



Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature



Penghai Wu^a, Huanfeng Shen^{b,*}, Liangpei Zhang^{c,*}, Frank-Michael Göttsche^d

^a School of Resources and Environmental Engineering, Anhui University, Hefei, China

^b School of Resource and Environmental Sciences, Wuhan University, Wuhan, China

^c The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

^d Institute of Meteorology and Climate Research (IMK), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

ARTICLE INFO

Article history:

Received 8 March 2014

Received in revised form 8 September 2014

Accepted 8 September 2014

Available online xxx

Keywords:

Integrated fusion

Land surface temperature

Multi-scale

Geostationary satellite

Polar-orbiting satellite

Resolution

ABSTRACT

Land surface temperature (LST) and its diurnal variation are important when evaluating climate change, the land–atmosphere energy budget, and the global hydrological cycle. However, the available satellite LST products have either a coarse spatial resolution or a low temporal resolution, which constrains their potential applications. This paper proposes a spatio-temporal integrated temperature fusion model (STITFM) for the retrieval of high temporal and spatial resolution LST from multi-scale polar-orbiting and geostationary satellite observations. Compared with the traditional fusion methods for LST with two different sensors, the proposed method is able to fuse the LST from arbitrary sensors in a unified framework. The model was tested using LST with fine, moderate, and coarse-resolutions. Data from the Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) sensors, the Moderate Resolution Imaging Spectroradiometer (MODIS), the Geostationary Operational Environmental Satellite (GOES) Imager, and the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI) were used. The fused LST values were evaluated with in situ LST obtained from the Surface Radiation Budget Network (SURFRAD) and the Land Surface Analysis Satellite Application Facility (LSA SAF) project. The final validation results indicate that the STITFM is accurate to within about 2.5 K.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Land surface temperature (LST) is an important parameter of the weather and climate system, controlling surface heat and water exchange with the atmosphere (Anderson, Allen, Morse, & Kustas, 2012; Kalma, McVicar, & McCabe, 2008; Li et al., 2013; Sellers et al., 1997; Tierney et al., 2008). However, the available satellite LST products have either low spatial resolution (geostationary satellites) or low temporal resolution (polar-orbiting satellites), which constrains their potential applications (Inamdar, French, Hook, Vaughan, & Lockett, 2008). Satellite sensors such as the GOES Imager and MSG SEVIRI provide multiple spectral images of the same region of the earth at frequent time intervals (up to 15 min). Diurnal cycles of LST, which constitute an important element of the surface energy budget, can thus be obtained from the geostationary satellites. However, their spatial resolutions (3–5 km) are too coarse for many applications requiring a higher spatial resolution. For example, the use of high temporal/low spatial resolution geostationary platforms has been limited in land-surface hydrological

studies (Kalma et al., 2008). In contrast, polar-orbiting sensors such as Landsat TM/ETM+ and MODIS provide LST with spatial resolutions ranging from about 100 to 1000 m, which allow monitoring of heterogeneous areas with a finer pixel footprint. Nevertheless, the strong diurnal cycles of LST cannot be captured at the temporal resolution (between two views per day and one view every 16 days for MODIS and Landsat, respectively) of the polar-orbiting satellites (Sun, Pinker, & Kafatos, 2006). In other words, it is currently extremely difficult to directly acquire LST with high temporal and spatial resolutions, due to the trade-offs among these resolutions (Zhan et al., 2013).

A number of methods have been proposed to estimate LST with both high temporal and spatial resolutions (see Zhan et al., 2013 for a review). Most of these methods utilize LST from one sensor and other auxiliary data of a better spatial resolution. These auxiliary data may contain some specific physical or ecological parameters, such as topography, normalized difference vegetation index (NDVI), emissivity, soil water content index, and the reflectance of the visible and near-infrared bands (Kustas, Norman, Anderson, & French, 2003; Nichol, 2009; Yang, Pu, Huang, Wang, & Zhao, 2010; Zhan et al., 2012). Only in recent years have we witnessed the emergence of spatio-temporal fusion, using the LST from two different sensors to predict high-resolution LST. For example, Gao, Masek, Schwaller, and Hall (2006) and Zhu,

* Corresponding authors.

E-mail addresses: dbqwph2007@gmail.com (P. Wu), shenhf@whu.edu.cn (H. Shen), zip62@public.wh.hb.cn (L. Zhang), frank.goettsche@kit.edu (F.-M. Göttsche).

Chen, Gao, Chen, and Masek (2010) proposed a spatial and temporal adaptive reflectance fusion model (STARFM) and its enhanced version (ESTARFM) to predict daily Landsat-like reflectance data at the Landsat spatial scale. Emelyanova, McVicar, Van Niel, Li, and van Dijk (2013) proposed a framework for algorithm selection between STARFM and ESTARFM, according to how the variance is driven by the temporal or spatial component. Although these methods were originally proposed for reflectance, they have also opened up a new avenue for other biophysical parameters, such as LST (Liu & Weng, 2012). Kim and Hogue (2012) adopted a previously developed subtraction method that is similar to STARFM, but without the spatial weighting function, to blend Landsat and MODIS LST. Liu and Weng (2012) utilized STARFM to blend ASTER and MODIS LST for public health studies. Huang, Wang, Song, Fu, and Wong (2013) combined bilateral filtering and STARFM to generate high spatio-temporal resolution LST from Landsat and MODIS data for urban heat island monitoring. Weng, Fu, and Gao (2014) improved and modified ESTARFM to predict thermal radiance and LST data by considering the annual temperature cycle (ATC) and urban thermal landscape heterogeneity. These fusion methods have shown that LST products at the Landsat spatial scale can be predicted every day, which considerably extends the potential applications of LST.

Although infrequently used, geostationary thermal sensors hold great potential because of their widespread availability and their higher temporal density (Zhan et al., 2013). Geostationary satellites may be a cost-effective data source to enhance the temporal resolution from daily to hourly. For instance, Inamdar et al. (2008) and Inamdar and French (2009) merged GOES LST, MODIS LST, and MODIS NDVI or emissivity data to yield half-hourly LST values at a 1-km spatial resolution. Using various auxiliary datasets, Zaksek and Ostir (2012) downsampled SEVIRI LST to a 1-km spatial resolution and a temporal resolution of 15 min. Wu, Shen, Ai, and Liu (2013) extended STARFM to blend Landsat TM/ETM+ and GOES LST to retrieve hourly LST at the Landsat TM/ETM+ spatial resolution. However, it is difficult to capture accurate spatio-temporal change information at such large-scale differences, which can cause instability of the local variables. A possible solution named “scale transition” is to add medium-scale sensors to the process of fusion between the fine- and coarse-scale sensors. For instance, in order to aid the passing of change information from GOES (4-km spatial) to Landsat (about 100-m resolution), MODIS data with their moderate resolution (1000-m spatial) can be added to the fusion process of Landsat and GOES.

However, many of the existing methods were developed to fuse data from only two sensors (Huang et al., 2013; Kim & Hogue, 2012; Liu & Weng, 2012; Weng et al., 2014; Wu et al., 2013), and little work has been done on fusing data from three or more sensors. In this study, we present the spatio-temporal integrated temperature fusion model (STITFM) for remote sensed LSTs which blends arbitrary observations from multi-scale polar-orbiting and geostationary satellites. Based on the proposed method, the objective is to fuse the LST data from arbitrary sensors by the use of a unified framework. The fused LSTs are integrated at the highest spatial and temporal resolutions among the input LST satellite data. The performance of the new method was evaluated with field observations from the Surface Radiation Budget Network (SURFRAD) and the Land Surface Analysis Satellite Application Facility (LSA SAF) project.

The next section briefly describes the primary sources of data. STITFM is introduced in Section 3. The experimental results and quantitative evaluations are provided in Section 4. In Section 5 and Section 6, we provide a discussion and concluding remarks.

2. Data

The primary data used in this study comprised in situ LSTs, cloud-free Landsat TM/ETM+ LSTs, Terra MODIS LSTs, GOES Imager LSTs, and MSG SEVIRI LSTs. These are described below.

2.1. In situ LST

In situ LSTs from two geographically different locations characterized by different surface types were used (see Fig. 1). The network of surface radiation measurement sites (SURFRAD) was established in 1993 through the support of the NOAA Office of Global Programs. The SURFRAD stations provide high-quality in situ measurements of up- and downwelling long-wave radiations within the U.S. (Wang & Liang, 2009; Yu et al., 2009). On 1 Jan 2009, data sampling changed from 3-min to 1-min resolution, and the data can be downloaded from the NOAA Global Monitoring Division (GMD) (<http://www.esrl.noaa.gov/gmd/dv/data/>). As in our previous study (Wu et al., 2013), data from the Desert Rock (DRA) site in Nevada (116.02 W, 36.63 N) were used for validating the LST retrieval. The land-cover type in the DRA site is open shrubland. The in situ surface skin temperatures or LSTs were estimated according to the method of Inamdar et al. (2008).

Another set of in situ measurements was acquired from the LST validation site established and maintained by the Karlsruhe Institute of Technology (KIT), Germany. KIT operates permanent validation stations, supported by the LSA SAF project, which were specifically designed for the validation of LST retrieved from TIR satellite measurements, such as the MSG/SEVIRI LST product (Xu, Yu, Tarpley, Gottsche, & Olesen, 2013). For the present study, the Evora site data in Portugal (8.00 W, 38.54 N), covering natural vegetation compounds of dispersed oak and cork trees interlaced with open grassland, were used for validating the LST retrieval. As in the previous studies (Inamdar et al., 2008; Kabsch, Olesen, & Prata, 2008; Sun & Pinker, 2003; Sun, Pinker, & Basara, 2004; Xu et al., 2013), the in situ LSTs were also used to directly evaluate the predictions (or spatial averages) from the fusion of the multi-sensor LST products.

2.2. Landsat TM/ETM+ LST

Landsat images offer the longest continuous global record of the earth's surface, and are a unique resource for global change research and applications in agriculture, cartography, geology, forestry, regional planning, surveillance, education, and national security (Gucluer et al., 2010). Landsat has a minimum 16-day revisit cycle, which is sometimes extended due to cloud contamination. The Landsat TIR channels have a spatial resolution of about 100 m (Landsat TM 120 m, Landsat ETM+ 60 m), and are available from the Next Generation Earth Science Discovery Tool (<http://reverb.echo.nasa.gov/reverb/>). This is an important spatial resolution because it is detailed enough to characterize human-scale processes such as the dynamics of urban growth (Masek, Lindsay, & Goward, 2000) or urban heat islands (Kumak, Bhaskar, & Padmakumari, 2012). We retrieved the Landsat LST using a generalized single-channel method from the Landsat TIR window channel 6 (10.4–12.5 μm), based on the work of Jimenez-Munoz and Sobrino (2003). The details of the Landsat, MODIS, and geostationary satellite data used in this study are summarized in Table 1.

2.3. MODIS 1 km LST

The MODIS sensors on board the Terra and Aqua satellites have 36 spectral bands ranging from the visible to the thermal-infrared spectrum. The spatial resolution varies from 250 m (bands 1 and 2) to 500 m (bands 3 to 7) and 1000 m (bands 8 to 36). The sensors view the entire earth every one or two days. The MODIS LST product was proposed by the MODIS Science Team as a daily daytime and nighttime 1-km global land product, and includes derivative products at lower temporal frequencies and spatial resolutions. In the range from 10 °C to 50 °C, the accuracy of the MODIS daily LST product is better than 1 °C, as validated with in situ measurements collected during field campaigns between 2000 and 2002 (Wan, Zhang, Zhang, & Li, 2004). We used the MODIS LST/Emissivity Daily Level 3 Global 1-km SIN grid product known as MOD11A1, which is available from the Geospatial Data

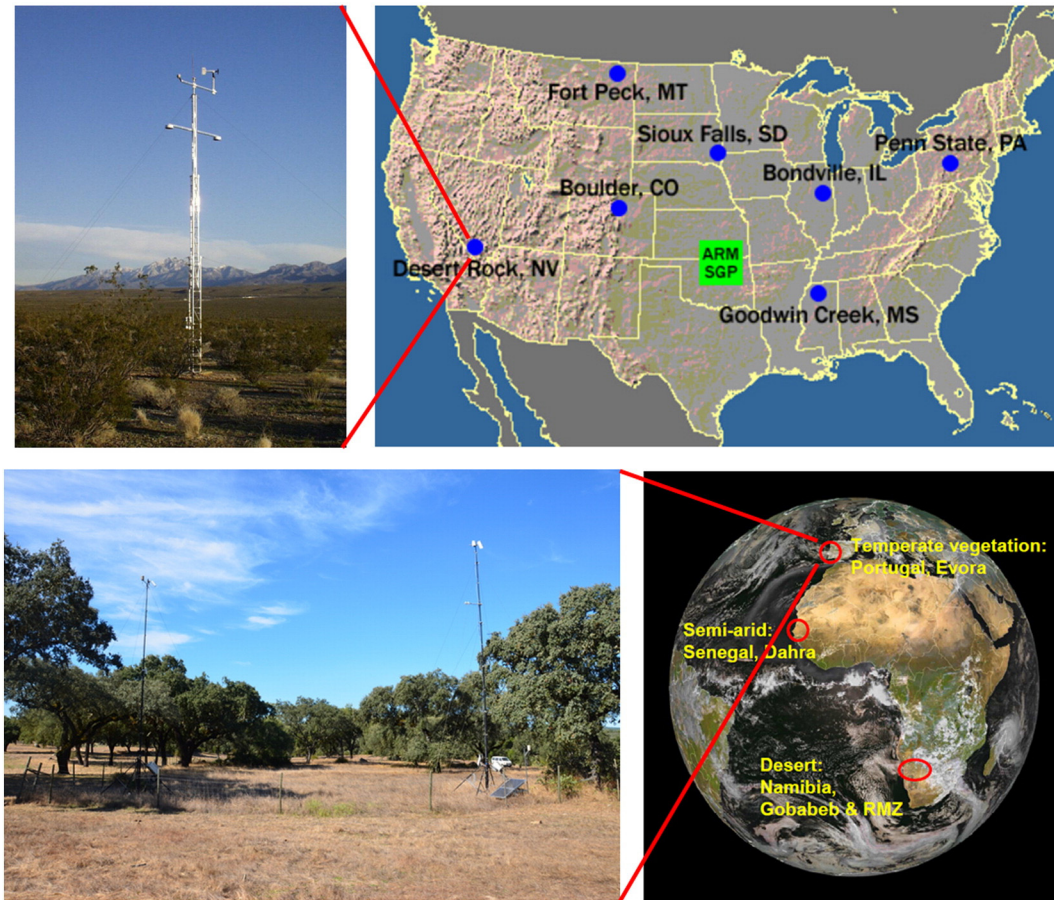


Fig. 1. Locations of SURFRAD's validation stations, and a photo of the DRA site (upper); and locations of the KIT validation stations, and a photo of the Evora site (below). From <http://www.esrl.noaa.gov/gmd/grad/surfrad/> and <http://www.imk-asf.kit.edu/>, respectively.

Cloud, the Computer Network Information Center, CAS (<http://www.gscloud.cn>), and the Next Generation Earth Science Discovery Tool (<http://reverb.echo.nasa.gov/reverb/>). The 1-km LST was produced by a generalized split window algorithm that uses radiances from the MODIS TIR window channels 31 (10.8–11.3 μm) and 32 (11.8–12.3 μm) (Wan & Dozier, 1996).

2.4. GOES Imager LST

GOES provides approximately half-hourly temporal resolution earth observation data, which constitute an important element of the land-surface processes. GOES Imager instruments consist of five spectral channels (one visible band and four infrared bands) (Sun & Pinker, 2003), which measure the radiant and reflected solar energy from sampled areas of the earth. GOES 10 Imager data are available from a web-based data archive and distribution system known as NOAA's

Comprehensive Large Array-data Stewardship System (<http://www.class.noaa.gov/nsaa/products>). The GOES 10 IR channels provide earth observation data at intervals of approximately every 30 min, at a nadir resolution of about 4 km. Infrared image data stored as GOES Variable Format (GVAR) counts packaged in 10-bit words in Network Common Data Format (netCDF) were used. The 10-bit (0–1023) GVAR count value was converted to brightness temperatures for channels 2 and 4, utilizing the calibration coefficients for the GOES 10 imager (Weinreb et al., 2011). GOES LST data at a 30-min temporal resolution were retrieved according to the algorithms described in Sun and Pinker (2003, 2005).

2.5. MSG SEVIRI LST

SEVIRI is the main instrument on board the MSG satellite series, which are the geostationary meteorological satellites developed by the

Table 1
The details of the satellite data used in this study.

Satellite data	Acquisition date	Acquisition time (UTC)	Number
Landsat ETM+ (path: 40, row: 35)	4 Aug 2002	18:09	2
MOD11A1 (h08v05)	20 Aug 2002	18:36	2
GOES 10 Imager	4 Aug 2002	18:36	2
Landsat TM (path: 203, row: 35)	20 Aug 2002	Starting at 00:00 and every 30 min thereafter	45
MOD11A1 (h17v05)	20 May 2010	10:59	1
MSG SEVIRI	18 May 2010	10:41	2
	18 May 2010	Starting at 00:00 and every 15 min thereafter	89

European Space Agency (ESA) and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) (Govaerts, Wagner, & Clerici, 2005; Schmets et al., 2002). The SEVIRI sensor scans the full earth disk every 15 min in 11 spectral bands, with a spatial resolution of 3 km. It consists of three visible and near-infrared channels, centered at 0.6, 0.8, and 1.6 μm , and eight infrared channels, centered at 3.9, 6.2, 7.3, 8.7, 9.7, 10.8, 12.0, and 13.4 μm . The LSA SAF operational LST product was used (Gottsche, Olesen, & Bork-Unkelbach, 2013), which is generated by the use of a generalized split window algorithm to operationally retrieve LST from the SEVIRI TIR channels at 10.8 and 12.0 μm . SEVIRI-derived LSTs are generated for four different geographical areas within the Meteosat disk: Europe, Northern Africa, Southern Africa, and South America. The general accuracy of the SEVIRI LST product is better than 2 K (Freitas, Trigo, Bioucas-Dias, & Gottsche, 2010; Trigo, Monteiro, Olesen, & Kabsch, 2008), and the SEVIRI LST product is freely available from the LSA SAF website (<http://landsaf.meteo.pt>).

3. Methodology

The premise of multi-source data fusion is that the preprocessed and corrected data from the different sensors are consistent and comparable (Pohl & Genderen, 1998). All the different sensor data need to be reprojected and resampled to a common spatial grid and coordinate system. If the different sensors scan the same scene within a short time interval, we treat the data as simultaneous, even though this assumption may cause some additional error (Liu & Weng, 2012). For example, the Terra/MODIS platform crosses the equator at about 10:30 am local solar time, which is roughly 30 min after Landsat ETM+ (Gao et al., 2006). For the purposes of this study, we consider the LST from the two sensors to be simultaneous observations (Table 1).

3.1. Fusion framework of arbitrary sensors

Neglecting preprocessing errors and differences in the retrieval methods, for a homogenous moderate spatial resolution pixel, the corresponding fine spatial resolution LST F can be expressed as (Huang et al., 2013; Wu et al., 2013):

$$F(i, t_1) = M(i, t_1) + \varepsilon_1, \tag{1}$$

where M denotes the moderate spatial resolution LST resampled to the fine-resolution grid, i denotes the i -th pixel, t_1 is the acquisition time, and ε_1 is the difference between the LST observed at the fine and moderate resolutions, which is mainly due to the spectral differences between the sensors and the radiometric differences caused by solar geometry. If the moderate spatial resolution LST $M(i, t_2)$ at another time t_2 is a known value among the inputs, then the fine spatial resolution LST at time t_2 can be predicted as:

$$F(i, t_2) = M(i, t_2) + \varepsilon_2. \tag{2}$$

Supposing that the land-cover type and sensor calibration do not change between dates t_1 and t_2 , then the residual would also not change from dates t_1 and t_2 . Furthermore, random noise is neglected. We have the residual $\varepsilon_2 = \varepsilon_1$, so the following equation can be obtained:

$$F(i, t_2) = F(i, t_1) + M(i, t_2) - M(i, t_1). \tag{3}$$

When a coarse spatial resolution sensor with a sufficiently high temporal resolution is also available, the fine spatial resolution LST at time t_3

can be predicted from the fine spatial resolution LST at time t_2 and the coarse spatial resolution LST $C(i, t_2)$ and $C(i, t_3)$:

$$F(i, t_3) = F(i, t_2) + C(i, t_3) - C(i, t_2). \tag{4}$$

Combining Eqs. (3) and (4), the high spatial resolution LST predicted at time t_3 can be expressed as:

$$F(i, t_3) = F(i, t_1) - M(i, t_1) + M(i, t_2) - C(i, t_2) + C(i, t_3). \tag{5}$$

By analogy, fused LST from a series of sensors with successively coarser resolutions can be obtained:

$$F(i, t_m) = F(i, t_1) - M(i, t_1) + M(i, t_2) - C(i, t_2) + C(i, t_3) - \dots - X(i, t_{m-1}) + X(i, t_m), \tag{6}$$

where t_m is the time for which the high spatial resolution LST is predicted, and $F, M, C,$ and X denote the LST values from the different sensors. The other sensors are not written but are expressed as ellipses. F has the highest spatial resolution among all the sensors, while X is the sensor with the highest temporal resolution. $t_1, t_2, t_3, \dots, t_{m-1}$ denote the acquisition times of the corresponding sensors. Note that, for each sensor, we may acquire more than one LST at their revisit cycle. For instance, we may have n input pairs of $F(i, t_1)$ and $M(i, t_1)$, where $n \in [1, +\infty]$.

However, at coarser spatial scales, land surfaces are rarely homogeneous and usually include several land-cover types. In this case, additional information from neighboring pixels can be introduced as a reference (Gao et al., 2006). We compute the LST value for the central pixel at t_m with a filter operator, which is similar to the weighting function used by Gao et al. (2006).

$$F(i_{w/2}, t_m) = \sum_{i=1}^w \sum_{t_1=1}^w \sum_{t_2=1}^w \sum_{t_3=1}^w \dots \sum_{t_{m-1}=1}^w W_{it} * \left(\begin{matrix} F(i, t_1) - M(i, t_1) + M(i, t_2) \\ -C(i, t_2) + C(i, t_3) - \dots - X(i, t_{m-1}) + X(i, t_m) \end{matrix} \right), \tag{7}$$

where w is the size of the moving window, and $i_{w/2}$ is the central pixel within the sliding window. It should be noted that only similar pixels within the sliding window are selected to calculate the LST value (Wu et al., 2013). The similar pixels are LST-similar pixels, as in Wu et al. (2013), and can be selected using:

$$\left| F(i, t_1) - F(i_{w/2}, t_1) \right| \leq \sigma \times 2/m \tag{8}$$

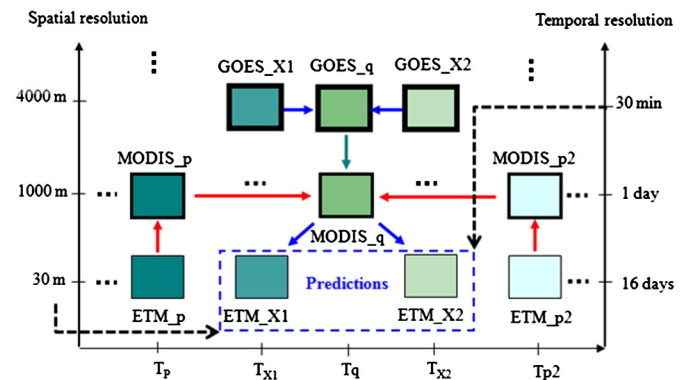


Fig. 2. Schematic of STITFM with Landsat ETM+, MODIS, and GOES data. The left vertical axis is the spatial resolution, and the right vertical axis is the temporal resolution, while the horizontal axis denotes the observation time. The predictions (within the dashed frame) are generated by considering two cloud-free ETM-MODIS pairs acquired at T_p and T_{p2} , the MODIS-GOES pairs acquired at T_q , and the GOES data at the predicted times T_{X1} and T_{X2} . The ellipses denote that STITFM allows for the input of more base LST pairs at different times. The predictions have a 30-m spatial resolution (ETM+) and a 30-min temporal resolution (GOES).

where $F(i, t_1)$ is the neighbor pixel LST, $F(i_{w/2}, t_1)$ is the central pixel LST, σ is the standard deviation for the LST data, and m is the estimated number of classes.

Although Eq. (7) allows for the input of more base sets of LST at different times ($t_1, t_2, t_3, \dots, t_{m-1} \in [1, +\infty]$), we generate a synthetic LST based on a single base set of LST ($t_1, t_2, t_3, \dots, t_{m-1} = 1$). Preliminary trials conducted using one base set of LST have indicated that the results are satisfactory, which were shown for two sensors in our previous study (Wu et al., 2013) and also by other researchers (Walker et al., 2011).

3.2. Spatio-temporal weighting function

The spatio-temporal weighting function W_{it} determines how much each similar pixel contributes to the central pixel and, therefore, has to be accurately known. It is determined by the following three measures:

- 1) The similarity difference between pixels in a sliding window and the central pixel at a given window location is:

$$S_i = |F(i, t_1) - F(i_{w/2}, t_1)|. \tag{9}$$

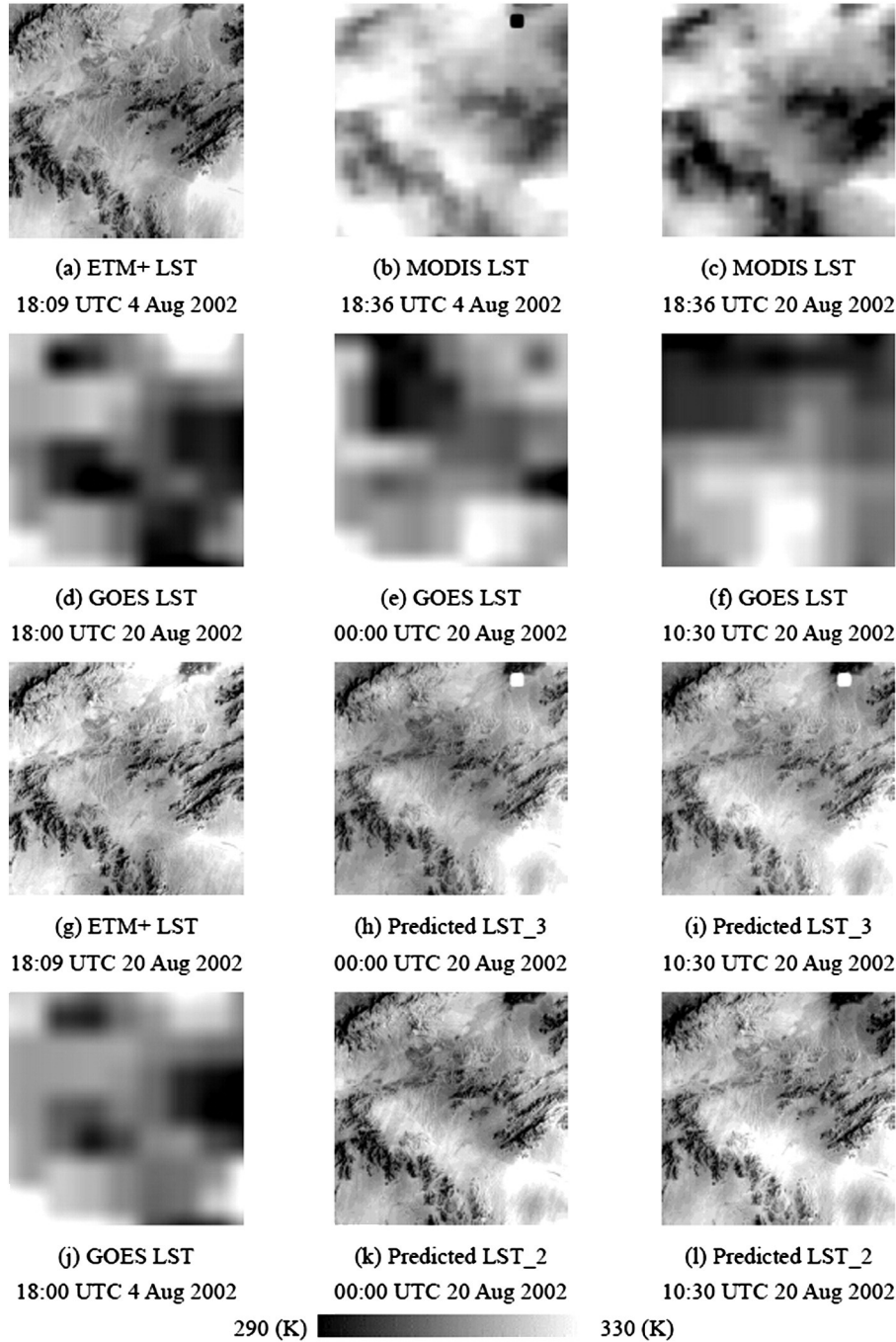


Fig. 3. Observed and predicted LSTs. The area is a 900 km² subset of remote sensed data. (a) observed ETM + LST on 4 August 2002; (b) and (c) observed MODIS LSTs on 4 August 2002 and 20 August 2002; (d)–(f) observed GOES LSTs at 18:00, 00:00, and 10:30 UTC on 20 August 2002; (g) observed ETM + LST on 20 August 2002; (h) and (i) predicted “Landsat-like” LSTs using three sensors (with MODIS LST); (j) observed GOES LST at 18:00 UTC on 4 August 2002; and (k) and (l) are the predicted “Landsat-like” LSTs using two sensors (without MODIS LST).

This is a simple measure to determine the similarity between the pixels in the sliding window and the central pixel at the fine spatial resolution observation. It is a nonnegative number, and was usually less than 2 K in our experiments. Even when similar pixels are chosen, their degree of similarity to the central pixel usually varies. A smaller value of S_i implies that the similar pixel has a higher degree of similarity to the central pixel and also contributes more additional information. The degree of similarity (SD_i) can be written as:

$$SD_i = \frac{\exp(-S_i)}{\sum \exp(-S_i)} \quad (10)$$

Therefore, the pixel with a higher SD_i should be assigned a higher weight in Eq. (7).

- 2) The scale difference between the different resolution data at their corresponding acquisition times is:

$$R_{it} = |F(i, t_1) - M(i, t_1) + M(i, t_2) - C(i, t_2) + C(i, t_3) - \dots - X(i, t_{m-1})|. \quad (11)$$

This is an approximate measure to determine the homogeneity of a lower-resolution pixel at the corresponding acquisition time. It is a nonnegative number, and was usually less than 4 K in our

experiments. The measure of homogeneity given by Eq. (11) is limited by differences in projection and geolocation errors. A smaller value of R_{it} means that the higher spatial resolution pixel has closer radiance features to the averaged surrounding pixels at the lower resolution; thus, the change at the higher resolution should be comparable to that of the averaged surrounding pixels.

- 3) The geometric distance between the similar pixel (x_i, y_i) and the central pixel $(x_{w/2}, y_{w/2})$ at a given window location is:

$$d_i = \sqrt{(x_i - x_{w/2})^2 + (y_i - y_{w/2})^2}. \quad (12)$$

This measures the spatial distance between the central predicted pixel and the surrounding similar pixels. A closer similar pixel should be assigned a higher weight according to spatial autocorrelation. To match the magnitude of the similarity difference and the scale difference, we convert the geometric distance to a relative distance through the function:

$$D_i = (A + d_i)/A \quad (13)$$

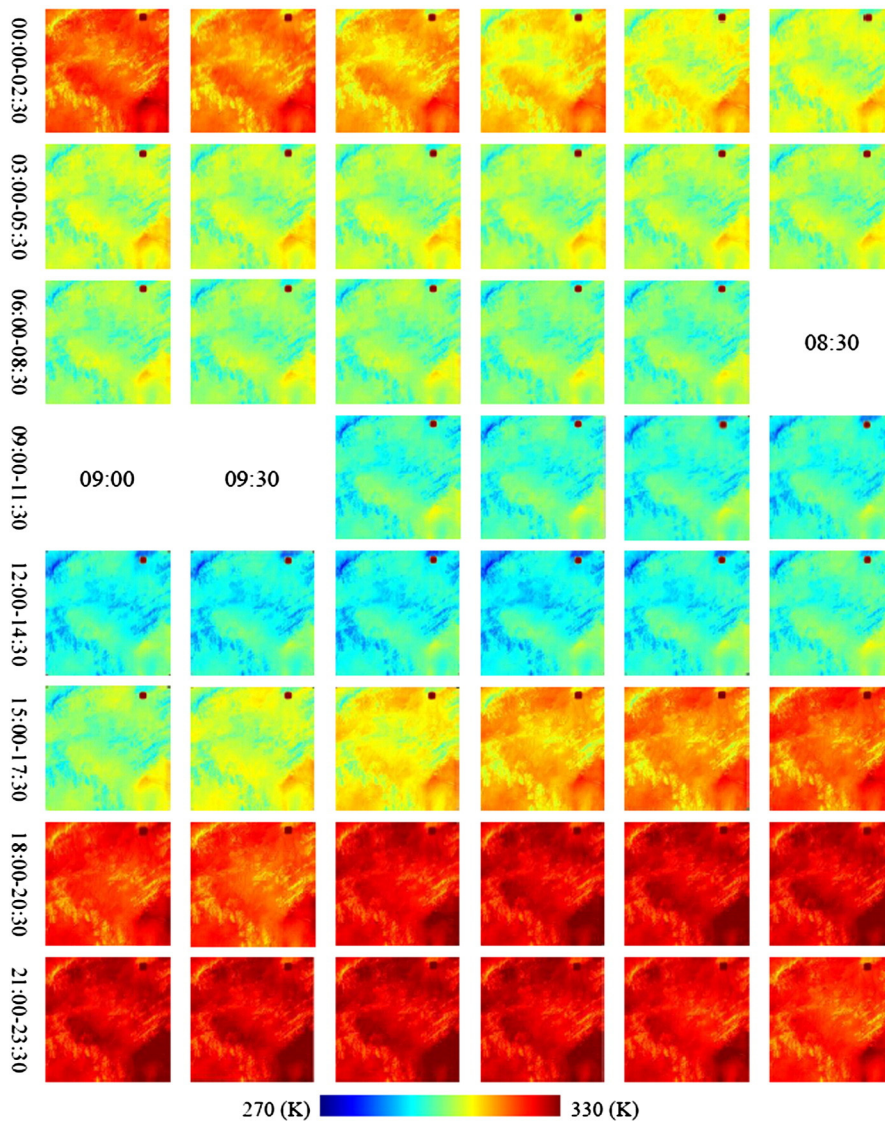


Fig. 4. The 45 predicted LSTs at a half-hourly temporal resolution at the Landsat spatial scale on 20 August 2002 (local time = UTC time – 7.7 h). The LSTs between 08:30 and 09:30 UTC could not be predicted due to missing GOES observations, and the LSTs with brown spaces are caused by the missing MODIS LST (the first column denotes the observation time of each line).

where A is a constant that is determined by the spatial resolution of the fine resolution and the window size (Gao et al., 2006).

The combined scale difference and the relative distance can be computed with:

$$E_{it} = \ln(R_{it} * 100 + 1) * D_i \tag{14}$$

To match the SD_i , we normalize E_{it} as:

$$V_{it} = E_{it} / \sum E_{it} \tag{15}$$

Ensuring that the sum of all weights is 1, W_{it} is normalized as follows:

$$W_{it} = (1 / (V_{it} * SD_i)) / \sum (1 / (V_{it} * SD_i)) \tag{16}$$

Eqs. (15) and (16) mean that a higher degree of similarity, a smaller scale difference, and a smaller distance from a similar pixel to the central pixel should produce a higher spatio-temporal weight. One special case is $R_{it} = 0$, where the spatio-temporal weight W_{it} is then set to the maximum value. The predicted value is $F(i_{w/2}, t_m) = X(i, t_m)$.

Fig. 2 presents a flowchart of STITFM with Landsat, MODIS, and GOES data.

4. Experimental results

We tested the performance of STITFM by fusing three LST products from the aforementioned satellite data, although the model is applicable to an arbitrary number of products. Two scenarios were investigated: the first scenario fuses Landsat, MODIS, and GOES data for North America, while the second scenario uses data from Landsat, MODIS, and SEVIRI for Europe. The predicted LST was obtained at the Landsat spatial resolution. Satellite data with a similar spatial resolution (such as ASTER, with a 90-m spatial resolution) would be ideal for evaluating the predictions; unfortunately, such data were not available (with a similar overpass time). The only option was the comparison of the satellite-based LSTs with in situ measurements, even if this implies a large scale factor (Yu et al., 2009). Therefore, corresponding ground measurements from SURFRAD and KIT were used, in line with the approach used in a number of previous studies (Gottsche et al., 2013; Inamdar et al., 2008; Wang & Liang, 2009; Xu et al., 2013; Yu et al., 2012, 2009).

Two experiments (EXP1 and EXP2) were conducted. For EXP1, Landsat ETM+, MODIS, and GOES LST were used as the input data (Table 1). We evaluated the fusion results by comparing them with in situ LST from SURFRAD. For EXP2, the in situ LST from KIT was used to evaluate the predictions obtained from the Landsat TM, MODIS, and MSG SEVIRI LSTs (Table 1). For a quantitative evaluation of the agreement between the in situ LSTs and the LST predictions, we used the coefficient of determination (R^2), the bias, and the root-mean-square error (RMSE).

4.1. Fusion of ETM+, MODIS, and GOES LST

To investigate the practical feasibility of STITFM, the presented experiments focused on deriving half-hourly LST at a Landsat spatial resolution over the southwestern U.S. The area was a 900 km² subset, encompassing an elevation range of 725–1813 m (McVicar & Körner, 2013). The choice of this area was based on the following reasons: (1) the availability of high-quality satellite data; (2) the accessibility to a surface LST validation site (the DRA site); (3) the opportunity to address problems created by low emissivities over the semi-arid and arid regions; and (4) the feasibility and credibility of the region for LST studies have already been evaluated (Inamdar et al., 2008).

The ETM+ LST observed at 18:09 UTC on 4 August 2002, the MOD11A1 LST observed at 18:36 UTC on 4 August 2002 and 20 August

2002, and the GOES LST observed at 18:00 UTC 20 August 2002 were used as the base data, while the other GOES LSTs were used as the input data for generating the LST predictions. It should be noted that the Landsat TIR channel had been resampled to 30 m when we downloaded it. Therefore, the predictions are synthetic LSTs with 30-m spatial resolution at a 30-min temporal interval. Fig. 3 shows the observed ETM+ LST, the MODIS LST, parts of the GOES LST, and the predicted results. The base data, Fig. 3(a)–(d), were observed ETM+ LST on 4 August 2002, the observed MODIS LST on 4 August 2002, the observed MODIS LST on 20 August 2002, and the observed GOES LST at 18:00 UTC on 20 August 2002, respectively. We only show two of the 44 GOES LSTs that were available for predicting the LST (Fig. 3(e) and (f)). The corresponding predicted “Landsat-like” LSTs at 00:00 and 10:30 UTC on 20 August 2002 are shown in Fig. 3(h) and (i). The predicted LSTs have a white dot in the top-right corner, which was caused by the missing MODIS LST shown in Fig. 3(b).

Note that we can also obtain “Landsat-like” predictions using only two sensors (i.e. without MODIS LST as an input), but the accuracy may be lower. LST predictions obtained from only GOES and ETM+ LST as the base data (Fig. 3(a)–(d) and (j)) are shown in Fig. 3(k) and

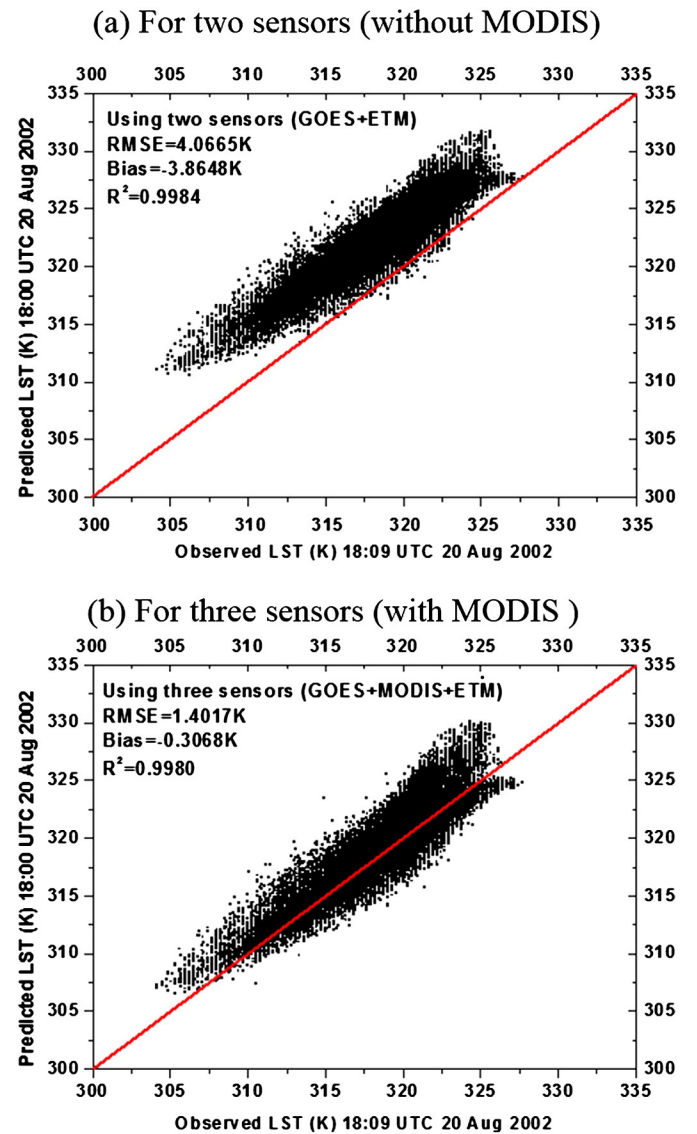


Fig. 5. Scatter plots of observed LST against LSTs predicted with STITFM for different numbers of sensors: (a) for two sensors (without MODIS); and (b) for three sensors (with MODIS). The solid line is the 1:1 line.

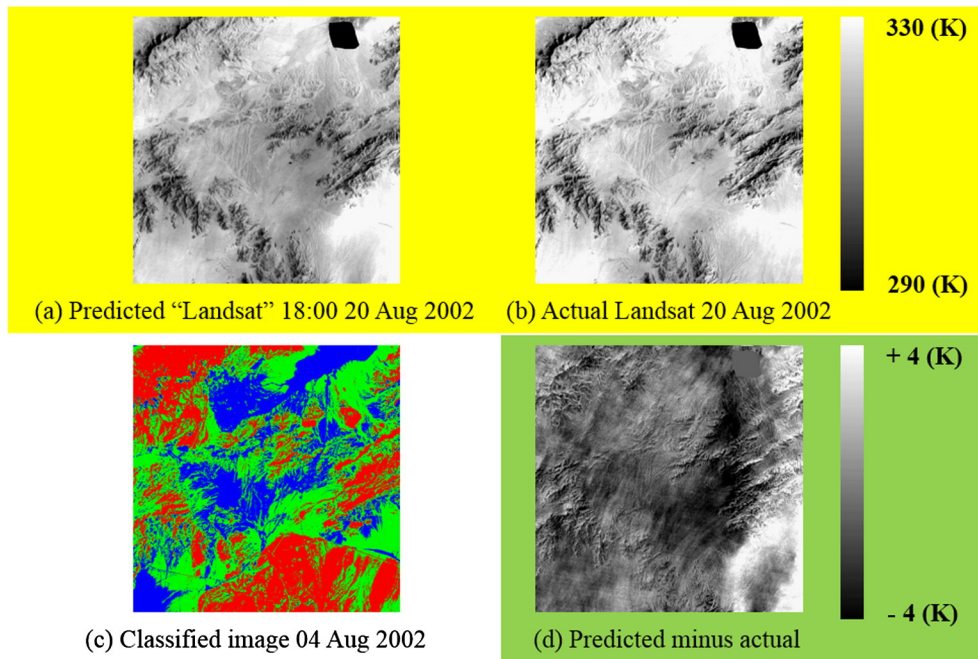


Fig. 6. Predicted (a), actual (b), and their difference (d) LSTs acquired at about 18:00 UTC on 20 Aug 2002. Panel (c) is the classified image of Landsat acquired on 04 Aug 2002. Open shrubland, desert rock, and mixed forest are marked in red, green, and blue, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(1). The observed ETM + LST (Fig. 3(g)) observed at 18:09 UTC on 20 August 2002 can be used to evaluate the predicted results. Using GOES LSTs at other times on 20 August 2002, we can obtain predicted Landsat-like LST at all available GOES acquisition times. For convenience, Fig. 4 shows the 45 predicted LSTs obtained from the three sensors during 20 August 2002. Fig. 4 also illustrates the temperature variation over the course of the day.

The ETM + LST at 18:09 UTC on 20 August 2002 can be observed; therefore, we can compare the observed ETM + LST with the predictions at the similar time. Fig. 5 shows scatter plots between the observed ETM + LST at 18:09 UTC and predicted ETM + LST at 18:00 UTC on 20 August 2002. Fig. 5(a) shows the observed LST vs. the LST predicted with STITFM when using two sensors (i.e. without MODIS), while Fig. 5(b) shows the observed vs. the predicted LST for three sensors (with MODIS). The data in Fig. 5(b) fall more closely onto the 1-1 line than those in Fig. 5(a), which indicates that STITFM with three sensors improves the prediction. Compared to the prediction obtained at 18:00 UTC from two sensors, the prediction at the same time using three sensors improves the RMSE from 4.06 K to 1.40 K and the bias from -3.86 K to -0.31 K. This is primarily because the moderate-resolution MODIS LST plays the important role of “scale transition” by narrowing the spatial resolution difference between the ETM + and GOES LST.

Moreover, to show the relationship between land cover and the proposed method, the errors for different land-cover types were calculated. Firstly, a mask (for the missing pixels) was built and applied for both the predicted “Landsat” LST from three sensors and the actual Landsat LST (Fig. 6(a) and (b)). The unsupervised IsoData classification method was then employed for the actual Landsat LST acquired on 4 Aug 2002, and the number of classes was set as three in the experiments (Fig. 6(c)). The land-cover classes consisted of a mixed type of desert

rock and open shrubland (in red), desert rock (in green), and open shrubland (in blue). Finally, the RMSE and bias of each land-cover type were calculated and are listed in Table 2, according to the classified image of Fig. 6(c) and the difference image of Fig. 6(d).

From Table 2, we can see that the errors for different land-cover types are different. The RMSE and bias of the mixed type (in red) are similar to the whole scene. The RMSE and bias of desert rock (in green) are lower than for the whole scene, while the RMSE and bias of open shrubland (in blue) are the highest. In other words, the capability of the spatio-temporal fusion method may be affected by different land-cover types. These findings have also been confirmed by previous studies (Hilker et al., 2009). Furthermore, the errors also differ for the same land-cover type, which indicates that the proposed method may also be affected by topography, such as elevation, slope, and aspect. Overall, we found that higher accuracies were obtained in the homogeneous areas of each category.

Since the Landsat satellite only observes the surface at its overpass time, in situ LSTs were used to validate the predicted LST at the other times on 20 August 2002. Ground-truth data from the DRA site in Nevada, U.S., were used to verify the effectiveness of STITFM. Fig. 7 compares the in situ LSTs from the DRA site with the LST predictions. From the figure, it is clear that there is an overall agreement between the LST predictions and the ground-truth LST. However, the predicted LSTs obtained from three sensors are closer to the in situ LSTs than the LSTs obtained from two sensors. The prediction error for three sensors (RMSE = 2.5 K) is lower than that for two sensors (RMSE = 3.7 K). These results were obtained by the use of Landsat LST and GOES LST acquired on different days (the Landsat LST was observed on 4 August 2002 and the GOES LST was observed on 20 August 2002). In our previous study (Wu et al., 2013), we obtained predictions by the use of Landsat LST and GOES LST observed on the same day (20 August 2002). The

Table 2
The errors for different land-cover types.

Class	Mixed type	Desert rock	Open shrubland	No classification (whole scene)
RMSE (K)	1.48	1.10	1.72	1.40
Bias (K)	-0.41	-0.20	-0.91	-0.31

prediction error for two sensors (observed on the same day) was 1.87 K (Wu et al., 2013), which is better than the result from different days (for both three sensors and two sensors: RMSE = 2.5 K for three sensors and RMSE = 3.7 K for two sensors). This is assumed to be due to the greater consistency of the surface parameters and, consequently, better predictions can be obtained when all the data are from the same day.

4.2. Fusion of TM, MODIS, and SEVIRI LST

In this experiment (EXP2), STITFM was tested on LST retrieved over Evora, Portugal, where one of KIT's permanent validation stations is located. Since SEVIRI data have a 15 min temporal resolution, we can theoretically obtain 96 SEVIRI LST data per day. However, due to clouds or missing data, we actually obtained 89 SEVIRI LST data for EXP2.

Fig. 8(a)–(h) shows the observed TM LST on 20 May 2010, the observed MODIS LST on 20 May 2010, the observed MODIS LST on 18 May 2010, the observed SEVIRI LST at 11:00 UTC on 18 May 2010, the observed SEVIRI LST at 00:15 UTC on 18 May 2010, the predicted “Landsat-like” LST at 00:15 UTC on 18 May 2010, and the predicted “Landsat-like” LST at 00:45 UTC on 18 May 2010, respectively. Fig. 8(a), (b), (c), and (d) shows the input base data, while Fig. 8(e) and (f) shows two of the 89 available SEVIRI LSTs. The “Landsat-like” LST predictions corresponding to the two SEVIRI acquisition times are shown in Fig. 8(g) and (h). Similarly, the other 87 LST predictions at SEVIRI acquisition times for 18 May 2010 can be obtained, but are not reproduced here. Predictions with the same resolution were obtained by fusing only SEVIRI and TM LSTs. The data in Fig. 8(a) and (i) served as input base data, while Fig. 8(e) and (f) shows the input SEVIRI LST at the predicted time. The corresponding predictions are shown in Fig. 8(j) and (k). Although there are no obvious visual differences between Fig. 8(g) and (h) obtained for three sensors and Fig. 8(j) and (k) obtained for two sensors, some detailed regions are shown in Fig. 9 for the convenience of visual judgment.

For a better analysis, Fig. 9(a),(b),(d), and (e) shows the LST predictions from Fig. 8(j),(k),(g), and (h) in color, while Fig. 9(g) shows the observed TM LST from Fig. 8(a). When comparing the subsets of the observed TM LST with the predicted Landsat-like LST in the right column, it is clear that the prediction from three sensors is considerably closer to the observations than the prediction obtained from two sensors. Since

there are no TM LSTs at 11:15 UTC on 18 May 2010, Landsat TM LST acquired at a similar time (10:59 UTC) on 20 May 2010 is shown in Fig. 9(g), which introduces some additional uncertainty. Fig. 10 shows in situ LSTs from the Evora site and LST predictions using two and three sensors as the input. It is clear that the predicted LSTs obtained by fusing the data from three sensors are closer to the in situ LSTs than those produced by fusing the data from two sensors. The prediction error for three sensors (RMSE = 2.2 K) is smaller than that for two sensors (RMSE = 3.3 K). This means that the predictions obtained from three sensors are closer to the ground truth, which indicates that the base concept of STITFM, i.e. to use the data from more than two sensors as input, is meaningful and leads to improved results.

5. Discussion

The available satellite LST products have either a coarse spatial resolution or a low temporal resolution, which constrains their potential applications. To better utilize the complementary information in different data, it is necessary to develop techniques to fuse this useful information. Considering the limitations of the existing fusion methods, this paper proposes a spatio-temporal integrated temperature fusion model (STITFM) that uses multi-scale polar-orbiting and geostationary satellite observations to predict higher temporal and spatial resolution LST. The multi-scale polar-orbiting satellites such as Landsat and Terra can provide fine-resolution spatial details, while data from geostationary platforms (GOES or SEVIRI) can be obtained with a much better temporal resolution (images every 15 or 30 min). This was accomplished by blending LSTs from three sensors, although STITFM is able to fuse data from an arbitrary number of sensors.

This study has examined in detail the theoretical basis of the proposed method and included two series of experiments that predicted LSTs at the Desert Rock (DRA) site in Nevada, U.S., and the Evora site in Evora, Portugal. In general, the resulting high spatial and temporal resolution LSTs obtained from three sensors were better than those obtained from two sensors. The predictions were evaluated against both observed high-resolution LST data and in situ LSTs. The experiments indicated that LST can be predicted using data from three sensors with an accuracy (RMSE) of better than 2.5 K, but the accuracy is affected by different land-cover types. Other sources of errors may include the data preprocessing, the retrieval methods of the different sensors.

The primary reason for the superior performance of the three-sensor STITFM is that the moderate resolution LST improves the distribution of the spatial information during the “scale transition” by narrowing the spatial resolution difference between the fine- and coarse-resolution LST. Many details that cannot be shown in geostationary satellite pixels (GOES or SEVIRI) can be captured by MODIS. Unlike the previous fusion models for Landsat and MODIS LST (Huang et al., 2013; Weng et al., 2014), the model proposed in this paper provides a potential solution that utilizes a geostationary satellite to obtain “Landsat-like” LSTs outside the Landsat overpass time, which may be useful for dynamic surface heat island analysis (Yuan & Bauer, 2007). On the other hand, in contrast with our original work, in which predictions were obtained by fusing only Landsat and geostationary satellite data (Wu et al., 2013), STITFM is suitable for the data from different days, and can also obtain more accurate predictions from more sensors in a unified framework.

However, it should be noted that there are several potential limitations to the proposed method. The main limitation of the proposed method is its relatively strict conditions, such as the quality of the input data, orbital parameters, overpass times, viewing geometry, imaging registration, and the availability of biophysical parameters. Considering that the differences in the overpass time and the viewing geometry between the Landsat and MODIS sensors were small for the study area, this study did not attempt to correct for the difference in diurnal temperature change. However, this correction or normalization should be conducted to extend the model further, especially for LST

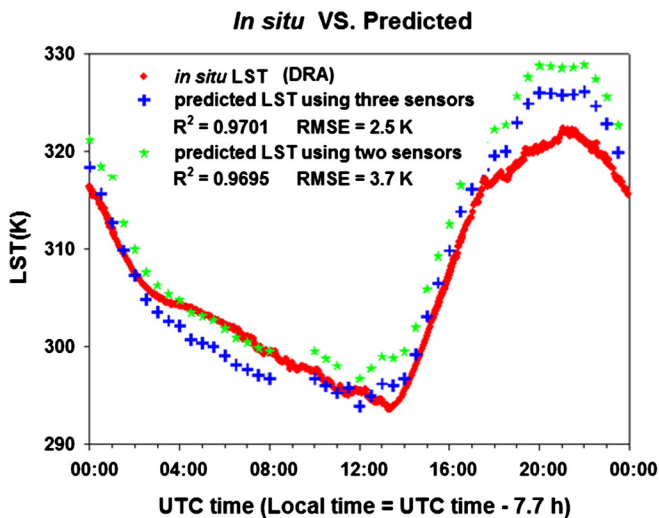


Fig. 7. Diurnal cycle of LSTs predicted using three sensors (blue pluses) and two sensors (green stars), compared with that of the in situ LSTs (red dotted line) from the DRA site on 20 August 2002. The data gap for GOES between 08:30 and 09:30 UTC is due to missing observations, which means that there are also no predictions.

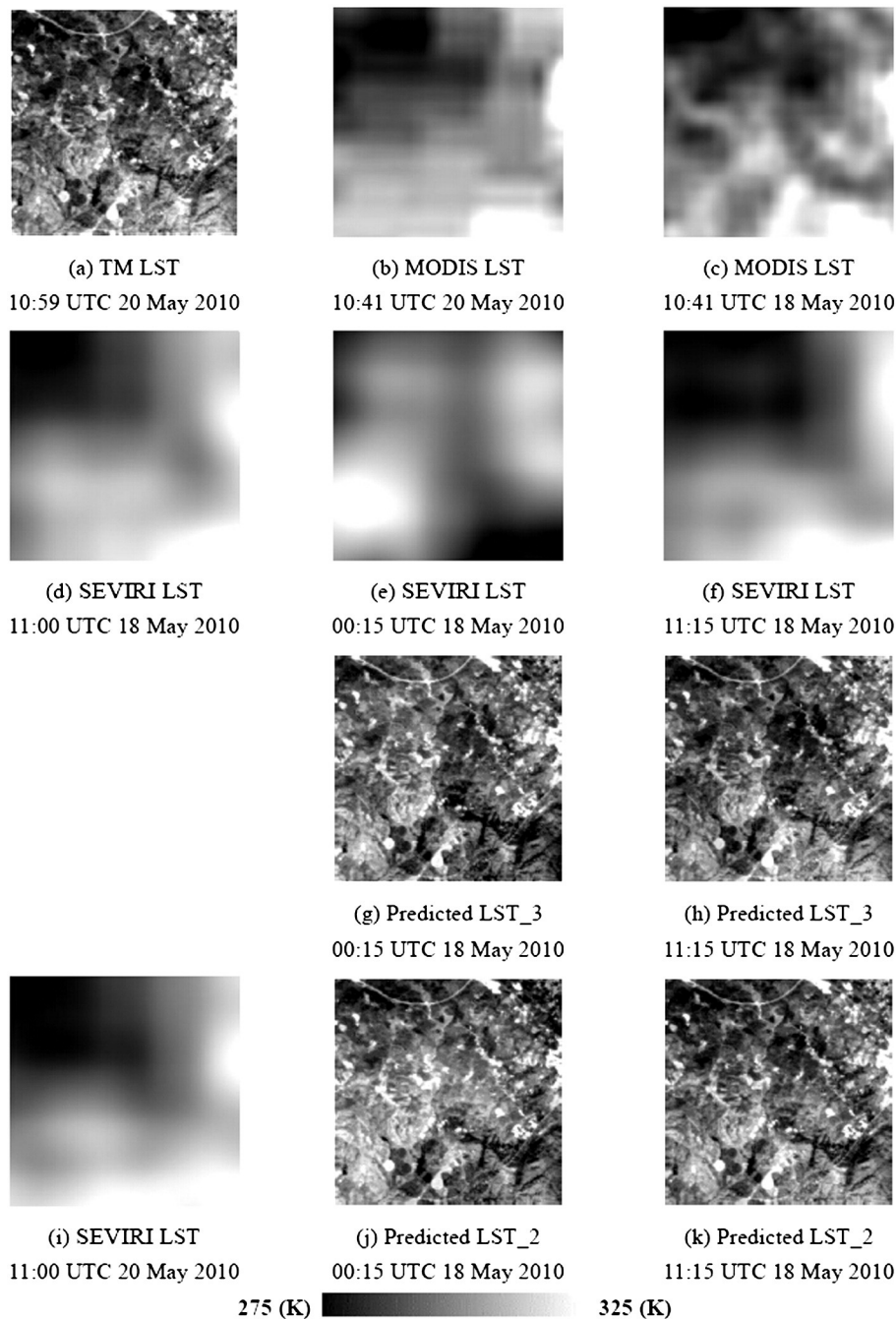


Fig. 8. Observed and predicted LSTs. The area is a 900 km² subset of remote sensed data. (a) observed TM LST on 20 May 2010; (b) and (c) observed MODIS LSTs on 20 May 2010 and 18 May 2010; (d)–(f) observed GOES LSTs at 11:00, 00:15, and 11:15 UTC on 18 May 2010; (g) and (h) predicted “Landsat-like” LSTs using three sensors (with MODIS LST); (i) observed GOES LST at 11:00 UTC on 20 May 2010; and (j) and (k) are the predicted “Landsat-like” LSTs using two sensors (without MODIS LST).

(Duan, Li, Tang, Wu, & Tang, 2014). Secondly, as in the previous studies (Gao et al., 2006; Huang et al., 2013; Weng et al., 2014; Zhu et al., 2010), STITFM does not have the ability to predict changes that are not reflected in all the sensor pixels. An improved fusion model that could combine the decomposition of mixed pixels and sub-pixel mapping would be useful for predicting changes. In addition, a number of parameters are sensitive to the different scenes, e.g. the window size and the number of land-cover classes, and must be carefully set.

For the design of the experiments, a number of issues need to be explained, and should be investigated further in the future.

Only one in situ site was selected for each series of experiments (EXP1: DRA; EXP2: Evora). The main reason for this is that the distance between the different sites (provided by SURFRAD or KIT) is significant,

and cannot be covered by one Landsat scene. Moreover, a large number of pixels in the MODIS LST (MOD11A1) product were masked out due to cloud, heavy aerosols, etc., especially in EXP1 (see Fig. 3(b)). Image mosaicing and gap-filling techniques using multi-temporal data could help to improve this situation.

The two experiments were conducted only for a single day (24 UTC hours). In theory, more consistent predictions (more than 24 UTC hours) could have been obtained. However, satellite-derived time series of LST (such as the daily MODIS LST) vary with satellite viewing angles and sun geometry, which cause serious bias between predictions and in situ LST (Vinnikov et al., 2012). Production of such a time series requires addressing angular anisotropy when using asynchronous satellite observations (Vinnikov et al., 2012). Although angular variations in LST

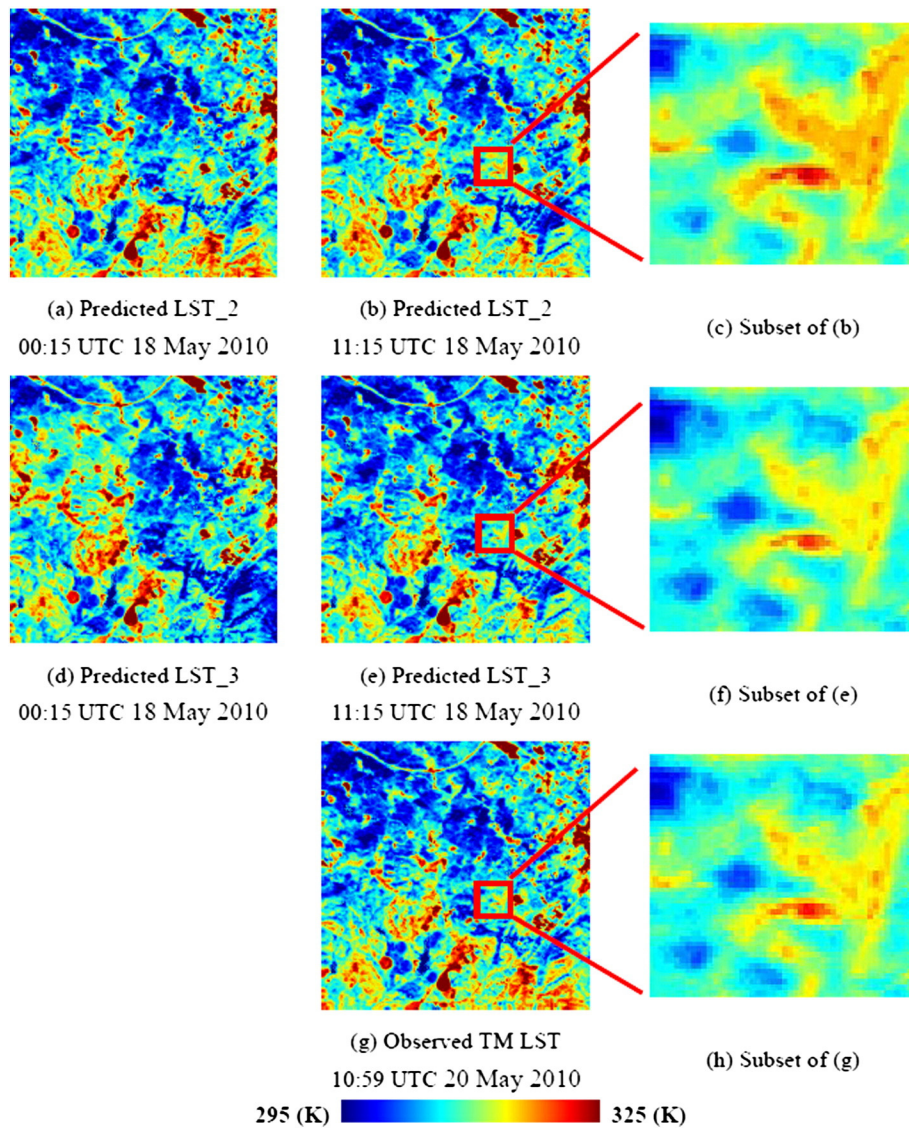


Fig. 9. Comparison between observed LSTs and LSTs predicted using a different number of sensors. Panels (a) and (b) show the predictions corresponding to Fig. 8(j) and (k) obtained from two sensors; panels (d) and (e) are the predictions retrieved from three sensors and correspond to Fig. 8(g) and (h); sub-figure (g) shows the observed TM LST as in Fig. 8(a); and panels (c), (f), and (h) show subsets of (b), (e), and (g) in more detail.

have been investigated by several researchers (Pinheiro, Privette, & Guillevic, 2006; Vinnikov et al., 2012), there is currently no practical way to perform angular normalization of satellite-derived LSTs (Li et al., 2013). To eliminate or at least limit the angular effect, only MODIS LST acquired under similar satellite viewing angles (less than 10°) was selected.

The final issue is the evaluation of the predictions with in situ LST. The most important factor for choosing any LST validation site is the homogeneity of its surface temperature. In such a case, evaluation with an in situ site can be acceptable; however, how to evaluate the other pixels of the predictions is still a real challenge. A possible solution is to achieve spatial modeling for the in situ LSTs while taking into account the influence of topography, land-cover types, and so on (McVicar et al., 2007).

6. Conclusion

The STITFM algorithm was successfully used in this study to obtain LST at a Landsat spatial resolution and 30 min or even 15 min temporal resolution. The predicted LSTs from three sensors (Landsat, MODIS, and GOES/SEVIRI) were better than those from two sensors (Landsat and

GOES/SEVIRI), which were evaluated using the observed high-resolution LST data and in situ LSTs. Land cover was shown to be associated with algorithm performance and was intricately linked with the site domain characteristics. Although STITFM was tested with LST in this study, it advances the capability for producing other remotely sensed data products with both high spatial resolution and frequent coverage from multi-scale polar-orbiting and geostationary satellite observations. Such a capability will be beneficial for monitoring diurnal land-surface and ecological dynamics at the spatial scales most relevant to human activities, such as urban heat islands (Zaksek & Ostir, 2012) or water turbidity measurement (Vanhellemont et al., 2014).

Acknowledgments

The authors would like to thank NASA, USGS, NOAA, LSA SAF and KIT for the satellite data and the in situ data. This work was supported by the Major State Basic Research Development Program (973 Program) under Grant 2011CB707103, the National High Technology Research and Development Program (863 Program) under Grant 2013AA12A301, the National Natural Science Foundation of China

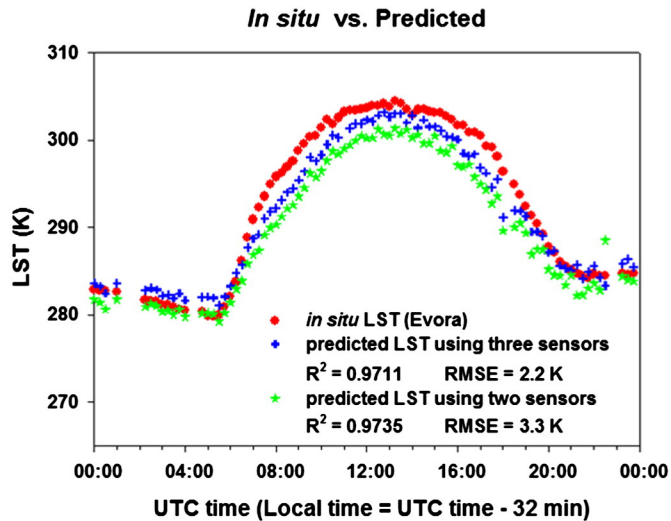


Fig. 10. Diurnal cycle of LSTs predicted using three sensors (blue pluses) and two sensors (green stars) and in situ LSTs (red solid circles) at the Evora site on 18 May 2010. The data gaps correspond to missing SEVIRI LST. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

under Grant 41422108 the Program for Changjiang Scholars and Innovative Research Team in University under Grant IRT1278, and the Startup Project of Doctor in Anhui University. Many thanks to the anonymous reviewers for their constructive comments and suggestions. The main work of the first author was completed in Wuhan University.

References

- Anderson, M. C., Allen, R. G., Morse, A., & Kustas, W. P. (2012). Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sensing of Environment*, 122, 50–65.
- Duan, S. B., Li, Z. L., Tang, B. H., Wu, H., & Tang, R. (2014). Generation of a time-consistent land surface temperature product from MODIS data. *Remote Sensing of Environment*, 140, 339–349.
- Emelyanova, I. V., McVicar, T. R., Van Niel, T. G., Li, L. T., & van Dijk, A. I. J. M. (2013). Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sensing of Environment*, 133, 193–209.
- Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., & Gottsche, F. -M. (2010). Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1), 523–534.
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44(8), 2207–2218.
- Gottsche, F. -M., Olesen, F. -S., & Bork-Unkelbach, A. (2013). Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. *International Journal of Remote Sensing*, 34(9–10), 3069–3083.
- Govaerts, Y., Wagner, S., & Clerici, M. (2005). *SEVIRI native format pre-processing toolbox user's guide version 2.2*. EUMETSAT, report no. EUM/OPS-MSG/TEN/03/0011.
- Gucluer, D., Bayram, B., & Maktav, D. (2010). *Land cover and coast line change detection by using object oriented image processing in Alacati*. Turkey: 158. Dolunay: IOS Press.
- Hilker, T., Wulder, M., Coops, N., Seitz, N., White, J., Gao, F., et al. (2009). Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing of Environment*, 113(9), 1988–1999.
- Huang, B., Wang, J., Song, H., Fu, D., & Wong, K. (2013). Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring. *IEEE Geoscience and Remote Sensing Letters*, 10(5), 1011–1015.
- Inamdar, A. K., French, A., Hook, S., Vaughan, G., & Lockett, W. (2008). Land surface temperature retrieval at high spatial and temporal resolutions over the southwestern United States. *Journal of Geophysical Research*, 113(D07). <http://dx.doi.org/10.1029/2007JD009048>.
- Inamdar, A. K., & French, A. (2009). Disaggregation of GOES land surface temperatures using surface emissivity. *Geophysical Research Letters*, 36(2), L02408. <http://dx.doi.org/10.1029/2008GL036544>.
- Jimenez-Munoz, J. C., & Sobrino, J. A. (2003). A generalized single-channel method for retrieving land surface temperature from remote sensing data. *Journal of Geophysical Research*, 108(D22), 4688. <http://dx.doi.org/10.1029/2003JD003480>.
- Kabsch, E., Olesen, F., & Prata, F. (2008). Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements. *International Journal of Remote Sensing*, 29(17–18), 5329–5345.
- Kalma, J. D., McVicar, T. R., & McCabe, M. F. (2008). Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surveys in Geophysics*, 29(4–5), 421–469.
- Kim, J., & Hogue, T. S. (2012). Evaluation and sensitivity testing of a coupled Landsat-MODIS downscaling method for land surface temperature and vegetation indices in semi-arid regions. *Journal of Applied Remote Sensing*, 6(1) (063569-063561-063569-063517).
- Kumak, K. S., Bhaskar, P. U., & Padmakumari, K. (2012). Estimation of land surface temperature to study urban heat island effect using LANDSAT ETM+ image. *International Journal of Engineering Science*, 4(2), 771–778.
- 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 Sensing of Environment*, 85(4), 429–440.
- Li, Z. -L., Tang, B. -H., Wu, H., Ren, H., Yan, G., Wan, Z., et al. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, 131, 14–37.
- Liu, H., & Weng, Q. (2012). Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007. *Remote Sensing of Environment*, 117, 57–71.
- Masek, J., Lindsay, F., & Goward, S. (2000). Dynamics of urban growth in the Washington DC metropolitan area, 1973–1996, from Landsat observations. *International Journal of Remote Sensing*, 21(18), 3473–3486.
- McVicar, T. R., & Körner, C. (2013). On the use of elevation, altitude and height in the ecological and climatology literature. *Oecologia*, 171(2), 335–337. <http://dx.doi.org/10.1007/s00442-012-2416-7>.
- McVicar, T. R., Van Niel, T. G., Li, L. T., Hutchinson, M. F., Mu, X. M., & Liu, Z. H. (2007). Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. *Journal of Hydrology*, 338(3–4), 196–220. <http://dx.doi.org/10.1016/j.jhydrol.2007.02.018>.
- Nichol, J. (2009). An emissivity modulation method for spatial enhancement of thermal satellite images in urban heat island analysis. *Photogrammetric Engineering and Remote Sensing*, 75(5), 547–556.
- Pinheiro, A. C., Privette, J. L., & Guillevic, P. (2006). Modeling the observed angular anisotropy of land surface temperature in a savanna. *IEEE Transactions on Geoscience and Remote Sensing*, 44(4), 1036–1047.
- Pohl, C., & Genderen, J. V. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5), 823–854.
- Schmetz, J., Pili, P., Tjemkes, S. A., Just, D., Kerkmann, J., Rota, S., et al. (2002). An introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, 83, 977–992.
- Sellers, P., Dickinson, R., Randall, D., Betts, A., Hall, F., Berry, J., et al. (1997). Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. *Science*, 275(5299), 502–509. <http://dx.doi.org/10.1126/science.275.5299.502>.
- Sun, D., & Pinker, R. T. (2003). Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8). *Journal of Geophysical Research*, 108(D11), 4326. <http://dx.doi.org/10.1029/2002JD002422>.
- Sun, D., & Pinker, R. T. (2005). Implementation of GOES-based land surface temperature diurnal cycle to AVHRR. *International Journal of Remote Sensing*, 26(18), 3975–3984.
- Sun, D., Pinker, R. T., & Basara, J. B. (2004). Land surface temperature estimation from the next generation of Geostationary Operational Environmental Satellites: GOES M-Q. *Journal of Applied Meteorology*, 43, 363–372.
- Sun, D., Pinker, R. T., & Kafatos, M. (2006). Diurnal temperature range over the United States: A satellite view. *Geophysical Research Letters*, 33(5), L05705. <http://dx.doi.org/10.1029/2005GL024780>.
- Tierney, J. E., Russell, J. M., Huang, Y., Damste, J. S. S., Hopmans, E. C., & Cohen, A. S. (2008). Northern hemisphere controls on tropical southeast African climate during the past 60,000 years. *Science*, 322(5899), 252–255. <http://dx.doi.org/10.1126/science.1160485>.
- Trigo, I. F., Monteiro, I. T., Olesen, F., & Kabsch, E. (2008). An assessment of remotely sensed land surface temperature. *Journal of Geophysical Research*, 113(D17). <http://dx.doi.org/10.1029/2008JD010035>.
- Vanhellemont, Q., Neukermans, G., & Ruddick, K. (2014). Synergy between polar-orbiting and geostationary sensors: Remote sensing of the ocean at high spatial and high temporal resolution. *Remote Sensing of Environment*, 146, 49–62.
- Vinnikov, K. Y., Yu, Y., Goldberg, M. D., Tarpley, D., Romanov, P., Laszlo, I., et al. (2012). Angular anisotropy of satellite observations of land surface temperature. *Geophysical Research Letters*, 39(23), L23802. <http://dx.doi.org/10.1029/2012GL054059>.
- Walker, J., de Beurs, K., Wynne, R., & Gao, F. (2011). Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sensing of Environment*, 117, 381–393.
- Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on Geoscience and Remote Sensing*, 34(4), 892–905.
- Wan, Z., Zhang, Y., Zhang, Q., & Li, Z. -L. (2004). Quality assessment and validation of the MODIS global land surface temperature. *International Journal of Remote Sensing*, 25(1), 261–274.
- Wang, K., & Liang, S. (2009). Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sensing of Environment*, 113(7), 1556–1565.
- Weinreb, P., Johnson, J., & Han, D. (2011). *Conversion of GVAR infrared data to scene radiance or temperature*. NOAA technical memorandum, NOAA NESDIS office of satellite operations. <http://www.oss.noaa.gov/goes/goes-calibration/gvar-conversion.htm>
- Weng, Q., Fu, P., & Gao, F. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment*, 145, 55–67.
- Wu, P., Shen, H., Ai, T., & Liu, Y. (2013). Land-surface temperature retrieval at high spatial and temporal resolutions based on multi-sensor fusion. *International Journal of Digital Earth*, 6(51), 113–133.

- Xu, H., Yu, Y., Tarpley, D., Gottsche, F., -M., & Olesen, F. -S. (2013). Evaluation of GOES-R land surface temperature algorithm using SEVIRI satellite retrievals with in situ measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 52(7), 3812–3822.
- Yang, G., Pu, R., Huang, W., Wang, J., & Zhao, C. (2010). A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 48(4), 2170–2178.
- Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., et al. (2012). Validation of GOES-R satellite land surface temperature algorithm using SURFRAD. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3), 704–713.
- Yu, Y., Tarpley, D., Privette, J. L., Goldberg, M. D., Rama Varma Raja, M., Vinnikov, K. Y., et al. (2009). Developing algorithm for operational GOES-R land surface temperature product. *IEEE Transactions on Geoscience and Remote Sensing*, 47(3), 936–951.
- Yuan, F., & Bauer, M. E. (2007). Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. *Remote Sensing of Environment*, 106(3), 375–386.
- Zaksek, K., & Ostir, K. (2012). Downscaling land surface temperature for urban heat island diurnal cycle analysis. *Remote Sensing of Environment*, 117, 114–124.
- Zhan, W., Chen, Y., Wang, J., Zhou, J., Quan, J., Liu, W., et al. (2012). Downscaling land surface temperatures with multi-spectral and multi-resolution images. *International Journal of Applied Earth Observation and Geoinformation*, 18, 23–36.
- Zhan, W., Chen, Y., Zhou, J., Wang, J., Liu, W., Voogt, J., et al. (2013). Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats. *Remote Sensing of Environment*, 131, 119–139.
- Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114(11), 2610–2623.