs and fine-resolution reference LSTs were largely reduced with the use of aggregation datasets.

The spatial patterns of the 90 m LSTs downscaled from the aggregation datasets using the DTsEB, TsHARP, LMS, and GWR methods over the three study areas are displayed in Figure 6 and a scatterplot of the comparison between the downscaled LSTs and the reference fine-resolution ASTER LSTs is presented in Figure 7. Overall, although an overestimation of low LST extremes and an underestimation of high LST extremes were present, the DTsEB, TsHARP, LMS, and GWR methods could all effectively reconstruct subpixel spatial variations within coarse-resolution pixels. The spatial distribution and texture characteristics of the three downscaled LST results were similar and basically consistent with those of the 90 m reference ASTER LSTs. The accuracy of LST downscaling results obtained by the DTsEB method was higher than that obtained by the TsHARP, LMS, and GWR methods, while the LMS method slightly outperformed the TsHARP and GWR method. The RMSE decreased by 0.08~0.51 K (~17% on average) from 1.17~2.32 K for the TsHARP method, by 0.02~0.66 K (~13% on average) from 1.10~2.12 K for the LMS method, and by 0.03~0.51 K (~13% on average) from 1.12~2.36 K for the GWR method, to 0.95~1.85 K for the DTsEB method (Table 3). As expected, the normalized RMSE of 0.57~0.79 for the DTsEB method was lower than that for the TsHARP method (0.66~0.99), the LMS method (0.65~0.92) and the GWR method (0.65~1.00). Except for the results on August 20, 2015, over

study area C, where the MAE of the LMS method and GWR method was negligibly lower (0.03 K and 0.02 K, respectively) than that of the DTsEB method, the MAE decreased by 0.10~0.36 K (16% on average) from 0.90~1.73 K for the TsHARP method, by 0.03~0.31 K (9% on average) from 0.86~1.58 K for the LMS method, and by 0.09~0.34 K (12% on average) for the GWR method, to 0.72~1.45 K for the DTsEB method. The correlation coefficient (Table 4), which varied between 0.63 and 0.83, for the DTsEB method was comparable to that for the TsHARP method (between 0.56 and 0.84), the LMS method (between 0.60 and 0.85), and the GWR method (between 0.47 and 0.85), while the bias for the DTsEB method was slightly higher than that for the latter three methods. Furthermore, for all the DTsEB, TsHARP, LMS, and GWR methods, better LST downscaling results were obtained in study areas A and C. The RMSE values obtained in the evaluation of the 90 m LSTs downscaled by the DTsEB method on October 3, 2010 in study area A, for example, were lower than 1 K. However, in the LST downscaling results on April 15, 2005 in study area B, a larger bias was observed in the low-value ranges of LSTs for the TsHARP, LMS, and GWR methods (Figure 7).

Compared with the application to MODIS datasets, the use of aggregated ASTER LST datasets for the DTsEB, TsHARP, LMS, and GWR methods all resulted in improved accuracy of the downscaled results for study areas A, B and C. When using the aggregated ASTER datasets instead of the MODIS datasets, the mean RMSE (MAE) values of the DTsEB, TsHARP, LMS, and GWR downscaled results decreased by 0.59 K (0.48 K), 0.70 K (0.58 K), 0.72 K (0.63 K), and 0.74 K (0.62 K), respectively. For study area B on April 15, 2005, the RMSE and MAE values obtained by using the DTsEB decreased by 39% and 43%, respectively, while the TsHARP yielded a 28% decrease in RMSE and a 36% decrease in MAE, the LMS produced a 40% decrease in RMSE and a 45% decrease in MAE, and the GWR yielded a 27% decrease in RMSE and a 35% decrease in MAE compared to the application to the MODIS LSTs. Furthermore, the obvious underestimation in the high-value ranges of LSTs by using the MODIS datasets was effectively improved in the application to the aggregated ASTER LST datasets (see the scatter plot distribution in Figures 4 and 7).

Figure 8 displays the spatial patterns of the contributions of the four scaling factors (e.g., R_n, f_c, r_a, and r_s) to the final *dLST* when the DTsEB method was applied to the aggregated ASTER LST downscaling. Similar to the findings of the application to the MODIS datasets, the results showed that the contributions of the four parameters were different from each other and varied

both spatially and temporally in the application to the aggregated ASTER LSTs; the contributions of all four scaling factors varied from negative to positive. The contribution of r_s was most affected by environmental variables and had a broader value distribution than the other three scaling factors whereas the contribution of R_n had the narrowest value distribution. For example, in study area A on April 24, 2006 and October 3, 2010, the contribution of R_n to the final *dLST* was less than ± 1 K while in study area B on July 29, 2005 and in study area C on September 2, 2015, the greatest contributions of r_s to the final *dLST* were larger than 18 K and 14 K, respectively. Furthermore, the contributions of r_s (Figure 8c) were positively correlated with the *dLST* (Figure 8e) and had the greatest impact on the final *dLST* results (0.72 ± 0.84 K on average), which was similar to the findings in the application to the MODIS LSTs. Meanwhile, the contributions of f_c (Figure 8b) were negatively correlated with *dLST*, which is consistent with the negative correlation relationship between surface vegetation and LSTs. The lowest impact on the final *dLST* was found in the contributions of R_n (0.19 ± 0.18 K on average, Figure 8a). Compared with the application to the MODIS datasets, the spatial distribution of the contributions of the four scaling factors was smoother, and the contributions of r_a and r_s to the final *dLST* in high-value ranges were larger.

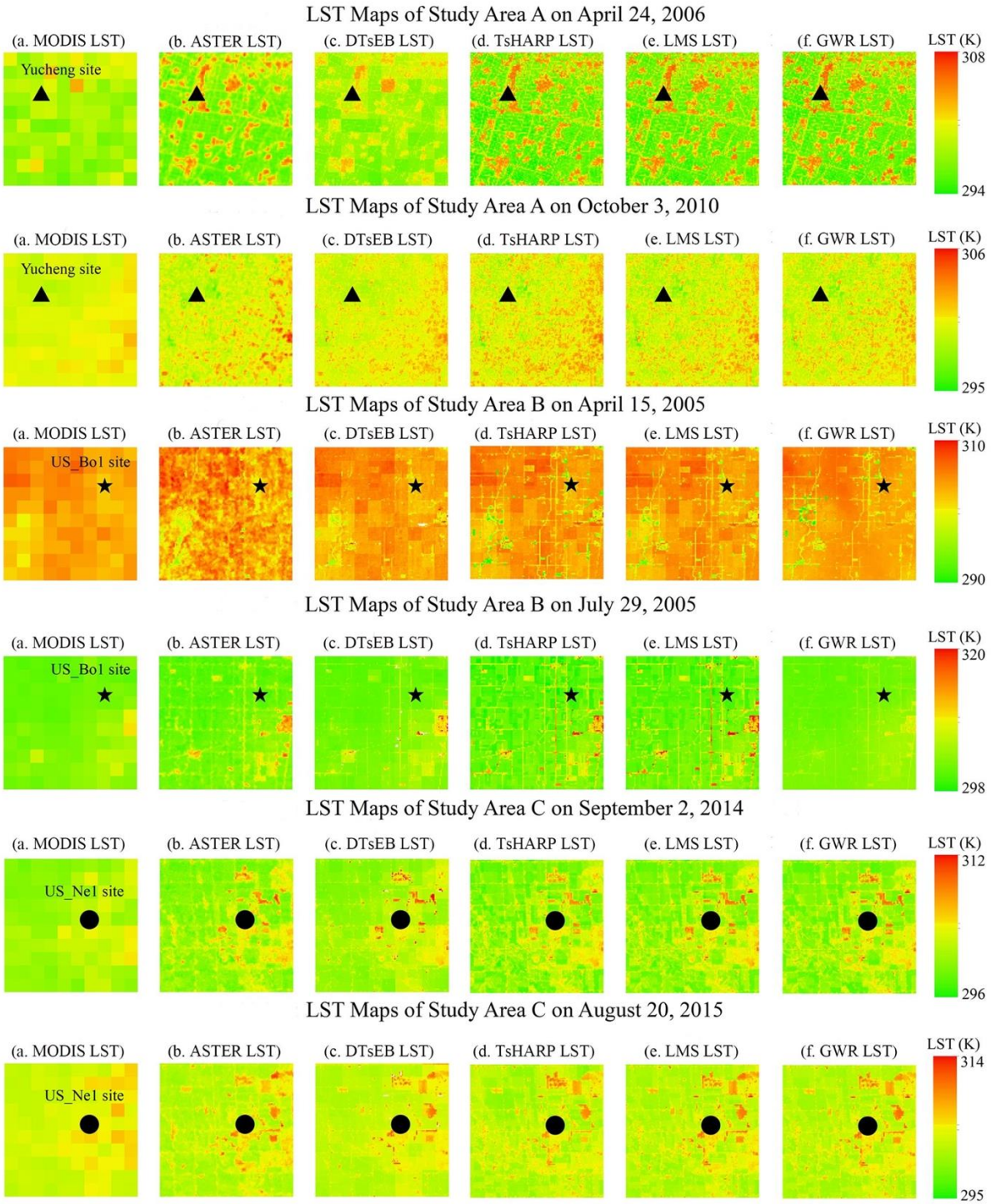

Figure 6. Spatial distributions of the (a) 990 m aggregated ASTER LSTs, (b) 90 m reference ASTER LSTs, (c) 90 m LSTs downscaled by the DTsEB method, (d) 90 m LSTs downscaled by the TsHARP method, (e) 90 m LSTs downscaled by the LMS method, and (f) 90 m LSTs downscaled by the GWR method for study areas A, B and C.

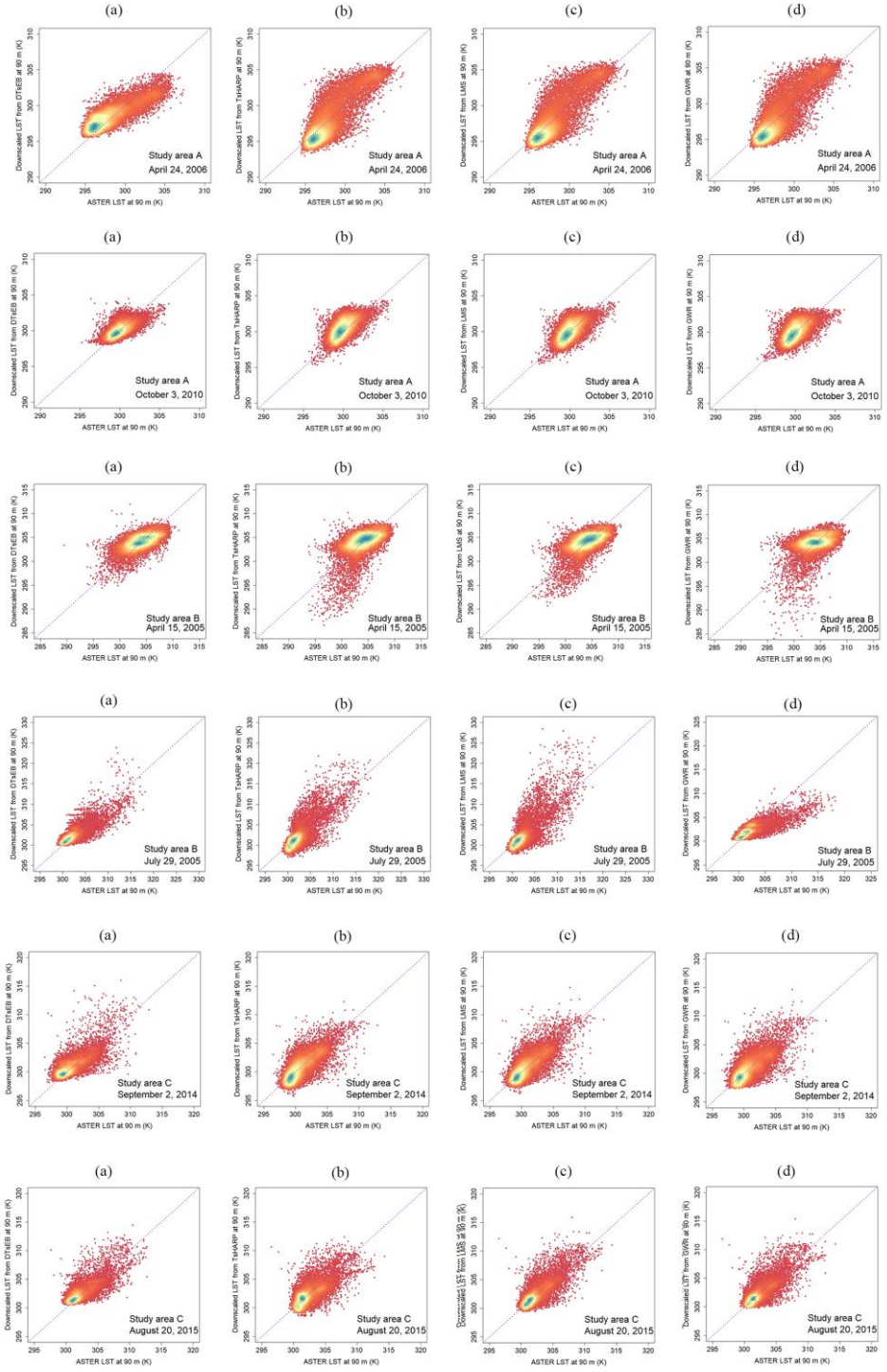

Figure 7. Comparisons of the 90 m LSTs downscaled from the 990 m aggregated ASTER LSTs using the DTsEB (a), TsHARP (b), LMS (c), and GWR (d) methods with ASTER LSTs for the three study areas.

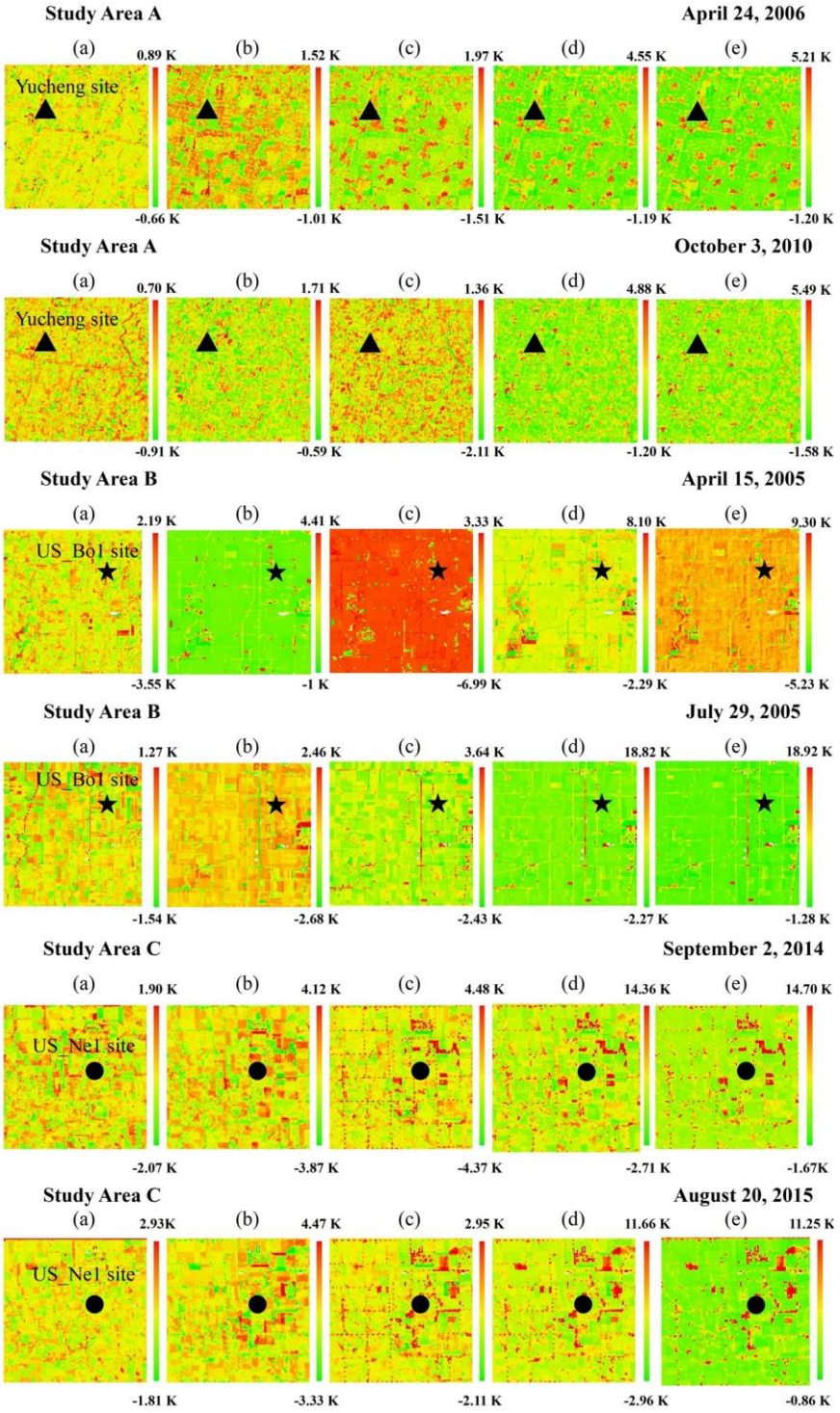

Figure 8. Spatial distribution of the contributions of four surface parameters to the final *dLST* results in study areas A, B, and C determined by using the aggregated ASTER datasets: (a) contribution of the surface net radiation, $\frac{\partial LST}{\partial R_n} dR_n$; (b) contribution of the fraction of vegetation, $\frac{\partial LST}{\partial f_c} df_c$; (c) contribution of the aerodynamic resistance, $\frac{\partial LST}{\partial r_a} dr_a$; (d) contribution of the surface resistance, $\frac{\partial LST}{\partial r_s} dr_s$; and (e) the estimated *dLST*.

4.3 Discussion

The improvements in LST downscaling methods made in recent years mainly include the following two aspects: 1) selecting more appropriate scaling factors and 2) establishing more accurate relationships between LSTs and scaling factors. Regrettably, many of the scaling factors chosen in previous works simply remedied the regressed negative correlation relationships between NDVI and LSTs *[Merlin et al., 2010; Yang et al., 2011; Bonafoni, 2016; Duan & Li, 2016]*, which is somewhat arbitrary and site-specific. Furthermore, most of the established relationships between LSTs and scaling factors were from statistical regressions *[Agam et al., 2007; Bindhu et al., 2013]*, which limited the applicability and robustness of the LST downscaling algorithm and even resulted in different LSTs for the same pixel under different regression equations and sizes of areas of interest. Instead of selecting the scaling factors subjectively and using statistical regression relationships with no explicit physical mechanism, the DTsEB method proposed in this paper improves the downscaling of coarse-resolution LSTs by proposing analytical equations. These equations make use of surface energy balance constraints to provide a physically intuitive mechanism for combining the thermal infrared spectrum data (coarse resolution) with the VNIR and SWIR spectrum data (fine resolution).

The performance improvements in the DTsEB method against the TsHARP, LMS, and GWR methods varied among different scenes and different underlying surface conditions. The DTsEB downscaled results were better than the TsHARP, LMS, and GWR results, especially in high LST ranges. The underestimations of TsHARP, LMS, and GWR over high LST value pixels (see the scatter plots in Figures 4 and 7) indicate the limitations in the extension of regression equations constructed with narrower NDVI ranges at coarse resolution to applications at wider NDVI ranges at fine resolution. Taking the concurrent 90 m ASTER LST product as the reference LST, the TsHARP method yielded average RMSE (MAE) values of 2.41 K (1.84 K) and 1.71 K (1.26 K) in the downscaling of the 990 m MODIS LSTs and the aggregated ASTER LSTs to 90 m, respectively. By contrast, for the LMS method the average RMSE (MAE) values were 2.35 K (1.80 K) and 1.63 K (1.17 K), respectively, and for the GWR method the average RMSE (MAE) values were 2.38 K (1.82 K) and 1.64 K (1.20 K), respectively. These LST downscaling accuracies of the TsHARP, LMS, and GWR methods are comparable to those achieved in previous studies. For example, *Hutengs & Vohland [2016]* applied TsHARP to downscale 960 m aggregated ETM+ LST to 240 m resolution and obtained an average RMSE of 1.48 K (referenced to 240 m ETM+

LST data). In the downscaling of 990 m MODIS LSTs to 90 m resolution, the TsHARP method achieved RMSEs of 3.62 K and 2.16 K (referenced to 90 m ASTER LST data) for two different study areas in *Wu & Li*'s work *[2019]*. The LMS method in *Mukherjee et al.*'s [*2014*] work produced an average RMSE of 1.43 K in the downscaling of 1000 m MODIS LSTs to 250 m (referenced to 250 m TM LST data) and in *Bisquert et al.*'s [*2016*] work generated average RMSE values of 1.80 K and 2.10 K in downscaling 960 m MODIS LSTs and aggregated ETM+ LSTs to 60 m (referenced to 60 m ETM+ LST data), respectively. *Duan & Li* [*2016*] introduced the GWR to downscale the 990 m MODIS LST to 90 m (referenced to 90 m ASTER LST data), and obtained an average RMSE (MAE) of 3.1 K (2.3 K). Compared to the TsHARP, LMS, and GWR, the DTsEB method in this study improved LST downscaling, with average RMSEs (MAEs) of 2.01 K (1.54 K) and 1.42 K (1.06 K) in the application to the 990 m MODIS datasets and aggregated ASTER datasets, respectively, indicating the effectiveness of this new proposed method. The average RMSE decrease achieved by using the DTsEB (17% and 17% compared to the TsHARP, 14% and 13% compared to the LMS, and 16% and 13% compared to the GWR in application to the MODIS datasets and aggregated ASTER datasets, respectively) in the LST downscaling compared favorably to those achieved by the Extended-RFD method (13% to 26% relative to the TsHARP) in *Hutengs & Vohland [2016]* and the regression tree-based method (averages of 20% and 25% relative to the TsHARP for an irrigated agricultural site and heterogeneous naturally vegetated area, respectively) in *Gao et al. [2012]*. Furthermore, compared to the downscaling methods suggested by *Merlin et al. [2010]*, *Chen et al. [2014]* and *Duan & Li [2016]*, the DTsEB method is also observed to produce similar or better LST accuracy in downscaling kilometer-resolution LSTs to fine resolution. In the work of *Merlin et al. [2010]*, broadband albedo was introduced into the TsHARP method to distinguish photosynthetically and nonphotosynthetically active vegetation and finally achieved average RMSEs of 3.81 K and 2.78 K in downscaling 1 km resolution MODIS LSTs and aggregated ASTER LSTs to 100 m resolution, respectively. *Chen et al. [2014]* combined the TsHARP method with thin-plate spline interpolation to downscale 1-km-resolution MODIS LSTs to 250 m resolution and obtained an RMSE of 2.38 K. *Duan & Li [2016]* introduced geographically weighted regression to the TsHARP method and obtained an average RMSE of 2.7 K in downscaling 990 m resolution MODIS LSTs to 90 m resolution.

The improved LST downscaling results obtained in the application to aggregated ASTER datasets compared to the application to MODIS datasets for the DTsEB, TsHARP, LMS, and GWR

methods mainly resulted from the smaller differences between the coarse-resolution aggregated ASTER LSTs and fine-resolution reference ASTER LSTs than between the coarse-resolution MODIS LSTs and fine-resolution reference ASTER LSTs, which is consistent with the previous findings that differences between coarse and reference fine resolution LSTs could directly affect evaluations of downscaling results *[Agam et al., 2007; Merlin et al., 2010]*. *Yang et al. [2011]* also found that downscaled LSTs often have a relatively high accuracy by using resampling and aggregation methods. Different from the intercalibration of the coarse- and fine-resolution LSTs and surface parameters (such as NDVI) that were obtained from different sensors in the works of *Bindhu et al. [2013]*, the datasets from the MODIS and ASTER sensors onboard the same satellite platform, which avoided errors caused by different satellite altitudes and overpass times, were used directly without extra processing in this study. In this aspect, reasonably enhancing the spatial details of original coarse-resolution MODIS LSTs (Real data) is crucial to LST downscaling methods. Given the relatively large differences between the LSTs from these two sensors, the better performance of the DTsEB method compared with the TsHAPR, LMS, and GWR methods, especially in the high-value ranges of LSTs, highlights the high robustness, generality, and accuracy of DTsEB. Nevertheless, both the coarse-resolution remotely sensed LST products and the reference fine-resolution LST products have an intrinsic bias, which is difficult to exclude in LST downscaling.

Another advantage of the DTsEB method is its ability to properly quantify the contributions of each scaling factor (e.g., surface net radiation, fraction of vegetation, aerodynamic resistance and surface resistance) within a physical framework. Although the values of these scaling factors are likely to vary with the spatial resolution of the VNIR/SWIR images, the physical relationship remains inviable, whereas the regression-based TsHARP, LMS, and GWR methods and others only attribute the subpixel spatial variations of LST to one or more vegetation indices and topographic variables (e.g., NDVI, NDWI, NDBI, EBBI, BI, TVDI, and DEM, see Introduction Section), and their regression relationships are different from one another. The test results in this study revealed that the surface resistance and aerodynamic resistance were, overall, the largest and second largest factors, respectively, which contributed to the subpixel spatial variations of coarse-resolution land surface temperature for the whole spatial domain in the three study areas. An exception, for which the largest contribution was from the aerodynamic resistance in the downscaling of MODIS surface temperatures, occurred and was possibly due to the intrinsic

difference in the surface temperature and reflectance measurements between the MODIS and ASTER sensors. Different from the DTsEB method, the regression-based downscaling techniques (e.g., TsHARP, LMS, GWR and other methods) were flawed in their attribution of the contributions and different attribution results could be obtained from these techniques with different independent variables (e.g., scaling factors), which clearly does not make sense.

Downscaled LSTs are often accompanied by the notorious "boxy effect" *[Agam et al., 2007, 2008; Duan & Li, 2016; Bindhu et al., 2013]*, which results from the addition of the constant residuals obtained at coarse resolutions. This addition is necessary and can help improve LST downscaling when there are LST differences at coarse resolution between the values calculated by the constructed relationship with the scaling factors and the values extracted from the remotely sensed image to be downscaled. The boxy effects become more pronounced when the residual errors are larger and disappear if LST downscaling is performed without adding the constant residual or the constructed relationship can perfectly (no residual error) represent the remotely sensed LSTs at coarse resolution. Compared to the TsHARP, LMS and GWR methods, the DTsEB method makes use of the dLST that represents the differences between LSTs at fine and coarse resolutions and is expressed as a function of the differences between surface parameters (e.g., dR_n, df_c, dr_a, and dr_s). The addition of residual field is actually not applied in the DTsEB method. Therefore, the DTsEB method can more effectively reduce the "boxy effect" and thus improve LST downscaling compared to the TsHARP, LMS, and GWR methods because of the higher accuracy of the physical LST equation relative to the regression equation.

In brief, the TsHARP, LMS, GWR and other regression-based methods are simple in model structure, do not require auxiliary near-surface data as input but are deficient in their poor spatiotemporal extensibility and in quantifying the contributions of influencing factors (namely, attribution analysis). In contrast, the DTsEB method, developed by theoretical derivations of surface energy balance and Penman-Monteith equation under the assumption of negligible spatial variations in atmospheric parameters over the subpixels within a coarse pixel, has the advantages of a solid physical foundation, the capability to separate the contributions of the influencing factors, and LST downscaling results with a high accuracy. The main limitation of the DTsEB method lies in the requirements for near-surface meteorological data (e.g., incoming solar radiation, air temperature, VPD, and wind speed), which may introduce a certain degree of uncertainty in LST downscaling, especially when the DTsEB method is applied regionally or globally, because pixel-

by-pixel meteorological data (e.g., sourced from reanalysis data) should be introduced to consider the spatial variation in near-surface meteorology under such conditions. This data requirement does not add much computational cost. For instance, in-situ meteorological data, such as the FLUXNET and AMERIFLUX datasets, can be used in small-scale study areas (e.g. the study areas (Figure 1) with spatial dimensions of 9.9 km by 9.9 km). As for regional or global study areas, interpolated meteorological data and reanalysis data (such as ERA5 datasets) can be used. Moreover, the uncertainty of the parameterization in the DTsEB algorithm also introduces biases in LST downscaling to some extent. For example, the scatters in the downscaling of high LSTs (primarily over built-up lands) in this study likely resulted from the uncertainty in the determination of roughness height (influencing aerodynamic resistance) and soil heat flux by following the general parameterizations over vegetated surfaces (e.g., cropland, grassland, forestland), indicating that improved parameterization of the DTsEB method for these two parameters is required over built-up lands (beyond the scope of this study). In particular, we did not distinguish the parameterization of surface resistance between crop and built-up lands but applied the same equation as shown in Appendix D to parameterize surface resistance for all land cover types, primarily because the focus of our study was not on the parameterization but on the development of DTsEB downscaling method. Parameterizing surface resistance differently are strongly recommended over cropland, built-up land and other land cover types. In addition, due to the complexities of LST downscaling that come from the uncertainty/error of coarse-resolution and fine-resolution LSTs and VNIR/SWIR reflectance products, downscaling algorithm, parameterization, and inputs, none of the three methods could obtain a normalized RMSE < 0.5 in this study, although these methods have reported RMSE values of similar magnitude to those from previous studies. In short, despite the great progress made in the past for LST downscaling, there remains a long way to go.

5. Summary and Conclusions

A physical LST downscaling method, DTsEB, has been developed to downscale coarse-resolution LST data to a fine resolution. By theoretical derivations of the surface energy balance equation and Penman-Monteith equation, analytical equations for combining thermal infrared data with visible and near-infrared data have been constructed in the newly proposed LST downscaling method. The differences in surface net radiation, fractional vegetation cover, aerodynamic resistance and surface resistance between coarse and fine resolutions are first calculated, and fine-

resolution LSTs can then be obtained by converting the differences between the LSTs at the two resolutions to the differences between these surface parameters. The surface energy balance constraint in the DTsEB method provides a robust and physical connection between the scaling factors and LSTs and thus avoids the subjective selection of scaling factors and the use of statistical regression relationships.

Because of the comprehensive consideration of various surface parameters related to LSTs, the DTsEB method can effectively reconstruct subpixel spatial variations within coarse-resolution pixels and achieve better downscaling accuracy than the widely adopted TsHARP, LMS, and GWR methods, when tested on 990 m MODIS and aggregated LST products collected between 2005 and 2015 over three 9.9 km by 9.9 km cropland (mixed by grass, tree, and built-up land) study areas. The average RMSE (MAE) values in DTsEB decreased by 17% (16%) relative to the TsHARP method, 14% (14%) relative to the LMS method, and 16% (15%) relative to the GWR method for application to 6 scenes of MODIS datasets and by 17% (16%) relative to the TsHARP method, 13% (9%) relative to the LMS method, and 13% (12%) relative to the GWR method for application to 6 scenes of aggregated ASTER datasets.

In summary, the DTsEB method has great potential in LST downscaling over various land cover types and satellite sensor data as long as the parameters are properly estimated, because 1) the solid physical foundation makes it robust and highly accurate and 2) the physical association between scaling factors and the LSTs can quantitatively separate the specific contributions of different scaling factors to the LST downscaling results. In the context that most existing LST downscaling methods are based on statistical regression, the physical DTsEB method proposed in this study is instructive and worthwhile. When other high-resolution satellite sensor (e.g. Landsat TM, ETM+, OLI) LST data are used to test the applicability of this new method, one may only perform a simulation of downscaling coarse-resolution aggregated LST to high-resolution LST (as in case 2 shown in Section 3.2) because no coarse-resolution satellite sensor LST concurrent with the high-resolution satellite sensor LST is available. To allow more general conclusions to be made, further work is recommended to evaluate the DTsEB method and the regression-based LST downscaling methods in more regions of the world that are characterized by a wider range of climates and land cover conditions.

Acknowledgments

We are grateful for the MODIS and ASTER data resources provided by the National Aeronautics and Space Administration (https://search.earthdata.nasa.gov/search/), for the meteorological data resources provided by Yucheng National Agriculture Ecosystem Observation and Research Station (http://yca.cern.ac.cn/) and to the U.S. Department of Energy's Office of Science (https://ameriflux.lbl.gov/). This work was supported by the National Natural Science Foundation of China under Grants 41922009, 42071332, 41971319, and 41921001, the National Key R&D Program of China under Grants 2018YFA0605401, 2018YFB050480304 and 2018YFB050480404, the Strategic Priority Research Program of Chinese Academy of Sciences under Grant XDA19040403, the Bureau of International Co-operation Chinese Academy of Sciences under Grant 181811KYSB20160040, and the Dragon 4 ESA-MOST Cooperation programme under Grant 32426_1.

Appendix

A. TsHARP method

For a comparative analysis of the LST downscaling performance of the DTsEB method, the widely used vegetation-based regression method, TsHARP (more specifically, the TsHARP version, named TsHARPfcS, which was recommended by Agam *et al*. *[2007]*), was applied in this study. The TsHARP method, a refinement of the disaggregation procedure for radiometric surface temperatures (DisTrad, proposed by Kustas *et al*. *[2003]*), assumes that the relationship between LSTs and NDVI-based transformed variables is scale invariant. A linear regression between LSTs and the NDVI-based transformed variable is first performed at coarse resolution, as follows:

$$f\left(NDVI_{CR}\right) = a_0 - a_1(1 - NDVI_{CR})^{0.625} \tag{A1}$$

where the subscript CR represents the coarse resolution.

Subsequently, the divergence (ΔLST) between the regressed LSTs and the source LSTs, which comes from the spatial variability in LSTs that is driven by factors other than the vegetation cover fraction at coarse resolution, can be calculated, as follows:

$$\Delta LST = LST_{CR} - f\left(NDVI_{CR}\right) \tag{A2}$$

This residual field is finally applied to derive the downscaled fine-resolution LSTs (LST_{TsHARP}), as follows:

$$\begin{aligned} LST_{TsHARP} &= f\left(NDVI_{FR}\right) + \Delta LST \\ &= a_0 - a_1(1 - NDVI_{FR})^{0.625} + \Delta LST \end{aligned} \tag{A3}$$

where the subscript FR represents the fine resolution.

B. LMS method

According to the work of Mukherjee *et al*. [*2014*], least median square regression downscaling (LMS), which is less sensitive to outliers than the ordinary least square regression algorithm (used in TsHARP method), could achieve a better accuracy in LST downscaling.

In the ordinary least square regression, the regression parameters slope and intercept are estimated by minimizing the sum of square residuals, as follows:

$$MinSSR = \sum_{i=1}^{n} (LST_i - f(NDVI_i))^2 \tag{B1}$$

In the LMS, the parameter slope and intercept are calculated to yield the least median of the square residuals, as follows:

$$MinMedSR = Median \begin{Bmatrix} (LST_1 - f(NDVI_1)), (LST_2 - f(NDVI_2)), \\ ..., (LST_n - f(NDVI_n)), \end{Bmatrix} \tag{B2}$$

The least median square regression between LST and NDVI is first performed at coarse resolution, and the divergence (ΔLST) between the regressed LSTs and the source LSTs can be subsequently calculated. Finally, this residual field ΔLST is added at fine resolution to obtain the fine-resolution LST.

C. GWR method

Compared with traditional regression method, geographically weighted regression (GWR) can fully consider the geographic similarity relationship between the dependent variables and the independent variables. According to the work of Duan *et al*. [*2016*], a nonstationary relationship at coarse-resolution is first established, which can be expressed as:

$$LST_i^{CR} = a_0^{CR}(\mu_i, \nu_i) + a_1^{CR}(\mu_i, \nu_i)NDVI_i^{CR} + a_2^{CR}(\mu_i, \nu_i)DEM_i^{CR} + \Delta_i^{CR} \quad \text{(C1)}$$

where the superscript *CR* represents the coarse resolution, the $\alpha_0^{CR}(\mu_i, \nu_i)$, $\alpha_1^{CR}(\mu_i, \nu_i)$, and $\alpha_2^{CR}(\mu_i, \nu_i)$ are the regression coefficients, and the Δ_i^{CR} is the residual at coarse resolution.

Subsequently, the coarse-resolution regression coefficients $\alpha_0^{CR}(\mu_i, \nu_i)$, $\alpha_1^{CR}(\mu_i, \nu_i)$ and $\alpha_2^{CR}(\mu_i, \nu_i)$ and the residual Δ_i^{CR} are interpolated to fine resolution by using the ordinary kriging interpolation technique (according to the work of Duan *et al*. [*2016*]).

Finally, the fine resolution downscaled LST can be obtained by using the fine resolution NDVI and DEM, as follows:

$$LST_i^{FR} = a_0^{FR}(\mu_i, \nu_i) + a_1^{FR}(\mu_i, \nu_i)NDVI_i^{FR} + a_2^{FR}(\mu_i, \nu_i)DEM_i^{FR} + \Delta_i^{FR} \quad \text{(C2)}$$

where the superscript *FR* represents the fine resolution, the $\alpha_0^{FR}(\mu_i, \nu_i)$, $\alpha_1^{FR}(\mu_i, \nu_i)$, $\alpha_2^{FR}(\mu_i, \nu_i)$, and Δ_i^{CR} are the regression coefficients and residual, respectively, which can be obtained with the ordinary kriging interpolation technique.

D. Parameterization of DTsEB

Table D1. Methods for estimating the intermediate variables/parameters in DTsEB

Parameters	Calculation formula	Description	References
r	for MODIS: $albedo = 0.160b_1 + 0.291b_2 + 0.243b_3 + 0.116b_4 + 0.112b_5 + 0.081b_7 - 0.015$ for ASTER: before April 2008, $albedo = 0.484b_1 + 0.335b_3 - 0.324b_5 + 0.551b_6 + 0.305b_8 - 0.367b_9 - 0.0015$ after April 2008, $albedo = 0.697b_1 + 0.298b_3 + 0.008$	r is the broadband albedo, b_i is the short-wave band spectral reflectances.	Liang, 2003; Mokhtari et al., 2013
G	$G = R_n\left[\Gamma_c + (1 - f_c)(\Gamma_s - \Gamma_c)\right]$	G is the soil heat flux, $\Gamma_c = 0.05$; $\Gamma_s = 0.4$.	Su, 2002
f_c	$f_c = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right)^2$	$NDVI_{min} = 0.2$; $NDVI_{max} = 0.86$; $NDVI > NDVI_{max}$, $f_c = 1$; $NDVI < NDVI_{min}$, $f_c = 0$.	Prihodko & Goward, 1997; Tang et al., 2010
r_a	$r_a = \frac{\left[\ln\left(\frac{Z_u - d}{Z_{om}}\right) - \Psi_m\right]\left[\ln\left(\frac{Z_t - d}{Z_{oh}}\right) - \Psi_h\right]}{k^2 u}$	r_a is the aerodynamic resistance, Ψ_m and Ψ_h are the stability correction functions for momentum and heat transfer, respectively.	Paulson, 1970; Li et al., 2009

Z_{om}	$Z_{om} = \exp(^{aNDVI}/_{albedo} + b)$	Z_{om} is the roughness length for momentum transfer	Tang et al., 2013;
$VegHeight$	$VegHeight = Z_{om} / 0.123$	a = 0.26; b = -2.21	Teixeira et al., 2009
d	$d = \frac{2}{3} VegHeight$	d is the zero-plane displacement height	Allen et al., 2007
Z_{oh}	$Z_{oh} = 0.1 Z_{om}$	Z_{oh} is the roughness length, governing the transfer of heat and vapour	Allen et al., 2007
r_s	$r_s = \frac{G_{s1} + G_{s2} + G_{cu}}{G_{s2} \times (G_{s1} + G_{cu}) \times LAI}$ $G_{s1} = C_L \times m(T\min) \times m(VPD) \times r_{corr}$	r_s is the surface resistance, G_{s1}, G_{s2} and G_{CU} is the stomatal conductance, leaf boundary-layer conductance, and leaf cuticular conductance, respectively. C_L is the mean potential stomatal conductance per unit leaf area, assumed to 0.007 for cropland	Mu et al., 2007; Mu et al., 2011;
ε_s	$\varepsilon_s = \varepsilon_v f_c R_v + \varepsilon_s (1 - f_c) R_s + d\varepsilon$ $R_v = 0.0585 f_c + 0.9332$ $R_s = 0.1068 f_c + 0.9902$	ε_s is the surface emissivity, ε_v=0.986 and ε_s=0.972 are the emissivities of bare soil and vegetation, respectively; Rv and Rs are the temperature ratio for vegetation and bare soil, respectively.	Qin et al., 2004
ε_a	$\varepsilon_a = 1.08(-\ln \tau_{sw})^{0.265}$	ε_a is the atmospheric emissivity, τ_{sw} is the atmospheric transmissivity for short wave radiation	Bastiaanssen, 1995

References

Agam, N., Kustas, W. P., Anderson, M. C., Li, F., & Neale, C. M. U., 2007. A vegetation index based technique for spatial sharpening of thermal imagery. Remote Sens. Environ. 107, 545–558. https://doi.org/10.1016/ j.rse.2006.10.006.

Agam, N., Kustas, W. P., Anderson, M. C., Li, F., & Colaizzi, P. D., 2008. Utility of thermal image sharpening for monitoring field-scale evapotranspiration over rainfed and irrigated agricultural regions. Geophys. Res. Lett. 35(L02402). https://doi.org/10.1029/ 2007GL032195.

Agathangelidis, I., & Cartalis, C., 2019. Improving the disaggregation of MODIS land surface temperatures in an urban environment: a statistical downscaling approach using high-resolution emissivity. Int. J. Remote Sens. 40, 5261-5286. https://doi.org/10.1080/01431161.2019.1579386.

Allen, R. G., Tasumai, M., & Trezza, R., 2007. Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) – Model. J. Irrigation Drainage E. 133(4), 380-394. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:4(380).

Anderson, M. C., Norman, J. M., Kustas, W. P., Houborg, R., Starks, P. J., & Agam, N., 2008. A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales. Remote Sens. Environ. 112, 4227-4241. http://doi:10.1016/j.rse.2008.07.009.

839 Atkinson, P. M., 2013. Downscaling in remote sensing. Int. J. Appl. Earth Observ. Geoinf. 22, 106–114.
840 http://dx.doi.org/10.1016/j.jag.2012.04.012. http://dx.doi.org/10.1016/j.jag.2012.04.012.

841 Bastiaanssen, W. G. M., 1995. Regionalization of surface flux densities and moisture indicators in composite
842 terrain: A remote sensing approach under clear skies in Mediterranean climates. Ph.D. Dissertation, CIP
843 Data Koninklijke Bibliotheek, Den Haag, The Netherlands.

844 Bechtel, B., Zakšek, K., & Hoshyaripour, G., 2012. Downscaling Land Surface Temperature in an Urban Area:A
845 Case Study for Hamburg, Germany. Remote Sens. 4(10), 3184-3200. https://doi.org/10.3390/rs4103184

846 Bindhu, V.M., Narasimhan, B., & Sudheer, K. P., 2013. Development and verification of a non-linear
847 disaggregation method (NL-DisTrad) to downscale MODIS land surface temperature to the spatial scale
848 of Landsat thermal data to estimate evapotranspiration. Remote Sens. Environ. 135, 118–129.
849 http://dx.doi.org/10.1016/j.rse.2013.03.023.

850 Bisquert, M., Sánchez, J. M., & Caselles, V., 2016a. Evaluation of Disaggregation Methods for Downscaling
851 MODIS Land Surface Temperature to Landsat Spatial Resolution in Barrax Test Site. IEEE J. Sel. Top.
852 Appl. Earth Observations Remote Sens. 9, 1430–1438.

853 Bisquert, M., Sánchez, J. M., López-Urreab, R., & Caselles, V., 2016b. Estimating high resolution
854 evapotranspiration from disaggregated thermal images. Remote Sens. Environ. 187, 423-433.
855 http://dx.doi.org/10.1016/j.rse.2016.10.049.

856 Bonafoni, S., 2016. Downscaling of Landsat and MODIS Land Surface Temperature Over the Heterogeneous
857 Urban Area of Milan. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. 9(5), 2019–2027.
858 http://doi:10.1109/JSTARS.2016.2514367.

859 Chen, X., Li, W., Chen, J., Rao, Y., & Yamaguchi, Y., 2014. A combination of TsHARP and thin plate spline
860 interpolation for spatial sharpening of thermal imagery. Remote Sens. 6, 2845–2863.
861 http://dx.doi.org/10.3390/rs6042845.

862 Chen, Y., Zhan, W., Quan, J., Zhou, J., Zhu, X., & Sun H., 2014. Disaggregation of remotely sensed land surface
863 temperature: A generalized paradigm. IEEE Trans. Geosci. Remote Sens. 52(9), 5952–5965,
864 doi:10.1109/TGRS.2013.2294031.

865 Daughtry, C. S. T., Kustas, W. P., Moran, M. S., Pinter, P. J., Jackson, R. D., Brown, P. W., Nichols, W. D., &
866 Gay, L. W., 1990. Spectral estimates soil heat flux of net radiation and soil heat flux. Remote Sens.
867 Environ. 32, 111−124. https://doi:10.1016/0034-4257(90)90012-B.

868 Dominguez, A., Kleissl, J., Luvall, J. C., & Rickman, D. L., 2011. High-resolution urban thermal sharpener
869 (HUTS). Remote Sens. Environ. 115(7), 1772–1780. https://doi:10.1016/ j.rse.2011.03.008.

Dozier, J., 1981. A method for satellite identification of surface temperature fields of subpixel resolution. Remote Sens. Environ. 11, 221–229. https://doi:10.1016/0034-4257(81)90021-3.

Duan, S. -B., & Li, Z. -L., 2016. Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: Case study in Northern China. IEEE Trans. Geosci. Remote Sens. 54(11), 6458-6469. https://doi.org/10.1109/ TGRS.2016.2585198.

Eckmann, T. C., Roberts, D. A., & Still, C. J., 2008. Using multiple endmember spectral mixture analysis to retrieve subpixel fire properties from MODIS. Remote Sens. Environ. 112, 3773–3783. https://doi.org/10.1016/ j.rse.2008.05.008.

Essa, W., Verbeiren, B., van der Kwast, J., van de Voorde, T., & Batelaan, O., 2012. Evaluation of the DisTrad thermal sharpening methodology for urban areas. Int. J. Appl. Earth Observ. Geoinf. 19, 163–172.

Gao, F., Kustas, W. P., & Anderson, M. C., 2012. A data mining approach for sharpening thermal satellite imagery over land. Remote Sens. 4, 3287–3319. http://dx.doi.org/10.3390/rs4113287.

Ghosh, A., R. Sharma, & P. K. Joshi., 2014. Random Forest Classification of Urban Landscape Using Landsat Archive and Ancillary Data: Combining Seasonal Maps with Decision Level Fusion. Appl. Geogr. 48, 31–41. doi: 10.1016/j.apgeog.2014.01.003.

Hutengs, C., & Vohland, M., 2016. Downscaling land surface temperatures at regional scales with random forest regression. Remote Sens. Environ. 178, 127–141. http://dx.doi.org/10.1016/j.rse.2016.03.006.

Inamdar, A. K., & French, A., 2009. Disaggregation of GOES land surface temperatures using surface emissivity. Geophys. Res. Lett. 36(L02408), http://dx.doi.org/10.1029/2008GL036544.

Jeganathan, C., Hamm, N. a S., Mukherjee, S., Atkinson, P. M., Raju, P. L. N., & Dadhwal, V. K., 2011. Evaluating a thermal image sharpening model over a mixed agricultural landscape in India. Int. J. Appl. Earth Observ. Geoinf. 13, 178–191. http://dx.doi.org/10.1016/j.jag.2010.11.001.

Kustas, W. P., Norman, J. M., Anderson, M. C., & French, A. N., 2003. Estimating subpixel surface temperatures and energy fluxes from the vegetation index–radiometric temperature relationship. Remote Sens. Environ. 85, 429–440. https://doi:10.1016/S0034-4257(03)00036-1.

Li, F. Q., & Lyons, T. J., 1999. Estimation of regional evapotranspiration through remote sensing. J. Appl. Meteorol. 38(11), 1644−1654. https://doi:10.1175/1520-0450(1999)038<1644: EORETR>2.0.CO;2.

Li, W., Ni, L., Li, Z. -L., Duan, S. -B., & Wu, H., 2019. Evaluation of Machine Learning Algorithms in Spatial Downscaling of MODIS Land Surface Temperature. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. 12(7), 2299–2307. http://doi:10.1109/JSTARS. 2019.2896923.

Li, Z. -L., Tang, B., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., & Sobrino, J. A., 2013. Satellite-derived land surface temperature: Current status and perspectives. Remote Sens. Environ. 131, 14-37. http://dx.doi.org/10.1016/j.rse.2012.12.008.

Liang, S. L., 2001. Narrowband to broadband conversions of land surface albedo I Algorithms. Remote Sens. Environ. 76(2), 213-238. http://dx.doi.org/10.1016/S0034-4257(00)00205-4.

Liu, D. S., & Pu, R. L., 2008. Downscaling thermal infrared radiance for subpixel land surface temperature retrieval. Sensors, 8, 2695–2706.

Liu, D. S., & Zhu, X. L., 2012. An enhanced physical method for downscaling thermal infrared radiance. IEEE Geosci. Remote Sens. Lett. 9(4), 690–694. http://doi:10.1109/LGRS.2011.2178814.

Liu, J. G., & Moore, J. M. (1998). Pixel block intensity modulation: adding spatial detail to TM band 6 thermal imagery. Int. J. Remote Sens. 19, 2477–2491. https://doi:10.1080/014311698214578.

Liu, K., Wang, S., Li, X., Li, Y., Zhang, B. & Zhai, R., 2020. The assessment of different vegetation indices for spatial disaggregating of thermal imagery over the humid agricultural region. Int. J. Remote Sens. 41, 1907–1926.

Merlin, O., Duchemin, B., Hagolle, O., Jacob, F., Coudert, B., Chehbouni, G., Dedieu, G., Garatuza, J., & Kerr, Y., 2010. Disaggregation of MODIS surface temperature over an agricultural area using a time series of Formosat-2 images. Remote Sens. Environ. 114(11), 2500–2512.

Merlin, O., Jacob, F., Wigneron, J. -P., Walker, J., & Chehbouni, G., 2012. Multidimensional disaggregation of land surface temperature using high-resolution red, near-infrared, shortwave-infrared, and microwave-l bands. IEEE Trans. Geosci. Remote Sens. 50(5), 1864–1880.

Meyers, T. P., & Hollinger, S. E., 2004. An assessment of storage terms in the surface energy balance of maize and soybean. Agric. For. Meteorol. 125, 105–115. http://dx.doi.org/10.1016/j.agrformet.2004.03.001.

Mokhtari, M. H., Busu, I., Mokhtari, H., Zahedi, G., Sheikhattar, L., & Movahed, M. A., 2013. Neural network and multiple linear regression for estimating surface albedo from ASTER visible and near-Infrared spectral bands. Earth Interact. 17(3), 1-20. https://doi.org/10.1175/2011EI000424.1.

Mu, Q., Heinsch, F. A., Zhao, M. & Running, S. W., 2007. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. Remote Sens. Environ. 111, 519-536. https://doi:10.1016/j.rse.2007.04.015.

Mu, Q., Zhao, M. & Running, S. W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. Remote Sens. Environ. 115, 1781-1800. https://doi:10.1016/j.rse.2011.02.019.

930 Mukherjee, S., Joshi, P. K., & Garg, R. D., 2014. A comparison of different regression models for downscaling
931 Landsat and MODIS land surface temperature images over heterogeneous landscape. Adv. Space Res.
932 54(5), 655-669. https://doi.org/ 10.1016/j.asr.2014.04.013.

933 Nichol J., 2009. An Emissivity Modulation Method for Spatial Enhancement of Thermal Satellite Images in
934 Urban Heat Island Analysis. Photogramm. Eng. Remote Sens. 75, 547-556.

935 Olivera-Guerra, L., Mattar, C., Merlin, O., Durán-Alarcón, C., Santamaría-Artigas, A., & Fuster, R., 2017. An
936 operational method for the disaggregation of land surface temperature to estimate actual
937 evapotranspiration in the arid region of Chile. ISPRS J. Photogramm. Remote Sens. 128, 170–181.
938 http://dx.doi.org/10.1016/j.isprsjprs.2017.03.014.

939 Paulson, C. A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable
940 atmospheric surface layer. J. Appl. Meteorol. 9(6), 857−861. https://doi:10.1175/1520-
941 0450(1970)0092.0.CO;2.

942 Prihodko, L., & Goward, S. N., 1997. Estimation of air temperature from remotely sensed surface observations.
943 Remote Sens. Environ. 60, 335−346. https://doi:10.1016/S0034-4257(96)00216-7.

944 Qin, Z. H., Li, W. J., Gao, M. F., & Zhang, H. O., 2006. Estimation of land surface emissivity for Landsat TM6
945 and its application to Lingxian Region in north China. Conf. Remote Sens. Environ. Mon., GIS Apps.,
946 Geo. VI. 636618. https://doi.org/10.1117/12.689310.

947 Quan, J., Chen, Y., Zhan, W., Wang, J., Voogt, J., & Wang, M., 2014. Multi-temporal trajectory of the urban
948 heat island centroid in Beijing, China based on a Gaussian volume model. Remote Sens. Environ. 149,
949 33–46. https://doi.org/10.1016/ j.rse.2014.03.037.

950 Stathopoulou, M., & Cartalis, C., 2009. Downscaling AVHRR land surface temperatures for improved surface
951 urban heat island intensity estimation. Remote Sens. Environ. 112, 2592–2605. https://doi:
952 10.1016/j.rse.2009.07.017.

953 Su, Z., 2002. The surface energy balance system (SEBS) for estimation of turbulent heat fluxes. Hydrol. Earth
954 Syst. Sci. 6(1), 85–99. https://doi:10.5194/hess-6-85-2002.

955 Tang, R., Li, Z. -L., & Tang, B., 2010. An application of the T_s–VI triangle method with enhanced edges
956 determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions:
957 Implementation and validation. Remote Sens. Environ. 114, 540-551.
958 https://doi.org/10.1016/j.rse.2009.10.012.

959 Tang, R., Li, Z. -L., Chen, K. –S., Jia, Y., Li, C., & Sun, M., 2013. Spatial-scale effect on the SEBAL model for
960 evapotranspiration estimation using remote sensing data. Agric. For. Meteorol. 174, 28–42.
961 http://dx.doi.org/10.1016/j.agrformet.2013.01.008.

962 Tang, R., & Li, Z. L., 2017a. An end-member-based two-source approach for estimating land surface
963 evapotranspiration from remote sensing data. IEEE Trans. Geosci. Remote Sens. 55(10), 5818-5832.
964 https://doi.org/10.1109/TGRS.2017.2715361.

965 Tang, R., & Li, Z. L., 2017b. Estimating daily evapotranspiration from remotely sensed instantaneous
966 observations with simplified derivations of a theoretical model. J. Geophys. Res., Atmos. 122, 10254-
967 10267. https://doi.org/10.1002/2017JD027094.

968 Teixeira, A., Bastiaanssen, W.G.M., Ahmad, M.D., & Bos, M.G., 2009. Reviewing SEBAL input parameters
969 for assessing evapotranspiration and water productivity for the low-middle São Francisco River basin,
970 Brazil. Part A: calibration and validation. Agric. For. Meteorol. 149. 462–476.
971 http://doi:10.1016/j.agrformet.2008.09.016.

972 Verma, S. B., Dobermann, A., Cassman, K. G., Walters, D. T., Knops, J. M., Arkebauer, T. J., Suyker, A. E.,
973 Burba, G. G., Amos, B., Yang, H. S., Ginting, D. Hubbard, K. G., Gitelson, A. A., & Walter-Shea, E.
974 A., 2005. Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems. Agric.
975 For. Meteorol. 131(1-2), 77-96. https://doi.org/ 10.1016/j.agrformet.2005.05.003.

976 Valor, E., & Caselles, V., 1996. Mapping land surface emissivity from NDVI: Application to European, African,
977 and South American areas. Remote Sens. Environ. 57(3), 167-184. https://doi.org/ 10.1016/0034-
978 4257(96)00039-9.

979 Weng, Q., Lu, D., & Schubring, J., 2004. Estimation of land surface temperature vegetation abundance
980 relationship for urban heat island studies. Remote Sens. Environ. 89(4), 467-483.
981 https://doi.org/10.1016/ j.rse.2003.11.005.

982 Wu, H., & Li, W., 2019. Downscaling land surface temperatures using a random forest regression model with
983 multitype predictor variables. IEEE Access. 7, 21904-21916.

984 Yang, G., Pu, R., Zhao, C., Huang, W., & Wang, J., 2011. Estimation of subpixel land surface temperature using
985 an endmember index based technique: A case examination on ASTER and MODIS temperature
986 products over a heterogeneous area. Remote Sens. Environ. 115(5), 1202–1219.
987 https://doi:10.1016/j.rse.2011.01.004.

988 Zakšek, K., & Oštir, K., 2012. Downscaling land surface temperature for urban heat island diurnal cycle analysis.
989 Remote Sens. Environ. 117, 114–124. https://doi:10.1016/j.rse.2011.05.027.

990 Zhan, W., Chen, Y., Zhou, J., Li, J., & Liu, W., 2011. Sharpening thermal imageries: A generalized theoretical
991 framework from an assimilation perspective. IEEE Trans. Geosci. Remote Sens. 49(2), 773–789.
992 https://doi:10.1109/TGRS.2010.2060342.

993 Zhan, W., Chen, Y., Wang, J., Zhou, J., Quan, J., Liu, W., & Li, J., 2012. Downscaling land surface temperatures
994 with multi-spectral and multi-resolution images. Int. J. Appl. Earth Observ. Geoinf. 18, 23–36.
995 https://doi:10.1016/j.jag.2012.01.003.

996 Zhan, W., Chen, Y., Zhou, J., Wang, J., Liu, W., Voogt, J., Zhu, X., Quan, J., & Li, J., 2013. Disaggregation of
997 remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats. Remote
998 Sens. Environ. 131, 119–139. https://dx.doi.org/10.1016/ j.rse.2012.12.014.

999 Zhan, W., Huang, F., Quan, J., Zhu, X., Gao, L., Zhou, J., & Ju, W., 2016. Disaggregation of remotely sensed
1000 land surface temperature: A new dynamic methodology. J. Geophys. Res., Atmos. 121, 10538-10554.
1001 https://doi:10.1002/2016JD024891.