Review

Resolution Enhancement of Remotely Sensed Land Surface Temperature: Current Status and Perspectives

Qi Mao, Jian Peng and Yanglin Wang *

Laboratory for Earth Surface Processes, College of Urban and Environmental Sciences, Peking University, Ministry of Education, Beijing 100871, China; maoqi@pku.edu.cn (Q.M.); jianpeng@urban.pku.edu.cn (J.P.)
* Correspondence: ylwang@urban.pku.edu.cn; Tel.: +86-010-6275-0926

Abstract: Remotely sensed land surface temperature (LST) distribution has played a valuable role in land surface processes studies from local to global scales. However, it is still difficult to acquire concurrently high spatiotemporal resolution LST data due to the trade-off between spatial and temporal resolutions in thermal remote sensing. To address this problem, various methods have been proposed to enhance the resolutions of LST data, and substantial progress in this field has been achieved in recent years. Therefore, this study reviewed the current status of resolution enhancement methods for LST data. First, three groups of enhancement methods—spatial resolution enhancement, temporal resolution enhancement, and simultaneous spatiotemporal resolution enhancement—were comprehensively investigated and analyzed. Then, the quality assessment strategies for LST resolution enhancement methods and their advantages and disadvantages were specifically discussed. Finally, key directions for future studies in this field were suggested, i.e., synergy between process-driven and data-driven methods, cross-comparison among different methods, and improvement in localization strategy.

Keywords: spatial resolution; temporal resolution; land surface temperature; image fusion; thermal remote sensing

1. Introduction

Land Surface Temperature (LST) is an important physical parameter of the earth surface system and significantly modulates surface energy flow at local and global scales. Knowledge about the temporal and spatial variations in LST is of fundamental importance when modeling surface processes and is vital for a wide range of applications, including soil moisture estimation [1], wildfire monitoring [2], and urban climate change mitigation [3]. Therefore, satellite-derived time series LSTs with high spatial resolutions provide valuable information that can be used to monitor the land–atmosphere interface dynamics in various landscapes.

The advent of thermal infrared (TIR) sensors has made it possible to obtain LSTs at various scales. The Thematic Mapper (TM) carried onboard Landsat 4/5 and the Enhanced Thematic Mapper Plus (ETM+) on Landsat 7 consist of a thermal infrared band, while the Thermal Infrared Sensor (TIRS) carried onboard Landsat 8 measures land surface temperature in two thermal bands. The National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra/Aqua satellites collects three thermal bands at a 1 km resolution twice per day and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) onboard the Terra satellite has five thermal infrared bands at 90 m every 16 days. The European Space Agency's Along-Track Scanning Radiometer (ATSR), Advanced Along-Track Scanning Radiometer (AATSR), and the Sea and Land Surface Temperature Radiometer (SLSTR) deliver thermal images at a spatial resolution of 1 km. The Visible and Infrared Radiometer (VIRR) onboard Fengyun-3 (B/C) managed by the National Meteorological Satellite Center of China and the Infrared Multispectral Scanner (IRMSS) onboard HJ-1B provide thermal

images at different spatial (1 km, 300 m, and 40 m) and temporal resolutions. In addition, geostationary satellites, such as GOES (Geostationary Operational Environmental Satellite), MSG (Meteosat Second Generation), and Fengyun-4A, could also provide thermal images of the same region of the earth at frequent time intervals (up to 15 min). Detailed information on commonly used sensors in LST retrieval are summarized in Table 1. Many of the aforementioned LST products have been validated over various sites under different atmospheric and surface conditions [4] and are accessible via the internet for free [5], which considerably extend their potential LST applications.

Table 1. Commonly used sensors in land surface temperature (LST) retrieval.

Sensor	Satellite	Spectral Ranges of TIR Bands (µm)	Spatial Resolution	Observation Frequency	Operational Period	Data Access [1]
AVHRR	NOAA	10.5–11.3; 11.5–12.5	1.1 km	Twice daily	Since 1979	https://earthexplorer.usgs.gov/
TM	Landsat 4	10.4–12.5	120 m	16 days	From 1982 to 1993	https://landsatlook.usgs.gov/
TM	Landsat 5	10.4–12.5	120 m	16 days	From 1984 to 2013	https://landsatlook.usgs.gov/
ETM+	Landsat 7	10.4–12.5	60 m	16 days	Since 1999	https://landsatlook.usgs.gov/
TIRS	Landsat 8	10.60–11.19; 11.50–12.51	100 m	16 days	Since 2013	https://landsatlook.usgs.gov/
MODIS	Terra/Aqua	10.78–11.28; 11.77–12.27	1 km	Twice daily	Since 1999/2002	https://modis.gsfc.nasa.gov/
ASTER	Terra	8.125–8.475; 8.475–8.825; 8.925–9.275; 10.25–10.95; 10.95–11.65	90 m	16 days	Since 1999	https://earthexplorer.usgs.gov/
AATSR	Envisat	11; 12 (center)	1 km	35 days	Since 2002	https://earth.esa.int/
SLSTR	Sentinel-3 (A/B)	10.85; 12.02 (center)	1 km	1 day	Since 2016/2018	https://sentinel.esa.int/
VIRR	Fengyun-3 (B/C)	10.3–11.3; 11.5–12.5	1.1 km	Twice daily	Since 2010/2013	http://satellite.nsmc.org.cn/
IRMSS	HJ-1B	10.5–12.5	300 m	31 days	Since 2008	www.chinageoss.cn
GOES Imager	GOES	10.2–11.2; 11.5–12.5	4 km	3 h (Geostationary)	Since 1975	https://www.star.nesdis.noaa.gov/GOES/index.php
SEVIRI	MSG	9.8–11.8; 11.0–13.0	3 km	15 min (Geostationary)	Since 2005	https://landsaf.ipma.pt/
AGRI	Fengyun-4A	8.0–9.0; 10.3–11.3; 11.5–12.5; 13.2–13.8	4 km	36 min (Geostationary)	Since 2016	http://satellite.nsmc.org.cn/

[1] Date access: 24 March 2021.

It is still difficult to obtain LST data with satisfactory temporal and spatial resolutions. On the one hand, the spatial resolutions of thermal bands are always much lower than the visible and near-infrared (VNIR) bands on the same platform because thermal infrared images are digitally sensed by radiance emitted directly by the surface whereas the VNIR images make use of the reflected radiance [6]. For example, the spatial resolutions of the VNIR bands for ASTER are 15 m, while the spatial resolutions of the TIR bands are 90 m. This may result in a serious mixture effect in heterogeneous areas where the grid cell of the TIR images is larger than the thermal elements. On the other hand, single sensors have not yet been able to provide both high temporal resolution and spatial resolution due to the trade-off between scanning swath and pixel size [7]. For example, the widely used Landsat TIR images could be used to generate relatively high spatial resolution LST data (100 m/120 m), but their long revisit interval (at least 16 days) has constrained the potential applications. The time interval between two usable observations could be even longer due to cloud contamination. MODIS, another widely used dataset, has a much higher revisit frequency, but may lose spatial details due to the low spatial resolution (1 km). These issues have meant that increasing attention has been paid to enhancement of the spatial and temporal resolutions by using auxiliary data or fusing data from multiple sensors.

Over the past decade, great success has been achieved in the field of LST resolution enhancement. This success is based on various data and methods inspired by Kustas et al. [8]

and Agam et al. [9]. Zhan et al. [10] systematically conducted a comprehensive review, in which the taxonomy, issues, and caveats of LST disaggregation methods were discussed in details. Since then, due to several technical advances (e.g., spatiotemporal image fusion methods) and increasing satellite data availability within the remote sensing community, there have been numerous publications on enhancing the resolution of LST data, and various ways to enhance the spatial and temporal resolution of remote-sensed LST data have been proposed. Consequently, there is an urgent need to present an overview of state-of-the-art resolution enhancement methods associated with LST data so that future research can be directed towards improving the performance of enhancement algorithms.

This review investigated the latest developments in enhancing spatial and temporal resolution of remote sensed LST data by surveying the progress and interdisciplinary status of the resolution enhancement methods, including spatial resolution enhancement methods, temporal resolution enhancement methods, and methods for enhancing spatial and temporal resolution simultaneously. We have also summarized commonly used reference data and assessment metrics for quality assessment, thereby discussing their advantages and disadvantages noted in the literature. Finally, we suggest directions for future research on this subject.

2. Resolution Enhancement Methods

Different terms are used in the literature to refer to the methods used to improve the spatial and temporal resolutions of LST images because researchers from different disciplines have used different names for their methods, including thermal downscaling [11–13], thermal sharpening [14–16], spatiotemporal image fusion of LST [17–19], and LST disaggregation [20,21]. These terms emphasize different aspects of the resolution improvements to the original LST images. For example, "sharpening" is a term commonly used in remote sensing studies. It combines the spectral information from relatively coarse multispectral images and spatial information from relatively fine panchromatic images to provide more spatial details. Therefore, thermal sharpening addresses the improvements in spatial resolution. In contrast, spatiotemporal image fusion of LST combines spatial information from the fine spatial resolution LST images and temporal information from coarse resolution LST images to generate high spatiotemporal LST images. Thus, it addresses the improvement in temporal resolution compared to the original LST images. Although disparate terms have been used, the purpose of all these studies is to enhance either the spatial or temporal resolution of LST images. From a practical perspective, we use the term resolution enhancement throughout this paper to refer to methods that aim to generate LST images with high spatial and temporal resolutions. This term highlights the improvement in the final products over the original LST images.

Accordingly, existing resolution enhancement methods are categorized into three groups based on their target products: spatial resolution enhancement, temporal resolution enhancement, and simultaneous spatiotemporal resolution enhancement (Figure 1). It should also be noted that diverse classification schemes are applied in the current literature. For example, Quan et al. [22] divided existing methods into two categories based on their input: using sole sensor LST or multi-sensor LST pairs, while Xia et al. [19] classified them into the kernel-driven methods and the fusion-based methods. As the aim of all resolution enhancement methods is to produce better LST products, we applied the output-based classification scheme in this review. Similar classification schemes can also be found in Weng et al. [23] and Weng and Fu [24].
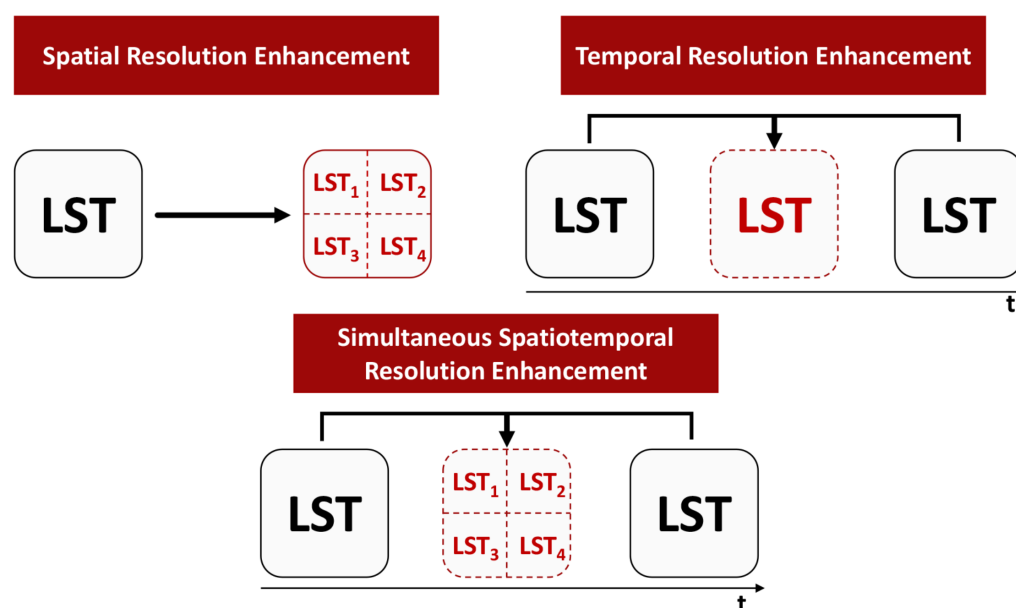
Figure 1. Three groups of LST resolution enhancement methods.

2.1. Spatial Resolution Enhancement

The enhancement of low-spatial-resolution LST images usually relies on high-spatial-resolution auxiliary data that are statistically correlated to LST [10]. The parameters used to construct the statistical regression relationship and portray the LST variations are often called "kernels" (or "scaling factors"), which originate from the kernel methods (widely used in algorithms such as the support vector machine) that are used for transforming nonlinear relationships into linear ones [25]. Therefore, spatial resolution enhancement methods are also often referred to as "kernel-driven methods" as well [19].

Note that in addition to kernel-driven methods that utilize auxiliary data to enhance the spatial resolution of LST data, traditional spatial interpolation methods based on the spatial autocorrelation among measured samples can also be used to enhance the spatial resolution of LST data [26]. However, spatial interpolated LSTs are merely a smooth expansion based on the coarse LST and are unable to reconstruct the rich details of the actual LST [27]. Therefore, it can only act as a complement for kernel-driven methods. For example, Chen et al. [28] combined a kernel-driven method and a spatial interpolation method to perform LST spatial enhancement. The process produced considerably better results than using only one method. The integration of interpolation methods could mitigate the block effect caused by kernel-driven methods [21,22].

2.1.1. General Process

Spatial Resolution Enhancement of LST data is a well-established research field. This section reviews the general process underlying the spatial resolution enhancement methods for low-resolution LST images from empirical and rational perspectives. The publication and the widespread use of the classical DisTrad [8] and TsHARP [9] sharpened the research focus and powered the development of LST spatial resolution enhancement. They were inspired by the well-known empirical relationship between vegetation index (VI) and LST. The procedure for enhancing coarse-resolution LSTs to the VI-pixel resolution can be summarized using a single equation:

$$LST_F = LST_C + [f(VI_F) - f(VI_C)] \tag{1}$$

where the subscripts "*F*" and "*C*" denote fine and coarse resolution, respectively, and *f* is the mapping relationship between *LST* and *VI* (regression kernels), which is usually

constructed at the coarse scale. Correlations between *LST* and *VI* are mainly caused by the general cooling effect by vegetated surfaces due to shading and increased evapotranspiration [12].

In addition to the aforementioned empirical generalization, Chen et al. [28] provided a thorough theoretical paradigm for LST spatial resolution enhancement methods. Based on the assumption that the errors associated with background temperature and indirect observations follow a normal distribution, the general spatial resolution enhancement procedures could be regarded as an optimized problem that tends to find the fine-resolution LST with the highest probability density using the observed background LST (un-enhanced temperature at coarse spatial resolutions) and indirect LST observations (estimated by regression kernels). When the regression kernels are linearly related to LST, the optimized enhanced LST could be extended to the following generalized form:

$$\mathbf{LST_F} = \mathbf{LST_b} + \Sigma\phi^T(\Omega + \phi\Sigma\phi^T)^{-1}(\mathbf{Y_{obs}} - \phi(\mathbf{T_b})) \tag{2}$$

where $\mathbf{LST_F}$ is the enhanced LST (fine-resolution), $\mathbf{LST_b}$ is the background LST, $\mathbf{Y_{obs}}$ is the indirect LST observations, ϕ is the functional representation of the statistical relationship between LST and regression kernels, and Σ and Ω denote the covariance matrices for $\mathbf{LST_b}$ and $\mathbf{Y_{obs}}$, respectively. Accordingly, Equation (1) could be regarded as a special case of Equation (2) [28].

Two key aspects have been highlighted in the empirical generalization and theoretical paradigm for spatial resolution enhancement of LST. (1) Whether the coarse resolution LST is an unbiased estimator of the target fine resolution LST within the grid cell. This mainly depends on the characteristics of the data source, such as georegistration errors between input and target LSTs and the sensor errors between different sensors. Extra attention should be paid to these characteristics when applying spatial resolution enhancement using data acquired from multiple sensors or at different times [29]. (2) Whether the indirect observation by the regression kernel is an unbiased estimator of the fine resolution LST. This mainly depends on the accuracy of the indirect observation. Currently, the indirect observation is generated by a regression model that establishes the statistical relationships between LST and regression kernels for certain areas. Consequently, suitable regression kernels, regression tools, and regression scales are desirable when attempting to improve the accuracy of spatial resolution enhancement of LST. The following sections will further discuss each of the key issues involved in designing spatial resolution enhancement methods.

2.1.2. Regression Kernels and Tools

Regression kernels are employed to bridge the gap between initial and target resolutions in LST spatial resolution enhancement methods. The basic question identified by Zhan et al. [30] is still pervasive in current literature: What are well-behaved kernels, and how should they be configured?

Table 2 shows an overview of commonly used regression kernels in LST spatial resolution enhancement. For remote sensing in the TIR spectrum, differences in the abilities of objects and landscape features to absorb shortwave radiation and emit energy are sensed. The collected TIR energy by imaging sensors is rather weak, and the spatial resolutions of the VNIR bands are commonly higher than those of the TIR bands on the same platform. Therefore, implementing LST spatial enhancement with VNIR bands and their derivatives is natural. The NDVI has a widely acknowledged trapezoid relationship with LST [31] and is the most popular regression kernel used in the current literature. In urban areas and other less vegetated areas, the NDBI can be used to portray the distribution pattern for LST [32]. Several kinds of auxiliary data that could be used to parameterize land use patterns [12], topographic characteristics [13], or energy balance conditions [33] have also been included to increase the accuracy of spatial enhancement.

Two threads are present in the selection of the ideal regression kernels: (1) From single to multivariate. In the classical algorithm used in [9] or [34], only one kernel was used to enhance the spatial resolution of the LST images, whereas recent studies [35] used more than

20 kernels to achieve high global operability. (2) From one-class to hybrid. Recent studies tend to employ different kinds of regression kernels, such as VNIR reflectance [33], band composites [36], or other LST-related variables derived from multiple data sources [37,38]. The use of combinations of kernels is largely dependent on prior knowledge of the study area as well as data availability and could result in higher accuracy under certain scenarios [39–41]. In addition to increasing the quantity and variety of regression kernels, some studies [19,42] applied higher degree variables (e.g., quadratic) [43] to better capture the complicated and nonlinear relationships between LST and different predictor variables. In general, the use of multiple forms of regression kernels tends to result in better enhancement across specific areas.

The performance of different regression kernels is highly variable and is affected by various factors, such as land cover, latitude and longitude, or season [30]. The regression kernels that perform well in specific areas may introduce large errors in other areas. This greatly limits the applicability of spatial enhancement methods across space and time and raises uncertainties when implementing LST spatial resolution enhancement in heterogeneous areas. Therefore, adequate prior knowledge has remained indispensable for spatial enhancement methods and it may still require time-consuming feature engineering to configure suitable regression kernels.

Table 2. Overview of the regression kernels used in LST spatial resolution enhancement methods.

Regression Kernel [1]	Category	Literature
VNIR Reflectance	Image-derived index	[12,33]
NDVI	Image-derived index	[8,9,11,13,15,16,18–20,25,28,29,32,33,36,38,40–51]
EVI	Image-derived index	[33,36,52–54]
NDMI	Image-derived index	[29,36]
NMDI	Image-derived index	[55,56]
SAVI	Image-derived index	[16,29,33,50,55–58]
NDBI	Image-derived index	[16,19,29,32,33,50,51,55,56,58,59]
NDWI	Image-derived index	[33,36,50,55]
MNDWI	Image-derived index	[29,33,56]
Albedo	Image-derived index	[18,20,43,52,53,60]
Emissivity	Image-derived index	[32,33,48,52–54,60,61]
Fractional vegetation cover	Image-derived index	[9,34,38]
DEM	Auxiliary data	[11–13,33,51,54,59–61]
Solar incidence angle	Auxiliary data	[12,33]
Land cover classification	Auxiliary data	[12,32,54,61]

[1] EVI (Enhanced Vegetation Index), NDMI (Normalized Difference Moisture Index), NMDI (Normalized Multi-band Drought Index), SAVI (Soil Adjusted Vegetation Index), NDBI (Normal Difference Built-up Index), NDWI (Normalized Difference Water Index), and MNDWI (Modified Normalized Difference Water Index).

Likewise, various regression tools have been employed to establish the relationships between regression kernels and LST. Several statistical models have been widely applied, including multivariate linear regression [34,58], piecewise linear regression [62], principal component regression [52,53], and geographically weighted regression [11]. These tools can achieve satisfactory enhancement results for rather homogeneous areas [19] and provide rapid enhancement due to their relatively low computational complexity [63].

In recent years, the emergence of machine learning algorithms has generated a step change in establishing the relationships between regression kernels and LST. Multiple machine learning algorithms, including the artificial neural network [47,63,64], the self-organizing feature map network [57], support vector regression [54,63,65], and random forest [35,36,66], have been used. These new tools have reduced uncertainties due to the high complexity caused by the variations in kernel–LST relationships across different regions and periods and can approximate the relationship more accurately. The rapid development in machine learning has meant that novel methods, such as convolutional neural networks and recurrent neural networks, have emerged. They could potentially underpin understanding of diverse behaviors and nonlinear interactions associated with LST spatial enhancement [67,68].

There has been substantial progress in the comparison of different regression kernels and tools. However, the efficiency and accuracy of the kernel/tool selection made in previous studies may not have always met the needs of LST spatial enhancement in practice [54].

Further, the applicability of conclusions were limited in terms of the land cover type, topography and the climate zone of the study area [48,63,65]. Recently, Dong et al. [35] systematically investigated the optimal choice for regression kernels/tools and their associated combinations over different land conditions from a global perspective. The study compared the combinations of seven regression kernel sets (albedo, seven Landsat 8 VNIR bands, and their twenty-one corresponding normalized difference indices and ratio indices) and five frequently used regression tools (multiple linear regression, partial least squares regression, artificial neural networks, support vector regression, and random forest). Their results showed the algorithms that used random forest as the regression model and 21 regression kernels derived from Landsat 8 data (after eliminating the redundant variables) achieved the best performance in terms of both accuracy and stability.

Overall, different regression kernels and tools should be utilized for different tasks due to the considerable variations in kernel–LST relationships across different regions and periods. Enhancement methods must account for the high heterogeneity of the relationships if they are to accomplish better enhancement results. Additional efforts should be exerted to improve the efficiency of the selection process for regression kernels and tools, especially with regards to global applicability.

2.1.3. Regression Scale

Conceptually, remote sensing imagery encapsulates two aspects of scale: grain and extent [69]. In the context of LST spatial enhancement methods, two corresponding issues in relation to scale are important: whether the kernel-LST relationships stay invariant at different spatial resolutions (i.e., scale invariance) and how to find a suitable regression window to obtain the kernel–LST relationships (i.e., localization strategy).

(1) Scale invariance

As the kernel–LST relationships cannot be obtained directly at high spatial resolutions, most LST spatial enhancement methods are based on the assumption that the relationships between regression kernels and LST are invariant at different resolutions. Nonetheless, the scale invariance assumption may be insufficient and is being debated in current literature. Jeganathan et al. [62] assessed the NDVI–LST relationships at different scales using multisource remote sensing data in northern India and found that R^2 for NDVI–LST relationships declined as the pixel size changed from 180 m to 990 m. Chen et al. [46] observed a similar phenomenon on the Inner Mongolia Plateau and the North China Plain. Ghosh and Joshi [48] conducted an experiment using hyperspectral data and Landsat-derived LST across an eastern Indian tropical landscape in different seasons. This study concluded that the scale dependence of the kernel–LST relationships is a function of the study area characteristics and regression tools. Additionally, Zhou et al. [42] developed an analytical model to quantify the scale variance of kernel–LST relationships and demonstrated that the scale effect is intrinsically due to the variance in the probability distribution for LST and the regression kernels at different resolutions. Therefore, removing the scale effect would not necessarily improve the accuracy of the enhanced LST.

Although the assumed scale invariance acts as a basis for LST spatial enhancement methods, attention should be paid to the feasibility, especially when there is a relatively large difference between the original and target resolutions. Proper selection of the regression kernel/tool is important, and the uncertainties caused by the scale effect must be considered on this basis.

(2) Localization strategy

In an attempt to take into account the varying kernel–LST relationships in space, researchers applied a localization strategy, which established regression relationships at the local geographical area, to improve the stability and accuracy of the regression for the local data in the previous literature. Therefore, the selection of the most suitable localization strategy for LST spatial enhancement is an important research topic.

The moving window technique is the most common localization strategy in this field. This technique constructs local kernel–LST relationships using data located within the

window sliding across the entire scene. Thus, the size of the moving window explicitly accounts for the performance of the enhancement methods [70]. Jeganathan et al. [62] found that the 7 × 7 window covered too many multiple land cover types while the 3 × 3 window led to unstable results. Therefore, the 5 × 5 window was the most appropriate choice because it balanced the trade-off between sample size and landscape heterogeneity. Similarly, Chen et al. [28] and Zhukov et al. [71] also suggested that the 5 × 5 was the most suitable for their case studies. In general, a too large window may contain redundant information and reduce the representativeness of the corresponding regression relationships, whereas a too small window yields insufficient information and reduce the stability of the corresponding regression relationships. Different methods have been proposed to determine the optimal window size for an arbitrary study area when applying LST spatial enhancement. Zhan et al. [72] demonstrated that the range in a semi-variance analysis of regression kernels could be used to determine the optimal window size. An indirect criterion based on aggregation–disaggregation was formulated by Gao et al. [20]. They suggested that there was a statistical linear relationship between the optimal window size and the resolution ratio between original and target LST. This empirical relationship could be used to choose the global and the local window strategies, and determine the optimal moving window size.

The shape of the moving window have also been taken into consideration by the current literature. Most previous studies used a square window when applying a localization strategy. However, this might be less accurate because the shape of the land cover objects is not always regular. The shape of the moving window could be flexible as long as it decreases the local anisotropic heterogeneity of the local area. Consequently, some studies applied a local strategy that aimed to determine optimal local area using adjacent pixel homogeneity. Lillo-Saavedra et al. [36] established the kernel–LST relationships by random forest regression within the superpixels generated by the simple linear iterative clustering (SLIC) algorithm. This method outperforms the TsHARP method on average by 0.218 K in terms of the root mean square error. Xia et al. [16] also determined the size and shape of the moving windows by SLIC algorithm and conducted LST spatial enhancement within these irregularly shaped windows. The proposed local strategy improved the mean accuracy by 0.44 K and 0.19 K, compared to the traditional global and local window strategies, respectively.

Overall, an appropriately configured local strategy could generate a higher accuracy level for LST spatial enhancement. In essence, the key issue is how to extract a homogeneous local area where kernel–LST relationships are stable. Therefore, future studies should focus on selecting similar samples for regression by identifying useful spectral, spatial, or temporal features from remote sensing images and auxiliary data while balancing both enhancement accuracy and computational complexity.

2.2. Temporal Resolution Enhancement

The temporal enhancement methods for LST used complementary information to reconstruct the original LST products at annual or diurnal scales and this could be achieved in the following two ways. (1) Interpolation-based methods. These methods utilize physical models, such as diurnal/annual temperature cycle (DTC/ATC) models, which can effectively describe the thermal behaviors and dynamics of land surface using a set of model parameters to interpolate temporally discrete LST observations [73]. After parameterization, the DTC/ATC models can be further used to produce dense time series LST data. (2) Fusion-based methods. The methods utilize the spatiotemporal data fusion methods, which blend together multisource remote sensed data [74]. Two types of LST products, frequent coarse-resolution LST products (e.g., MODIS/geostationary satellites) and sparse fine-resolution LST products (e.g., Landsat/ASTER), are commonly used to produce synthesized LST products with both high spatial and temporal resolutions.

2.2.1. Interpolation-Based Methods

Various DTC/ATC models have been developed to describe the diurnal/annual variations of LST. The aim is to interpolate the missing values and to generate continuous time series LST data. The LST dynamics at diurnal and annual scales are fundamentally controlled by all the components contained in the surface energy balance [73,75]. Temperature-related parameters, including surface properties (e.g., surface geometry, emissivity, albedo, and thermal diffusivity), meteorological variables (e.g., wind, air temperature, humidity, and pressure), and other implicit parameters linked to surface and atmospheric status (e.g., temperature amplitude, residual temperature, and total optical thickness), could be used to model temperature cycle [76,77]. With the aid of different assumptions or constraints on solar irradiation and the surface energy balance equation, diverse functions, such as harmonic functions or a combination of harmonic terms with an exponential decay function, have been applied to fit the thermal behavior of the surface [78]. These temperature cycle models could be further used to interpolate temporally discrete LST observations derived from satellite.

Interpolated-based temporal resolution enhancement methods for LST have been widely discussed in the literature. Jin and Dickinson [79] and Sun and Pinker [80] developed DTC models using LST data obtained by geostationary satellites. They then applied the model to polar-orbiting observations to enable frequent time series estimates at a diurnal LST scale at relatively high spatial resolutions. Furthermore, many other methods, including those proposed by Inamdar et al. [44], Duan et al. [81], and Huang et al. [76], have modified the model forms and introduced additional terms that are related to the dynamics of surface property and atmospheric status. Additionally, the annual temperature cycle could also be modeled to analyze the LST variability at an annual scale and mitigate the contamination caused by clouds and other atmospheric conditions [82,83].

The increasing availability of multiple polar-orbiting and geostationary thermal remote sensors means that the data requirements needed by temperature cycle models to achieve stable solutions over one cycle can be readily satisfied. However, the following drawbacks are also evident. As the diurnal temperature cycle is defined under different constraints, determining temperature cycle model parameters is largely dependent on a prior knowledge of the study regions and models may fail when the surface and atmospheric status invalidate the certain constraints. The computational cost may become considerably large when attempting to establish temperature cycle models in highly heterogeneous areas [77]. In addition, interpolation-based methods are based on the assumption that the LST series are cyclically-stationary. This simplification is hardly feasible in reality and causes unexpected uncertainties in temporally interpolated LSTs because of multiple factors, such as changes in land cover or human disturbances [84]. Therefore, some studies have discussed the possibility of utilizing additional spectral analyses to improve model stability. Quan et al. [85] combined a DTC model with a linear temperature mixing model. This new model employed multi-temporal and multi-angular LST observations based on mixed pixels and their corresponding component fractions retrieved by VNIR reflectance to fit the DTC model for each component. The LST was then reconstructed by a weighted linear combination of the DTC for each component. Although the interpolated-based methods have their drawbacks, they can still potentially be combined with other temporal enhancement methods. The fusion of various LST temporal enhancement methods will be briefly presented in the following section.

2.2.2. Fusion-Based Methods

Recently, there have been breakthroughs in spatiotemporal image fusion in the remote sensing field. The spatiotemporal image fusion methods are designed to address the trade-off between the spatial and temporal resolutions of a single satellite sensor [86]. Two types of satellite images—frequent coarse-resolution images and sparse fine-resolution images—are used in these methods to produce synthesized images with both high spatial and temporal resolutions [87]. Spatiotemporal image fusion has been successfully used in

a wide range of applications, including land cover classification, crop growth monitoring, and ecosystem dynamics monitoring [88]. This review will concentrate on the use of spatiotemporal image fusion methods to advance LST resolution enhancement. Broader reviews on the principles and applications of general spatiotemporal image fusion methods can be found in Zhu et al. [7] and Belgiu and Stein [89].

Although most spatiotemporal image fusion methods were initially designed for surface reflectance, they can still be used to enhance LST temporal resolutions. Liu and Weng [90] directly applied the spatial and temporal adaptive reflectance fusion model (STARFM) proposed by Gao et al. [91], which assumes that changes in reflectance are consistent at coarse and fine resolutions. It estimates daily Landsat-like reflectance using a weighted sum of reflectance changes derived from coarse pixels and fine-resolution reflectance in similar neighboring pixels to fuse ASTER and MODIS data and create dense time series ASTER-like LSTs. Wu et al. [17] extended the STARFM and proposed the spatiotemporal integrated temperature fusion model (STITFM), which integrates LSTs from multi-scale polar-orbiting and geostationary satellite observations to generate diurnal Landsat-like LSTs. However, the STARFM and its variants suffer from low accuracy and weak robustness in complex and heterogeneous landscapes. Therefore, several studies modified the original STARFM and improved the model performance over landscapes with mixed and changing land cover types. These modified methods have also been used to fuse LST data. For example, the enhanced STARFM (ESTARFM) proposed by Zhu et al. [74] and an unmixing-based spatiotemporal image fusion method proposed by Wu et al. [92], also known as the spatial and temporal data fusion approach (STDFA), have been used to fuse ASTER and MODIS LSTs [93]. The results showed that the daily synthetic LSTs generated by ESTARFM and STDFA had a closer match with actual ASTER observations than STARFM, especially over cultivated land areas. To improve the model performance in heterogeneous areas, the ESTARFM introduces a conversion coefficient calculated by a linear spectral mixing model, while the STDFA estimates reflectance change using unmixing endmember reflectance.

Several studies have attempted to develop advanced spatiotemporal fusion models specially for LSTs. The modifications are mainly based on physical models, such as temperature mixing models and temperature cycle models. These models are able to effectively express the thermal radiative transfer process and could be adopted to characterize the spatial and temporal patterns for enhanced LSTs. After considering the observation differences between sensors and the spatial correlations between pixels, Wu et al. [94] extended the STARFM and used a spatiotemporal fusion method to obtain daily and hourly LSTs at a Landsat spatial scale. Weng et al. [23] combined a linear spectral mixture model and an ATC model and proposed the Spatio-temporal Adaptive Data Fusion Algorithm for Temperature mapping (SADFAT) to predict LSTs at high temporal frequencies and medium spatial resolution by fusing LSTs generated by MODIS and Landsat. The SADFAT modified the STARFM using a conversion coefficient that related the thermal radiance change in a mixed pixel at a coarse resolution to that of a fine resolution pixel. Quan et al. [22] combined a linear temperature mixing model, a DTC model, and an ATC model, and proposed an integrated framework to BLEnd Spatiotemporal Temperatures (BLEST). In addition to using the conversion coefficients derived from temperature cycle models, the BLEST also interpolates the residuals of the weight function using the thin plate spline technique to reduce the biases from land cover type change. Compared to the STARFM, ESTARFM, and STITFM, the BLEST achieved higher accuracy with more spatial details and pronounced temporal evolutions when validated by in situ data. Apart from integrating physical-based models into spatiotemporal image fusion procedures, some studies have also applied machine learning algorithms to cope with the nonlinear nature of the LSTs at different scales. For example, Moosavi et al. [95] proposed a hybrid wavelet-artificial intelligence fusion approach (WAIFA), which coupled wavelet transformations with artificial neural networks to fuse MODIS and Landsat LST data. The enhancement results were a considerable improvement in terms of accuracy and efficiency, which implied that incorporating machine learning

algorithms could help the model deal with the non-stationary properties of LST and capture the complex relationships between observations from different sensors.

Spatiotemporal image fusion methods have successfully enhanced LST resolutions in recent years and this success may lead to rapid future developments [96]. However, fusion methods must account for the mixed state of the non-pure pixels and the sub-pixel changes in the land cover when they attempt to accomplish accurate LST temporal enhancement. Such methods should also investigate the comprehensive relationships between LST observations from different sensors and consider spatial/temporal dependence of LST distribution. Uncertainty assessment for spatiotemporal image fusion methods needs further explorations considering the instrumental noise, retrieval errors, and abrupt land surface changes.

2.3. Simultaneous Spatiotemporal Resolution Enhancement

The aforementioned resolution enhancement methods have their own limitations. The spatial enhancement methods provide spatial details at the scale of the regression kernels (mostly VNIR bands), but are limited by the temporal resolution of the original LSTs, while the temporal enhancement methods produce dense LST time series, but are limited by the spatial resolution of the available fine LST. Therefore, some studies concentrate on combining two kinds of enhancement methods to simultaneously realize spatial and temporal resolution enhancement. Different enhancement methods have been integrated. Bai et al. [97] first enhanced Landsat ETM+ LST from 60 m to 30 m using the extreme learning machine as the regression tool and VNIR bands as the regression kernels. Then, the enhanced LST data were fused with high-frequency MODIS LST observations by SADFAT [23] to derive synthetic LSTs at high spatiotemporal resolutions. Zhan et al. [30] combined interpolation-based methods and established temperature cycle models at the coarse resolution scale. They then used the regression relationships between the model parameters and scaling factors at the coarse resolution to further predict the model parameters using fine-resolution scaling factors and produced temporally continuous LSTs at a spatial resolution equivalent to that of the scaling factors.

Two kinds of strategies that combined the spatial and temporal enhancement methods were identified by Xia et al. [18]. The first kind was the "regression-then-fusion" (R-F) strategy. This strategy attempts to first enhance the spatial resolution of the LSTs into the fine resolution LSTs and then enhance the temporal resolution of the LSTs obtained in the previous step. This strategy was employed by the aforementioned enhancement methods proposed by Bai et al. [97]. In contrast, another kind is the fusion-then-regression (F-R) strategy, which tends to enhance the temporal resolution first and then enhance the spatial resolution of the obtained LST series. After testing the two strategies, Xia et al. [18] concluded that the R-F strategy performed better when the performance of the regression at the start time was better than that at the target time, whereas the F-R strategy performed better when the performance of the regression at the target time was better than that at the start time. On this basis, Xia et al. [19] further proposed the weighted Combination of Kernel-driven and Fusion-based Method (CKFM). One part of CKFM enhanced the spatial resolution of Landsat LSTs using Landsat VNIR bands. It then estimated the LSTs at the target date using the enhanced LSTs and MODIS LSTs under the invariant-sensor-error assumption. The other part utilized the STARFM and TPS interpolation to derive estimated fine-resolution LSTs at the target date. Finally, the results of the two parts were combined via weights calculated from error estimations.

Although several methods have been proposed to integrate the spatial and temporal enhancement methods, additional work is still required in this field. Future studies should consider how to select proper enhancement methods that take into account the spatiotemporal variations and model performances at different scales. In addition, the integration strategy and uncertainty analysis of multiple enhancement methods should be further explored. Most existing studies have presented an overall accuracy assessment of enhanced LSTs. However, there should be a separate uncertainty analysis for different enhancement

procedures, especially when the assumptions and restrictions involved in the enhancement methods cannot be met simultaneously.

3. Quality Assessment

Quality assessment is an important requirement when developing reliable enhancement methods. It is used to assess the uncertainty and applicability of the proposed methods. The results can provide potential users and future developers with comprehensive information about the strength and weaknesses of the enhanced LST products. The resolution enhancement methods aim to produce LST data that cannot be observed directly. Therefore, any assessment of the enhancement methods may need indirect reference data (Figure 2). Furthermore, the closeness between the enhanced and reference LSTs, and the spatiotemporal details reserved in the enhanced LSTs should be considered when assessing the enhancement quality [7]. Therefore, diverse assessment metrics have been proposed to address this issue. The following sections will further introduce the commonly used reference data and assessment metrics in LST enhancement and discuss their advantages and disadvantages.
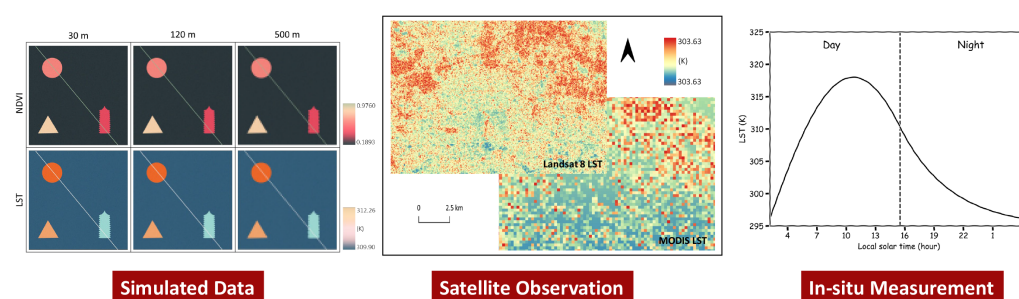
Figure 2. Examples of reference data to assess the LST enhancement methods.

3.1. Reference Data

3.1.1. Simulated Data

Simulated data allow us to assess the strengths and weaknesses of the enhancement methods in certain aspects. By controlling the spectral, structure, and evolution characteristics of the data, representations of the real-world objects, such as built-up area, bare soils, and vegetation, or incidents, such as construction/destruction and flood/drought, could be simulated. For example, Xia et al. [16] simulated linear, circular, and rectangular objects and assigned different ranges of NDVI values to different objects to assess whether the proposed LST spatial enhancement method could be adapted to different object shapes. Quan et al. [22] simulated multiple image series by assigning each pixel a digital number (DN) of 0–255. The image series were generated with different evolutionary characteristics, including gradual changes in DN value, changes in the size of objects, and combinations of the two changes. The simulated image series were used to test whether the proposed BLEST could restore the spatiotemporal variations and predict diurnal temperature cycles for areas where there were changes in the land cover type. In particular, the test indicated that there was a blurred linear feature inside the large rectangular object when conducting visual comparison. This drawback of the proposed BLEST was attributed to the smoothing effect by TPS interpolation.

Overall, simulated data have been widely used to evaluate the LST enhancement methods and provide a controlled environment for quality assessment. They have a natural advantage because they can interpret the model performance in detail. In practice, simulated data could be used in preliminary assessments and should be regarded as useful complementary data when assessing LST enhancement methods.

3.1.2. Satellite Observation

Comparing the actual satellite-derived LSTs and the enhanced LSTs at the target spatiotemporal scales is a direct way of assessing the resolution enhancement methods. In practice, it is helpful to evaluate the model performance under real conditions. However, the feasibility of an assessment based on satellite-derived LSTs crucially relies on the data availability. This requirement cannot always be met because of the spatial and temporal inconsistency between the observed and enhanced LSTs. On the one hand, spatial-enhanced LSTs are rarely available in practice because of the common differences between the spatial resolutions of the VNIR and TIR bands on the same platform. On the other hand, the fluctuant satellite overpass time may be not coincide with the time of interest. Furthermore, a direct comparison might not be suitable because the variations in LST during day could be more than 10 K over short time intervals.

Therefore, indirect assessments based on the actual satellite-derived LSTs have been performed in the previous studies. When high spatial resolution reference LST data are not available, then an indirect criterion based on aggregation–disaggregation [9,20] is commonly employed, wherein the original LST data are aggregated to a coarser resolution. The data are then disaggregated (i.e., enhanced) to the original resolution and subsequently validated by the original LST observations. Weng and Fu [24] applied a predefined DTC model to MODIS LSTs, and the estimated LST series was further used to validate the enhanced GOES LST (from 5 km to 1 km). Note that these indirect assessment methods based on satellite-derived LSTs may introduce additional uncertainties during the aggregation or temporal interpolation processes, thus assessment based on the satellite-derived LSTs are valuable, yet challenging procedures.

3.1.3. In-Situ Measurement

A geostationary satellite can provide high-frequency LST observations that can be used to evaluate the high temporal resolution enhanced LSTs, but data availability highly depends on the study area and period. Therefore, in situ measurement needs to be employed as another kind of reference data to assess the enhanced LST. In situ measurement can effectively evaluate hour-by-hour enhanced LST values from the ground level. For example, Zhou et al. [98] evaluated the hourly 1 km LST generated by the proposed diurnal temperature cycle genetic algorithm using the in situ measurements of LST at two meteorological stations. Quan et al. [22] also generated hourly Landsat-like LSTs from the proposed BLEST and used in-situ LSTs as a supplement when considering several other evaluation approaches that utilized polar-orbiting and geostationary satellite observations.

In situ LST measurements do not necessarily represent LST values at the pixel scale due to the large spatial variations in LSTs, especially during the day [99]. Accurate quantitative assessments of the enhanced LSTs require high-quality in situ LST observations over areas that are homogeneous at the pixel scale [4]. Therefore, it is necessary to consider the spatial homogeneity across the enhanced LST pixel when using in situ LST measurements to conduct quality assessment.

3.2. Assessment Metrics

Various quantitative metrics have been employed to compare the similarity between the enhanced and reference LSTs. Most existing studies used residual-related metrics to measure how far the enhanced LSTs were from the reference LSTs, such as the root mean square error (RMSE) [51], the mean absolute error (MAE) [29], and the Pearson's correlation coefficient (*r*) [22]. The spatial resolution difference between pre- and post-enhanced LSTs meant that the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS; which means relative dimensionless global error in synthesis), which includes a correction factor that compensates for the resolution difference, was employed as a supplement to residual-related metrics [100]. In addition, the metrics used in optical image fusion could also be extended to evaluate the LST resolution enhancement methods. For example, Rodriguez-Galiano et al. [101] used the structural similarity index (SSIM) [102] to measure

the luminance, contrast, and structure differences between pre- and post-enhanced LSTs. Subsequently, Zhou et al. [42] used the image quality index (Q), which was proposed by Wang and Bovik [103], to model the differences between the original and enhanced LSTs, such as the combined loss of correlation, luminance distortion, and contrast distortion.

Although quantitative assessment metrics have rapidly developed over recent years, the majority of them are intrinsically flawed. Three drawbacks associated with the assessment metrics for spatial resolution enhancement were identified by Gao et al. [21]. These were the inability to distinguish the retrieval errors, low competence in removing the process controls (e.g., differences in thermal contrasts, temperature units, and resolution ratios), and an inability to denote the sharpening statuses (e.g., under- or over-sharpening). Therefore, a Simple Yet Flexible Index (SIFI) was subsequently designed to mitigate the impacts of the aforementioned issues. At present, many types of assessment metrics are available [21]. However, a cross-comparison that can help users select the most suitable enhanced method for their specific cases is an ongoing challenge because of the differences in the assessment metrics. Furthermore, the assumptions or simplifications regarding the metrics may decrease both the feasibility and the applicability of LST enhancement algorithms and procedures. Therefore, improving the universality and reliability of the assessment metrics is a valuable, but difficult, area for future research.

4. Future Development and Perspectives

Acquiring LSTs at high spatiotemporal resolutions is important to many fields of study, including energy balance/exchange, water circulation, and climate change over the surface of the earth. Various methods have been proposed to enhance the spatial and temporal resolutions of LSTs. These often combine diverse satellite-derived/auxiliary data and state-of-the-art enhancement tools. Significant progress has been made over recent decades in instruments and techniques used to sense thermal radiance and retrieve LSTs. However, it seems unlikely that there will be any rapid development toward obtaining direct LST observations at both high spatial and temporal resolutions in the short-term. Therefore, it is necessary to explore new ideas and new paths in the LST resolution enhancement field (Figure 3).
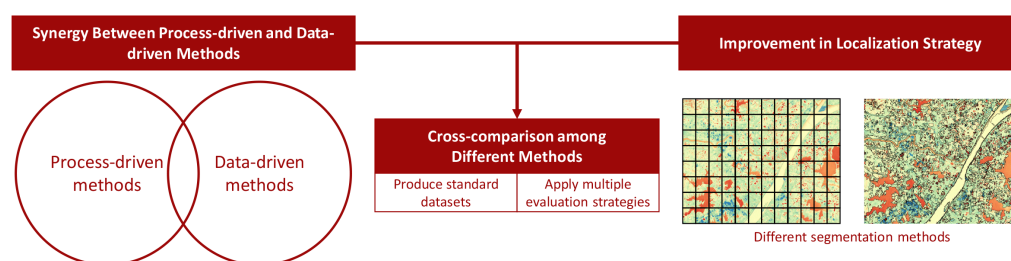
Figure 3. Key future directions for LST resolution enhancement.

4.1. Synergy between Process-Driven and Data-Driven Methods

As reviewed in Section 2, most existing resolution enhancement methods can be divided into two groups: process-driven and data-driven. Process-driven methods, such as the aforementioned DTC/ATC models, use systematic physical theories to depict the temperature variations at certain scales. In contrast, data-driven methods rely on statistical modeling and machine learning algorithms to generate enhanced LSTs by assimilating the observed data into the modeling system. The process-driven and data-driven methods have been considered as two different kinds of enhancement methods with distinct scientific paradigms. However, these two groups of methods are complementary in practice. Process-driven methods offer potential extrapolation beyond the observations, whereas data-driven methods can flexibly adapt to data [104]. In recent years, the synergy between the two groups of methods has gained extra attention in the field of earth system science [105].

Therefore, we argue that the combination of process-driven and data-driven methods could open a promising door for LST resolution enhancement.

The synergy between process-driven and data-driven methods deserves to be recommended for LST resolution enhancement. This synergy may be achieved in the following three main ways: (1) Data-driven parameterization of process-driven methods. Process-driven methods generally require parameters that cannot be easily derived at different scales for LST enhancement tasks. However, data-driven modeling could realize parameterization by using the statistical relationships between indirect observations and model parameters. For example, the simple quadratic functions were employed to relate the DTC parameters (i.e., daily mean temperature and thermal inertia) and the ATC parameters (i.e., annual mean LST and yearly amplitude) to the scaling factors (i.e., NDVI and albedo) at a coarse resolution [30]. The functions were then used to derive DTC/ATC parameters at the fine resolution and generate temporally continuous LST at the same resolution. (2) Process-guided design of data-driven methods. When the physical mechanisms and processes are taken into account, then certain data-driven methods could be specifically designed for LST enhancement. These models can capture the LST variations at different scales. For example, the regression kernels, which can represent the physical property and reflect ground moisture content and thermal inertia of the mixed pixels at their original resolutions, are highly recommended and could lead to better enhancement accuracy by spatial resolution enhancement methods [10]. Another example from temporal resolution enhancement method category was to incorporate the conversion coefficients derived by temperature cycle models into the proposed BLEST. This method fused the LST products from Landsat, MODIS, and FY-2F, and generated LST series at one hour intervals and 100 m resolution [22]. (3) Combining the process-driven and data-driven methods. Combinations of multiple models derived from different perspectives often produce more accurate results than a single model would [106]. For instance, Xia et al. [19] utilized the enhanced results from a kernel-driven enhancement procedure and a fusion-based procedure. The enhanced LST data, which combined the results from the two procedures using weights calculated from error estimations, showed higher robustness and more spatial details than the results from one single procedure. Notably, a simple interpolation based on the autocorrelation among measured samples could mitigate the block effect and reduce the biases during the resolution enhancement [22,28]. Overall, the combination of process-driven and data-driven methods may improve the enhancement accuracy and boost physical understanding of the thermal radiative process and its corresponding spatiotemporal variations.

4.2. Cross-Comparison among Different Methods

In recent decades, we have witnessed the rapid development of resolution enhancement methods for LST. The emergence of newly available LST datasets and novel image processing tools have generated a step change in our ability to capture the potential association between environmental variables and LSTs, which has led to a substantial improvement in enhancement accuracy. However, due to the differences in data, study areas, and evaluation strategies, the optimal choice of the best resolution enhancement method and its associated combinations over diverse land surfaces and weather conditions remains unclear and even controversial. To cope with this issue, a cross-comparison between different enhancement methods must be performed to help users make a better selection for their specific cases, especially from a global perspective.

Specifically, we recommend that cross-comparisons are conducted in the following two main ways: (1) Producing standard datasets. Standard datasets should be introduced to allow participants to compare the state-of-the-art enhancement methods on the same LST resolution enhancement task readily. Diverse standard datasets could be designed to represent heterogeneous land surface and thermal radiative conditions. In the field of land use and land cover classification of hyperspectral data, several well-validated public datasets, such as Pavia and Indian Pines datasets, have been made available to the scientific community. They, along with annotations to evaluate the model performances, have

boosted the model cross-comparison among researchers [107,108]. Such standard datasets could be developed from the existing studies to test the performance of different LST resolution enhancement methods. Specifically, the standard datasets should also be properly cropped, validated, and archived in a way that is easy to retrieve and process. (2) Applying multiple evaluation strategies. As reviewed in Section 3.2, there is no widely accepted assessment procedure for LST enhancement methods. In fact, using only one single metric to reflect different aspects of enhancement errors is highly challenging. An alternative possible solution is to design multiple evaluation strategies to test the relative strengths and weaknesses of the proposed enhancement methods, such as the above-mentioned evaluation strategies [19,22]. Specifically, developing an assessment metric with multiple steps [21] may improve uncertainty analysis and could be used as a starting point when attempting to design a future accuracy assessment method and procedure for LST resolution enhancement.

4.3. Improvement in Localization Strategy

Like other environmental parameters, LST varies in space and time and shows spatiotemporal autocorrelation at different scales. Therefore, the enhancement methods usually utilize neighboring information to improve the enhancement performance, such as using the moving-window regression in spatial resolution enhancement methods [62] and using the weighting procedures in fusion-based temporal resolution enhancement methods [17]. However, the effectiveness of the localization strategy remains ambiguous and might be less adaptive under complex surface conditions.

The practicability of the localization strategy significantly relies on the spatiotemporal dependence among neighboring pixels. Although it is natural to assume that nearby objects have similar thermal characteristics, additional work is required to improve the generalization and applicability of the localized enhancement methods. Specifically, two essential aspects should be considered when applying various localization strategies. First, both the spatiotemporal similarity and the spatiotemporal heterogeneity of the LST distributions need to be taken into account. Accordingly, designing a localization strategy from spatial, temporal, and spectral perspectives is considered to be a good choice. Furthermore, auxiliary data could be included to capture the spatiotemporal characteristics of LSTs, such as the topographic [51] and land cover information [93]. The determination of spatiotemporally localized similarity and heterogeneity needs further exploration [109]. The other aspect is the computational complexity, which is remarkably greater when establishing of geographically and temporally related LST enhancement methods. The efficiency must be considered in practice when scaling up the enhancement methods to regional/global scales or long-length time series. Advanced high-performance cloud computing platforms, such as Google Earth Engine [110] and PIE-Engine (https://engine.piesat.cn/, accessed on 24 March 2021), could potentially reduce the large computational cost of localized LST enhancement.

5. Conclusions

Obtaining LST data at concurrently high spatiotemporal resolutions is still difficult due to the trade-off between spatial and temporal resolutions in remote sensing. As a result, various methods have been proposed to enhance the resolution of LST data. In this paper, a systematical review of existing LST resolution enhancement methods and their quality assessment strategies are presented. Three groups of enhancement methods—spatial resolution enhancement, temporal resolution enhancement, and simultaneous spatiotemporal resolution enhancement—are discussed, and the investigation results showed that LST resolution enhancement methods have considerably advanced in recent decades. The commonly used reference data and assessment metrics for LST resolution enhancement methods and their advantages and disadvantages have also been reviewed. Finally, several new insights into LST resolution enhancement research that could be undertaken by future studies are discussed, including synergy between process-driven and data-driven

methods, cross-comparison among different methods, and improvement in localization strategy.

References

1. Shivers, S.W.; Roberts, D.A.; McFadden, J.P. Using Paired Thermal and Hyperspectral Aerial Imagery to Quantify Land Surface Temperature Variability and Assess Crop Stress within California Orchards. *Remote Sens. Environ.* **2019**, *222*, 215–231. [CrossRef]
2. Chuvieco, E.; Mouillot, F.; van der Werf, G.R.; San Miguel, J.; Tanase, M.; Koutsias, N.; García, M.; Yebra, M.; Padilla, M.; Gitas, I.; et al. Historical Background and Current Developments for Mapping Burned Area from Satellite Earth Observation. *Remote Sens. Environ.* **2019**, *225*, 45–64. [CrossRef]
3. Peng, J.; Qiao, R.; Liu, Y.; Blaschke, T.; Li, S.; Wu, J.; Xu, Z.; Liu, Q. A Wavelet Coherence Approach to Prioritizing Influencing Factors of Land Surface Temperature and Associated Research Scales. *Remote Sens. Environ.* **2020**, *246*, 111866. [CrossRef]
4. Duan, S.B.; Li, Z.L.; Li, H.; Göttsche, F.M.; Wu, H.; Zhao, W.; Leng, P.; Zhang, X.; Coll, C. Validation of Collection 6 MODIS Land Surface Temperature Product Using in Situ Measurements. *Remote Sens. Environ.* **2019**, *225*, 16–29. [CrossRef]
5. Zhu, Z.; Wulder, M.A.; Roy, D.P.; Woodcock, C.E.; Hansen, M.C.; Radeloff, V.C.; Healey, S.P.; Schaaf, C.; Hostert, P.; Strobl, P.; et al. Benefits of the Free and Open Landsat Data Policy. *Remote Sens. Environ.* **2019**, *224*, 382–385. [CrossRef]
6. Campbell, J.B.; Wynne, R.H. *Introduction to Remote Sensing*, 5th ed.; The Guilford Press: New York, NY, USA, 2011.
7. Zhu, X.; Cai, F.; Tian, J.; Williams, T. Spatiotemporal Fusion of Multisource Remote Sensing Data: Literature Survey, Taxonomy, Principles, Applications, and Future Directions. *Remote Sens.* **2018**, *10*, 527. [CrossRef]
8. Kustas, W.P.; Norman, J.M.; Anderson, M.C.; French, A.N. Estimating Subpixel Surface Temperatures and Energy Fluxes from the Vegetation Index–Radiometric Temperature Relationship. *Remote Sens. Environ.* **2003**, *85*, 429–440. [CrossRef]
9. Agam, N.; Kustas, W.P.; Anderson, M.C.; Li, F.; Neale, C.M. A Vegetation Index Based Technique for Spatial Sharpening of Thermal Imagery. *Remote Sens. Environ.* **2007**, *107*, 545–558. [CrossRef]
10. Zhan, W.; Chen, Y.; Zhou, J.; Wang, J.; Liu, W.; Voogt, J.; Zhu, X.; Quan, J.; Li, J. Disaggregation of Remotely Sensed Land Surface Temperature: Literature Survey, Taxonomy, Issues, and Caveats. *Remote Sens. Environ.* **2013**, *131*, 119–139. [CrossRef]
11. Duan, S.B.; Li, Z.L. Spatial Downscaling of MODIS Land Surface Temperatures Using Geographically Weighted Regression: Case Study in Northern China. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6458–6469. [CrossRef]
12. Hutengs, C.; Vohland, M. Downscaling Land Surface Temperatures at Regional Scales with Random Forest Regression. *Remote Sens. Environ.* **2016**, *178*, 127–141. [CrossRef]
13. Bartkowiak, P.; Castelli, M.; Notarnicola, C. Downscaling Land Surface Temperature from MODIS Dataset with Random Forest Approach over Alpine Vegetated Areas. *Remote Sens.* **2019**, *11*, 1319. [CrossRef]
14. Gao, F.; Kustas, W.; Anderson, M. A Data Mining Approach for Sharpening Thermal Satellite Imagery over Land. *Remote Sens.* **2012**, *4*, 3287–3319. [CrossRef]
15. Chen, X.; Li, W.; Chen, J.; Zhan, W.; Rao, Y. A Simple Error Estimation Method for Linear-Regression-Based Thermal Sharpening Techniques with the Consideration of Scale Difference. *Geo-Spat. Inf. Sci.* **2014**, *17*, 54–59. [CrossRef]
16. Xia, H.; Chen, Y.; Quan, J.; Li, J. Object-Based Window Strategy in Thermal Sharpening. *Remote Sens.* **2019**, *11*, 634. [CrossRef]
17. Wu, P.; Shen, H.; Zhang, L.; Göttsche, F.M. Integrated Fusion of Multi-Scale Polar-Orbiting and Geostationary Satellite Observations for the Mapping of High Spatial and Temporal Resolution Land Surface Temperature. *Remote Sens. Environ.* **2015**, *156*, 169–181. [CrossRef]
18. Xia, H.; Chen, Y.; Zhao, Y.; Chen, Z. "Regression-Then-Fusion" or "Fusion-Then-Regression"? A Theoretical Analysis for Generating High Spatiotemporal Resolution Land Surface Temperatures. *Remote Sens.* **2018**, *10*, 1382. [CrossRef]
19. Xia, H.; Chen, Y.; Li, Y.; Quan, J. Combining Kernel-Driven and Fusion-Based Methods to Generate Daily High-Spatial-Resolution Land Surface Temperatures. *Remote Sens. Environ.* **2019**, *224*, 259–274. [CrossRef]
20. Gao, L.; Zhan, W.; Huang, F.; Quan, J.; Lu, X.; Wang, F.; Ju, W.; Zhou, J. Localization or Globalization? Determination of the Optimal Regression Window for Disaggregation of Land Surface Temperature. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 477–490. [CrossRef]
21. Gao, L.; Zhan, W.; Huang, F.; Zhu, X.; Zhou, J.; Quan, J.; Du, P.; Li, M. Disaggregation of Remotely Sensed Land Surface Temperature: A Simple yet Flexible Index (SIFI) to Assess Method Performances. *Remote Sens. Environ.* **2017**, *200*, 206–219. [CrossRef]

22. Quan, J.; Zhan, W.; Ma, T.; Du, Y.; Guo, Z.; Qin, B. An Integrated Model for Generating Hourly Landsat-like Land Surface Temperatures over Heterogeneous Landscapes. *Remote Sens. Environ.* **2018**, *206*, 403–423. [CrossRef]

23. Weng, Q.; Fu, P.; Gao, F. Generating Daily Land Surface Temperature at Landsat Resolution by Fusing Landsat and MODIS Data. *Remote Sens. Environ.* **2014**, *145*, 55–67. [CrossRef]

24. Weng, Q.; Fu, P. Modeling Diurnal Land Temperature Cycles over Los Angeles Using Downscaled GOES Imagery. *ISPRS J. Photogramm. Remote Sens.* **2014**, *97*, 78–88. [CrossRef]

25. Zhan, W.; Chen, Y.; Zhou, J.; Li, J.; Liu, W. Sharpening Thermal Imageries: A Generalized Theoretical Framework From an Assimilation Perspective. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 773–789. [CrossRef]

26. Metz, M.; Andreo, V.; Neteler, M. A New Fully Gap-Free Time Series of Land Surface Temperature from MODIS LST Data. *Remote Sens.* **2017**, *9*, 1333. [CrossRef]

27. Zhu, X.; Helmer, E.H.; Gao, F.; Liu, D.; Chen, J.; Lefsky, M.A. A Flexible Spatiotemporal Method for Fusing Satellite Images with Different Resolutions. *Remote Sens. Environ.* **2016**, *172*, 165–177. [CrossRef]

28. Chen, X.; Li, W.; Chen, J.; Rao, Y.; Yamaguchi, Y. A Combination of TsHARP and Thin Plate Spline Interpolation for Spatial Sharpening of Thermal Imagery. *Remote Sens.* **2014**, *6*, 2845–2863. [CrossRef]

29. Wang, J.W.; Chow, W.T.L.; Wang, Y.C. A Global Regression Method for Thermal Sharpening of Urban Land Surface Temperatures from MODIS and Landsat. *Int. J. Remote Sens.* **2020**, *41*, 2986–3009. [CrossRef]

30. Zhan, W.; Huang, F.; Quan, J.; Zhu, X.; Gao, L.; Zhou, J.; Ju, W. Disaggregation of Remotely Sensed Land Surface Temperature: A New Dynamic Methodology. *J. Geophys. Res. Atmos.* **2016**, *121*, 10538–10554. [CrossRef]

31. Burdun, I.; Bechtold, M.; Sagris, V.; Komisarenko, V.; De Lannoy, G.; Mander, Ü. A Comparison of Three Trapezoid Models Using Optical and Thermal Satellite Imagery for Water Table Depth Monitoring in Estonian Bogs. *Remote Sens.* **2020**, *12*, 1980. [CrossRef]

32. Zawadzka, J.; Corstanje, R.; Harris, J.; Truckell, I. Downscaling Landsat-8 Land Surface Temperature Maps in Diverse Urban Landscapes Using Multivariate Adaptive Regression Splines and Very High Resolution Auxiliary Data. *Int. J. Digit. Earth* **2020**, *13*, 899–914. [CrossRef]

33. Agathangelidis, I.; Cartalis, C. Improving the Disaggregation of MODIS Land Surface Temperatures in an Urban Environment: A Statistical Downscaling Approach Using High-Resolution Emissivity. *Int. J. Remote Sens.* **2019**, *40*, 5261–5286. [CrossRef]

34. Merlin, O.; Duchemin, B.; Hagolle, O.; Jacob, F.; Coudert, B.; Chehbouni, G.; Dedieu, G.; Garatuza, J.; Kerr, Y. Disaggregation of MODIS Surface Temperature over an Agricultural Area Using a Time Series of Formosat-2 Images. *Remote Sens. Environ.* **2010**, *114*, 2500–2512. [CrossRef]

35. Dong, P.; Gao, L.; Zhan, W.; Liu, Z.; Li, J.; Lai, J.; Li, H.; Huang, F.; Tamang, S.K.; Zhao, L. Global Comparison of Diverse Scaling Factors and Regression Models for Downscaling Landsat-8 Thermal Data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 44–56. [CrossRef]

36. Lillo-Saavedra, M.; García-Pedrero, A.; Merino, G.; Gonzalo-Martín, C. TS2uRF: A New Method for Sharpening Thermal Infrared Satellite Imagery. *Remote Sens.* **2018**, *10*, 249. [CrossRef]

37. Stathopoulou, M.; Cartalis, C. Downscaling AVHRR Land Surface Temperatures for Improved Surface Urban Heat Island Intensity Estimation. *Remote Sens. Environ.* **2009**, *113*, 2592–2605. [CrossRef]

38. Amazirh, A.; Merlin, O.; Er-Raki, S. Including Sentinel-1 Radar Data to Improve the Disaggregation of MODIS Land Surface Temperature Data. *ISPRS J. Photogramm. Remote Sens.* **2019**, *150*, 11–26. [CrossRef]

39. Maeda, E.E. Downscaling MODIS LST in the East African Mountains Using Elevation Gradient and Land-Cover Information. *Int. J. Remote Sens.* **2014**, *35*, 3094–3108. [CrossRef]

40. Zhang, X.; Zhao, H.; Yang, J. Spatial Downscaling of Land Surface Temperature in Combination with TVDI and Elevation. *Int. J. Remote Sens.* **2019**, *40*, 1875–1886. [CrossRef]

41. Liu, K.; Wang, S.; Li, X.; Li, Y.; Zhang, B.; Zhai, R. The Assessment of Different Vegetation Indices for Spatial Disaggregating of Thermal Imagery over the Humid Agricultural Region. *Int. J. Remote Sens.* **2020**, *41*, 1907–1926. [CrossRef]

42. Zhou, J.; Liu, S.; Li, M.; Zhan, W.; Xu, Z.; Xu, T. Quantification of the Scale Effect in Downscaling Remotely Sensed Land Surface Temperature. *Remote Sens.* **2016**, *8*, 975. [CrossRef]

43. Dominguez, A.; Kleissl, J.; Luvall, J.C.; Rickman, D.L. High-Resolution Urban Thermal Sharpener (HUTS). *Remote Sens. Environ.* **2011**, *115*, 1772–1780. [CrossRef]

44. Inamdar, A.K.; French, A.; Hook, S.; Vaughan, G.; Luckett, W. Land Surface Temperature Retrieval at High Spatial and Temporal Resolutions over the Southwestern United States. *J. Geophys. Res. Atmos.* **2008**, *113*. [CrossRef]

45. Bechtel, B.; Zakšek, K.; Hoshyaripour, G. Downscaling Land Surface Temperature in an Urban Area: A Case Study for Hamburg, Germany. *Remote Sens.* **2012**, *4*, 3184–3200. [CrossRef]

46. Chen, X.; Yamaguchi, Y.; Chen, J.; Shi, Y. Scale Effect of Vegetation-Index-Based Spatial Sharpening for Thermal Imagery: A Simulation Study by ASTER Data. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 549–553. [CrossRef]

47. Bindhu, V.; Narasimhan, B.; Sudheer, K. Development and Verification of a Non-Linear Disaggregation Method (NL-DisTrad) to Downscale MODIS Land Surface Temperature to the Spatial Scale of Landsat Thermal Data to Estimate Evapotranspiration. *Remote Sens. Environ.* **2013**, *135*, 118–129. [CrossRef]

48. Ghosh, A.; Joshi, P.K. Hyperspectral Imagery for Disaggregation of Land Surface Temperature with Selected Regression Algorithms over Different Land Use Land Cover Scenes. *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 76–93. [CrossRef]

49. Cho, K.; Kim, Y.; Kim, Y. Disaggregation of Landsat-8 Thermal Data Using Guided SWIR Imagery on the Scene of a Wildfire. *Remote Sens.* **2018**, *10*, 105. [CrossRef]

50. Pan, X.; Zhu, X.; Yang, Y.; Cao, C.; Zhang, X.; Shan, L. Applicability of Downscaling Land Surface Temperature by Using Normalized Difference Sand Index. *Sci. Rep.* **2018**, *8*, 9530. [CrossRef]

51. Wang, R.; Gao, W.; Peng, W. Downscale MODIS Land Surface Temperature Based on Three Different Models to Analyze Surface Urban Heat Island: A Case Study of Hangzhou. *Remote Sens.* **2020**, *12*, 2134. [CrossRef]

52. Zakšek, K.; Oštir, K. Downscaling Land Surface Temperature for Urban Heat Island Diurnal Cycle Analysis. *Remote Sens. Environ.* **2012**, *117*, 114–124. [CrossRef]

53. Jiang, Y.; Fu, P.; Weng, Q. Downscaling GOES Land Surface Temperature for Assessing Heat Wave Health Risks. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1605–1609. [CrossRef]

54. Sismanidis, P.; Keramitsoglou, I.; Kiranoudis, C.T.; Bechtel, B. Assessing the Capability of a Downscaled Urban Land Surface Temperature Time Series to Reproduce the Spatiotemporal Features of the Original Data. *Remote Sens.* **2016**, *8*, 274. [CrossRef]

55. Yang, G.; Pu, R.; Zhao, C.; Huang, W.; Wang, J. Estimation of Subpixel Land Surface Temperature Using an Endmember Index Based Technique: A Case Examination on ASTER and MODIS Temperature Products over a Heterogeneous Area. *Remote Sens. Environ.* **2011**, *115*, 1202–1219. [CrossRef]

56. Yang, Y.; Cao, C.; Pan, X.; Li, X.; Zhu, X. Downscaling Land Surface Temperature in an Arid Area by Using Multiple Remote Sensing Indices with Random Forest Regression. *Remote Sens.* **2017**, *9*, 789. [CrossRef]

57. Yang, G.; Pu, R.; Huang, W.; Wang, J.; Zhao, C. A Novel Method to Estimate Subpixel Temperature by Fusing Solar-Reflective and Thermal-Infrared Remote-Sensing Data With an Artificial Neural Network. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2170–2178. [CrossRef]

58. Bonafoni, S. Downscaling of Landsat and MODIS Land Surface Temperature Over the Heterogeneous Urban Area of Milan. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2019–2027. [CrossRef]

59. Peng, Y.; Li, W.; Luo, X.; Li, H. A Geographically and Temporally Weighted Regression Model for Spatial Downscaling of MODIS Land Surface Temperatures Over Urban Heterogeneous Regions. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5012–5027. [CrossRef]

60. Sismanidis, P.; Keramitsoglou, I.; Bechtel, B.; Kiranoudis, C.T. Improving the Downscaling of Diurnal Land Surface Temperatures Using the Annual Cycle Parameters as Disaggregation Kernels. *Remote Sens.* **2017**, *9*, 23. [CrossRef]

61. Keramitsoglou, I.; Kiranoudis, C.T.; Weng, Q. Downscaling Geostationary Land Surface Temperature Imagery for Urban Analysis. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1253–1257. [CrossRef]

62. Jeganathan, C.; Hamm, N.; Mukherjee, S.; Atkinson, P.; Raju, P.; Dadhwal, V. Evaluating a Thermal Image Sharpening Model over a Mixed Agricultural Landscape in India. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 178–191. [CrossRef]

63. Li, W.; Ni, L.; Li, Z.; Duan, S.; Wu, H. Evaluation of Machine Learning Algorithms in Spatial Downscaling of MODIS Land Surface Temperature. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2299–2307. [CrossRef]

64. Kolios, S.; Georgoulas, G.; Stylios, C. Achieving Downscaling of Meteosat Thermal Infrared Imagery Using Artificial Neural Networks. *Int. J. Remote Sens.* **2013**, *34*, 7706–7722. [CrossRef]

65. Ebrahimy, H.; Azadbakht, M. Downscaling MODIS Land Surface Temperature over a Heterogeneous Area: An Investigation of Machine Learning Techniques, Feature Selection, and Impacts of Mixed Pixels. *Comput. Geosci.* **2019**, *124*, 93–102. [CrossRef]

66. Wu, H.; Li, W. Downscaling Land Surface Temperatures Using a Random Forest Regression Model With Multitype Predictor Variables. *IEEE Access* **2019**, *7*, 21904–21916. [CrossRef]

67. Zhu, X.; Tuia, D.; Mou, L.; Xia, G.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [CrossRef]

68. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep Learning in Remote Sensing Applications: A Meta-Analysis and Review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *152*, 166–177. [CrossRef]

69. Marceau, D.J.; Hay, G.J. Remote Sensing Contributions to the Scale Issue. *Can. J. Remote Sens.* **1999**, *25*, 357–366. [CrossRef]

70. Mukherjee, S.; Joshi, P.K.; Garg, R.D. A Comparison of Different Regression Models for Downscaling Landsat and MODIS Land Surface Temperature Images over Heterogeneous Landscape. *Adv. Space Res.* **2014**, *54*, 655–669. [CrossRef]

71. Zhukov, B.; Oertel, D.; Lanzl, F.; Reinhackel, G. Unmixing-Based Multisensor Multiresolution Image Fusion. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1212–1226. [CrossRef]

72. Zhan, W.; Chen, Y.; Wang, J.; Zhou, J.; Quan, J.; Liu, W.; Li, J. Downscaling Land Surface Temperatures with Multi-Spectral and Multi-Resolution Images. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 23–36. [CrossRef]

73. Göttsche, F.M.; Olesen, F.S. Modelling the Effect of Optical Thickness on Diurnal Cycles of Land Surface Temperature. *Remote Sens. Environ.* **2009**, *113*, 2306–2316. [CrossRef]

74. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model for Complex Heterogeneous Regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [CrossRef]

75. Göttsche, F.M.; Olesen, F.S. Modelling of Diurnal Cycles of Brightness Temperature Extracted from METEOSAT Data. *Remote Sens. Environ.* **2001**, *76*, 337–348. [CrossRef]

76. Huang, F.; Zhan, W.; Duan, S.B.; Ju, W.; Quan, J. A Generic Framework for Modeling Diurnal Land Surface Temperatures with Remotely Sensed Thermal Observations under Clear Sky. *Remote Sens. Environ.* **2014**, *150*, 140–151. [CrossRef]

77. Liu, X.; Tang, B.H.; Li, Z.L.; Zhou, C.; Wu, W.; Rasmussen, M.O. An Improved Method for Separating Soil and Vegetation Component Temperatures Based on Diurnal Temperature Cycle Model and Spatial Correlation. *Remote Sens. Environ.* **2020**, *248*, 111979. [CrossRef]

78. Duan, S.B.; Li, Z.L.; Wang, N.; Wu, H.; Tang, B.H. Evaluation of Six Land-Surface Diurnal Temperature Cycle Models Using Clear-Sky in Situ and Satellite Data. *Remote Sens. Environ.* **2012**, *124*, 15–25. [CrossRef]

79. Jin, M.; Dickinson, R.E. Interpolation of Surface Radiative Temperature Measured from Polar Orbiting Satellites to a Diurnal Cycle: 1. Without Clouds. *J. Geophys. Res. Atmos.* **1999**, *104*, 2105–2116. [CrossRef]

80. Sun, D.; Pinker, R.T. Implementation of GOES-based Land Surface Temperature Diurnal Cycle to AVHRR. *Int. J. Remote Sens.* **2005**, *26*, 3975–3984. [CrossRef]

81. Duan, S.B.; Li, Z.L.; Wu, H.; Tang, B.H.; Jiang, X.; Zhou, G. Modeling of Day-to-Day Temporal Progression of Clear-Sky Land Surface Temperature. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1050–1054. [CrossRef]

82. Weng, Q.; Fu, P. Modeling Annual Parameters of Clear-Sky Land Surface Temperature Variations and Evaluating the Impact of Cloud Cover Using Time Series of Landsat TIR Data. *Remote Sens. Environ.* **2014**, *140*, 267–278. [CrossRef]

83. Zhan, W.; Zhou, J.; Ju, W.; Li, M.; Sandholt, I.; Voogt, J.; Yu, C. Remotely Sensed Soil Temperatures beneath Snow-Free Skin-Surface Using Thermal Observations from Tandem Polar-Orbiting Satellites: An Analytical Three-Time-Scale Model. *Remote Sens. Environ.* **2014**, *143*, 1–14. [CrossRef]

84. Stine, A.R.; Huybers, P.; Fung, I.Y. Changes in the Phase of the Annual Cycle of Surface Temperature. *Nature* **2009**, *457*, 435–440. [CrossRef] [PubMed]

85. Quan, J.; Chen, Y.; Zhan, W.; Wang, J.; Voogt, J.; Li, J. A Hybrid Method Combining Neighborhood Information from Satellite Data with Modeled Diurnal Temperature Cycles over Consecutive Days. *Remote Sens. Environ.* **2014**, *155*, 257–274. [CrossRef]

86. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1301–1312. [CrossRef]

87. Schmitt, M.; Zhu, X.X. Data Fusion and Remote Sensing: An Ever-Growing Relationship. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 6–23. [CrossRef]

88. Pohl, C.; van Genderen, J. Remote Sensing Image Fusion: An Update in the Context of Digital Earth. *Int. J. Digit. Earth* **2014**, *7*, 158–172. [CrossRef]

89. Belgiu, M.; Stein, A. Spatiotemporal Image Fusion in Remote Sensing. *Remote Sens.* **2019**, *11*, 818. [CrossRef]

90. Liu, H.; Weng, Q. Enhancing Temporal Resolution of Satellite Imagery for Public Health Studies: A Case Study of West Nile Virus Outbreak in Los Angeles in 2007. *Remote Sens. Environ.* **2012**, *117*, 57–71. [CrossRef]

91. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [CrossRef]

92. Wu, M.; Niu, Z.; Wang, C.; Wu, C.; Wang, L. Use of MODIS and Landsat Time Series Data to Generate High-Resolution Temporal Synthetic Landsat Data Using a Spatial and Temporal Reflectance Fusion Model. *J. Appl. Remote Sens.* **2012**, *6*, 063507. [CrossRef]

93. Wu, M.; Li, H.; Huang, W.; Niu, Z.; Wang, C. Generating Daily High Spatial Land Surface Temperatures by Combining ASTER and MODIS Land Surface Temperature Products for Environmental Process Monitoring. *Environ. Sci. Process. Impacts* **2015**, *17*, 1396–1404. [CrossRef] [PubMed]

94. Wu, P.; Shen, H.; Ai, T.; Liu, Y. Land-Surface Temperature Retrieval at High Spatial and Temporal Resolutions Based on Multi-Sensor Fusion. *Int. J. Digit. Earth* **2013**, *6*, 113–133. [CrossRef]

95. Moosavi, V.; Talebi, A.; Mokhtari, M.H.; Shamsi, S.R.F.; Niazi, Y. A Wavelet-Artificial Intelligence Fusion Approach (WAIFA) for Blending Landsat and MODIS Surface Temperature. *Remote Sens. Environ.* **2015**, *169*, 243–254. [CrossRef]

96. Shao, Z.; Cai, J.; Fu, P.; Hu, L.; Liu, T. Deep Learning-Based Fusion of Landsat-8 and Sentinel-2 Images for a Harmonized Surface Reflectance Product. *Remote Sens. Environ.* **2019**, *235*, 111425. [CrossRef]

97. Bai, Y.; Wong, M.; Shi, W.Z.; Wu, L.X.; Qin, K. Advancing of Land Surface Temperature Retrieval Using Extreme Learning Machine and Spatio-Temporal Adaptive Data Fusion Algorithm. *Remote Sens.* **2015**, *7*, 4424–4441. [CrossRef]

98. Zhou, J.; Chen, Y.; Zhang, X.; Zhan, W. Modelling the Diurnal Variations of Urban Heat Islands with Multi-Source Satellite Data. *Int. J. Remote Sens.* **2013**, *34*, 7568–7588. [CrossRef]

99. Wang, W.; Liang, S.; Meyers, T. Validating MODIS Land Surface Temperature Products Using Long-Term Nighttime Ground Measurements. *Remote Sens. Environ.* **2008**, *112*, 623–635. [CrossRef]

100. Zurita-Milla, R.; Kaiser, G.; Clevers, J.G.P.W.; Schneider, W.; Schaepman, M.E. Downscaling Time Series of MERIS Full Resolution Data to Monitor Vegetation Seasonal Dynamics. *Remote Sens. Environ.* **2009**, *113*, 1874–1885. [CrossRef]

101. Rodriguez-Galiano, V.; Pardo-Iguzquiza, E.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Downscaling Landsat 7 ETM+ Thermal Imagery Using Land Surface Temperature and NDVI Images. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 515–527. [CrossRef]

102. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef] [PubMed]

103. Wang, Z.; Bovik, A.C. A Universal Image Quality Index. *IEEE Signal Process. Lett.* **2002**, *9*, 81–84. [CrossRef]

104. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature* **2019**, *566*, 195–204. [CrossRef]

105. Steffen, W.; Richardson, K.; Rockström, J.; Schellnhuber, H.J.; Dube, O.P.; Dutreuil, S.; Lenton, T.M.; Lubchenco, J. The Emergence and Evolution of Earth System Science. *Nat. Rev. Earth Environ.* **2020**, *1*, 54–63. [CrossRef]
106. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
107. Audebert, N.; Saux, B.L.; Lefevre, S. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 159–173. [CrossRef]
108. Vali, A.; Comai, S.; Matteucci, M. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sens.* **2020**, *12*, 2495. [CrossRef]
109. Murphy, J.M.; Maggioni, M. Unsupervised Clustering and Active Learning of Hyperspectral Images With Nonlinear Diffusion. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 1829–1845. [CrossRef]
110. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]