



# Consistent land surface temperature data generation from irregularly spaced Landsat imagery

Peng Fu <sup>b</sup>, Qihao Weng <sup>a,b,\*</sup>

<sup>a</sup> School of Geography, South China Normal University, Guangzhou 510631, PR China

<sup>b</sup> Center for Urban and Environmental Change, Department of Earth and Environmental Systems, Indiana State University, Terre Haute, IN 47809, USA



## ARTICLE INFO

### Article history:

Received 30 July 2015

Received in revised form 15 June 2016

Accepted 24 June 2016

Available online xxxx

### Keywords:

Consistent LST

Landsat TIR imagery

Temporal segmentation

Gaussian process

Time series

## ABSTRACT

Land surface temperature (LST), derived from satellite thermal infrared (TIR) sensors, is a key variable for characterization of urban heat island, modeling of surface energy balance, estimation of evapotranspiration and soil moisture, and retrieval of air temperature. Among the satellite TIR sensors in operation, Landsat TIR sensor provides the only feasibility for long-term reconstruction of a LST dataset for environmental applications. However, a holistic technique is not currently available to generate spatially and temporally continuous LSTs from Landsat due to its 16-day revisit frequency, impact of atmospheric conditions and the SLC (Scan Line Corrector) -off gap. Previous algorithms had been developed to overcome these limitations, it is still not possible to generate LSTs at any desired date with consistent accuracy and corrections. Therefore, this study aimed to devise an algorithm to reconstruct consistent, daily LSTs at Landsat spatial resolution based solely on Landsat imagery. By selecting Beijing, China, as the study area, a total of 512 images from 1984 to 2011 were downloaded from the USGS online portal and were consistently calibrated to surface reflectance and brightness temperature. The cloud-, cloud shadow-, and snow-contaminated pixels were excluded according to quality flags; and a further screening procedure based on temporal information of Landsat spectral bands 2, 4, and 5 was conducted. Brightness temperatures were converted to LSTs through the single channel algorithm with input of water vapor from the NCEP Reanalysis dataset. Field LSTs were collected from 11 weather stations in Beijing in the year of 2008, 2009, and 2010. The proposed algorithm included four modules: Data filtEr, temporal segmentation, periodic and trend modeling, and GAussian process (DELTA). Accuracy assessment showed that, compared with the *in situ* LSTs from weather stations, satellite-derived LSTs inverted through the single channel algorithm had an average accuracy of 2.3 K. Further comparison between LSTs reconstructed from the DELTA algorithm and those collected from weather stations in the year 2008 yielded a mean error of 3.5 K. Twelve LST maps reconstructed from the DELTA in 2000 showed that LSTs of different land covers exhibited similar seasonal patterns and reached their maximal values in June/July. Using LST of every August 15th as an example, the SUHI (surface urban heat island) intensity of Beijing was computed, which ranged from 3.3 K to 5.3 K from 1984 to 2011, with an increase pattern of LST in both rural and urban areas.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction

Land surface temperature (LST) data derived from satellite thermal infrared (TIR) imagery is a crucial valuable that has been utilized for quantifying surface urban heat island (SUHI) effect (Imhoff, Zhang, Wolfe, & Bounoua, 2010; Streutker, 2003; Tomlinson, Chapman, Thornes, & Baker, 2012), estimating soil moisture and evapotranspiration (Anderson, Allen, Morse, & Kustas, 2012; Carlson, 2007; Holzman, Rivas, & Piccolo, 2014), modeling surface energy fluxes (Friedl, 2002; Mallick et al., 2014), and retrieving spatially continuous air temperature

(Kloog, Nordio, Coull, & Schwartz, 2014; Shamir & Georgakakos, 2014; Zhu, Lü, & Jia, 2013). LST from satellite TIR imagery derived through the radiative transfer equation has aroused increasing attention since the 1970s (McMillin, 1975). Currently, a series of satellite sensors are in operation to deliver TIR data, such as AVHRR, Landsat TM/ETM +/TIRS, MODIS, ASTER, and GOES. However, thermal imagery provided by the Landsat series represents the only long-term TIR observations at the medium scales suitable for climatological and environmental applications (Schott et al., 2012; Weng, 2009). Nevertheless, to date, generating a LST dataset at daily interval is still highly challenging, even by applying data fusion algorithms such as STARFM (Gao, Masek, Schwaller, & Hall, 2006) and SADFAT (Weng, Fu, & Gao, 2014). This difficulty is owing to data gaps caused by poor atmospheric conditions (e.g., cloud contaminations), the SLC-off (ETM + sensor) gap and the 16-day revisit frequency of Landsat.

\* Corresponding author at: Center for Urban and Environmental Change, Department of Earth and Environmental Systems, Indiana State University, Terre Haute, IN 47809, USA.

E-mail address: [qweng@indstate.edu](mailto:qweng@indstate.edu) (Q. Weng).

Recent years witnessed the emergence of thermal downscaling algorithms that are developed to enhance the spatial and/or temporal resolution of TIR imagery. Thermal downscaling refers to the scaling process of converting remote sensing TIR data from low to high resolution and generally falls into categories of spatial and temporal sharpening (Weng et al., 2014). The generation of LSTs at high spatial resolution is usually fulfilled by employing auxiliary data of high spatial resolution through statistical or physical approaches. Kustas, Norman, Anderson, and French (2003) exploited the relationship between vegetation index and LST in a disaggregation procedure to derive LSTs at the agricultural field scale. The algorithm was further adapted to downscale LSTs over urban areas by using the relationship between LST and impervious fraction (Essa, van der Kwast, Verbeiren, & Batelaan, 2013). The physical method, such as the Pixel Block Intensity Modulation developed by Liu and Moore (1998), was refined to redistribute LSTs into fine pixel scale according to different scaling factors (Nichol, 2009; Stathopoulou & Cartalis, 2009). On the other hand, temporal sharpening, by utilizing the high temporal resolution of geostationary satellites, such as GOES and Meteosat Second Generation, can produce LST images to model diurnal temperature cycles (Inamdar, French, Hook, Vaughan, & Luckett, 2008; Weng & Fu, 2014b; Zakšek & Oštir, 2012).

Compared to the myriad studies in the spatial or temporal thermal sharpening, the retrieval of LSTs under cloudy conditions received much less attention. Cloud contaminations and other poor atmospheric conditions should be considered in generating a long-term LST dataset (Weng & Fu, 2014a). Undetected clouds may produce LST values as low as 230 K or generate extremely small discrepancies so that cloud contaminated pixels may be assumed valid (Bulgin, Sembhi, Ghent, Remedios, & Merchant, 2014). Accurate estimation of LSTs under cloudy conditions requires both the identification of cloud-contaminated pixels as well as effective techniques for inference. Jin (2000) developed a neighbor pixel (NP) technique to spatially and temporally interpolate MODIS LSTs under cloudy conditions from neighboring pixels based on the surface energy budget. Lu, Venus, Skidmore, Wang, and Luo (2011) refined the NP algorithm by including temporally neighboring pixels; their results showed that the temporal method was better than the original spatial technique. However, these interpolation methods are not developed specifically for the Landsat data, and thus, cannot be directly applied to Landsat imagery.

Reconstruction of a long-term LST dataset necessitates the development of an algorithm that can transcend the techniques of thermal sharpening and LST interpolation under cloudy conditions to produce LSTs of both high spatial and temporal resolutions. The emergence of the spatial-temporal fusion algorithm - STARFM (Gao et al., 2006) that blends different sensors to generate daily reflectance at fine spatial resolution - holds great potential for deriving a long-term LST dataset. Liu and Weng (2012) generated a series of synthetic reflectance and LSTs by using the STARFM for a time-dependent epidemiological study in Los Angeles. Huang, Wang, Song, Fu, and Wong (2013) applied the STARFM to predict daily LSTs by taking light reflection and refraction among ground objects and considering neighboring spatial effects by incorporating a bilateral filter. Weng et al. (2014) modified and improved the original STARFM algorithm to generate daily LSTs at Landsat resolution by considering annual temperature cycle and urban landscape heterogeneity. Wu, Shen, Zhang, and Götsche (2015) presented a spatio-temporal integrated temperature fusion model to extend the fusion method to fuse multiple satellite sensors, including Landsat TM/ETM+, Terra MODIS LSTs, GOES Imager, and MSG SEVIRI. Despite all these progresses, existing fusion algorithms are still subject to several key limitations and cannot directly be used for generating a consistent, long-term LST dataset. The first limitation is that LSTs under cloudy conditions cannot be interpolated if the input images are cloud-contaminated, which is common for areas experiencing frequent cloud coverage. In addition, uncertainties remain in selecting the best imagery pairs as the inputs for predictions. Thus, the accuracy of the data fusion algorithms (e.g., STARFM, SADFAT) for deriving LSTs has not been fully

assessed. The third limitation is that these algorithms are not effective in generating LSTs for areas where disturbance events, such as deforestation, forest degradation, desertification and other land cover and land use changes, occur (Hilker et al., 2009; Julien & Sobrino, 2012), since the corresponding LST variations are not stationary over time. Finally, the inter-annual trend within LST variations cannot be captured by these data fusion algorithms. The last issue does not pose a big challenge for predicting LSTs over a short time period; however, the maximum annual trend change may reach as high as 0.34 K (Julien & Sobrino, 2012). Therefore, it is highly desirable to develop a new technique that can overcome these limitations and generate consistent, long-term LSTs.

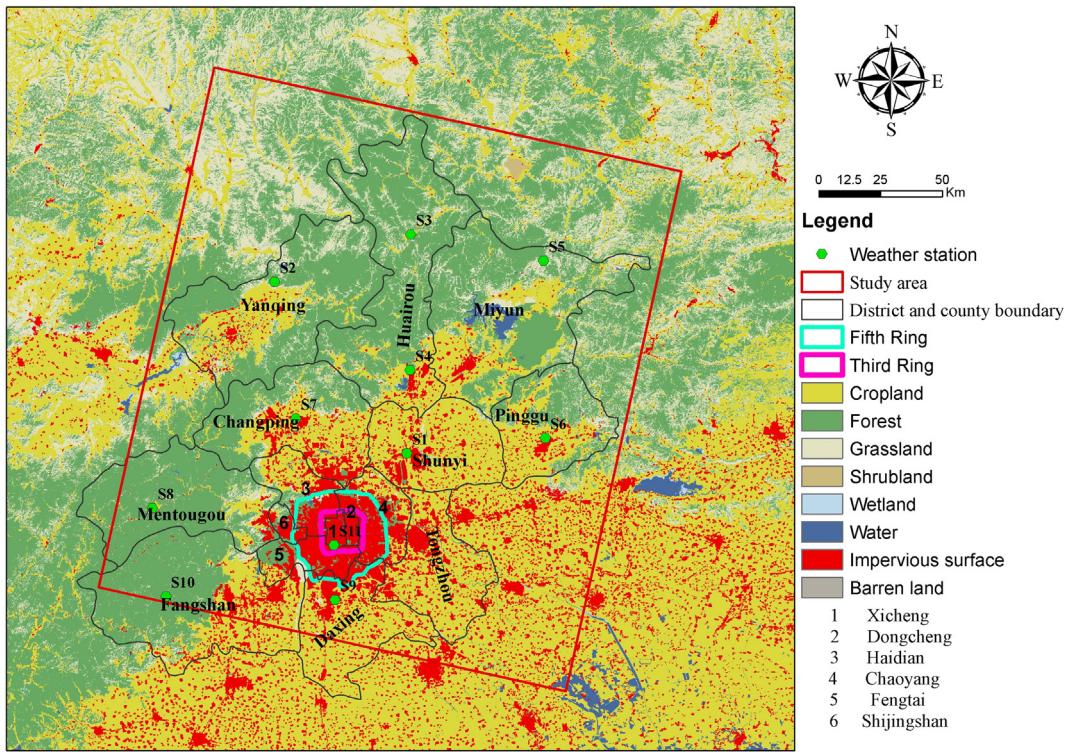
Consistent time series LSTs are of prime importance for assessing climate change of different scales (Jin & Dickinson, 2002; Jin, Dickinson, & Zhang, 2005; Sun, Pinker, & Kafatos, 2006). Recently, GEO Global Urban Supersite Initiative identified the time series analysis of the urban heat island effect and environmental impacts over “megacities” as one of the key activities (Weng, 2014). These efforts explicitly refer to the utilization of time series consistent LSTs, because of the synoptic coverage of remotely sensed data, in characterizing thermal landscape patterns from both inter- and intra-annual perspectives. A long-term LST dataset of high quality can benefit analyses of impact of urbanization on thermal characteristics. Therefore, the objective of this study is to develop an algorithm that allows reconstructing historical LST measurements at daily interval based solely on irregularly spaced Landsat imagery. Instead of blending data among different satellite sensors, this algorithm takes advantage of unevenly distributed time series Landsat imagery. The algorithm is then applied to Beijing, China, to reconstruct LSTs from 1984 to 2011, and to assess the change in the SUHI intensity using derived LSTs.

## 2. Study area and data-preprocessing

### 2.1. Study area

The study area consists of both metropolitan and rural areas of Beijing. The metropolis, located in the northern tip of the roughly triangular North China Plain, has 14 urban and suburban districts and 2 rural counties (Fig. 1). Beijing experiences elevation decrease from the northwest to the southwest with the mountains in the north and northwest shielding the city from the encroaching desert steppes. This region of China exhibits a typical temperate continental climate generally characterized by hot and wet summers and dry and cold winters. The study area covers >95% of the Beijing metropolis, captured by the Landsat scene of path/row 123/32. The global land cover mapping project (GlobalLand30) (Chen et al., 2015) identifies eight land covers including croplands, forest, grassland, shrubland, wetland, water, impervious surface, and barren land in the study area for the baseline of year 2010.

Since the 1980s, Beijing underwent rapid urban growth. The urban area of Beijing increased from 183.84 km<sup>2</sup> in 1973 to 1209.97 km<sup>2</sup> in 2005 with an annual expansion rate of built-up area at 32.07 km<sup>2</sup> (Mu et al., 2007). The population reached 21.51 million in 2014 and the average population density was 1311 persons/km (Beijing Municipal Statistical Bureau 2014). Beijing's gross domestic product (GDP) value was merely 10 billion in 1978 and soared to almost 1980 billion in 2013, ranking the most developed and prosperous in China (National Bureau of Statistics of China, 2013). The city now has a post-industrial economy dominated by tertiary sector diversified by financial services, information technology, and scientific research, etc. The intensive urbanization in the past decades has also caused a series of environmental issues, such as haze pollution, extreme rainstorms, and water contamination. It has been reported that surface temperature and the urban heat island (UHI) intensity in Beijing increased at the rate of 0.25 °C and 0.31 °C per decade, respectively, after 1981 (Lin & Yu, 2005). Since Beijing have been experiencing a serious UHI, many studies have been reported to analyze the thermal characteristics (Gong, Li, Wang, Chen, & Hu, 2006;



**Fig. 1.** The study area Beijing, China. It consists of 6 urban districts-Xicheng, Dongcheng, Haidian, Chaoyang, Fengtai, and Shijingshan, and 8 suburban districts-Changping, Daxing, Fangshan, Huairou, Mentougou, Pinggu, Shunyi, and Tongzhou, and 2 rural counties-Miyun and Yanqing. Fig. 1 also shows the land cover types identified by the GlobalLand30 (baseline year of 2010) and the distribution of weather stations (total 11 stations, named from S1–S11) for the study area.

Quan et al., 2014; Song & Zhang, 2003) and the adverse effect of high temperature (Ji, Liu, & Xuan, 2006; Liu et al., 2011). Therefore, generation of a time-series LST dataset can help understand the impact of urban growth on environment and public health, and is also of significance for sustainable urban development in Beijing.

## 2.2. Data and pre-processing

All L1 T Landsat images available for the study area were downloaded through the USGS Center Science Processing Architecture (ESPA) On Demand Interface. The datasets included surface and top-of-atmosphere (TOA) reflectance and brightness temperature processed by the Landsat Ecosystem Disturbance Adaptive Processing System (Masek et al., 2006). Using the quality flag, images with cloud coverage < 90% were selected, and a total of 512 images were acquired, covering the period from 1984 to 2011 (18 images on average per year). A further screening procedure *Tmask* based on the multi-temporal images were adopted to remove the outliers possibly caused by cloud, cloud shadow and snow contaminations (Zhu & Woodcock, 2014a). In addition, pixels of data gaps in the SLC-off images were removed in light of the metadata. Then, all the images were resampled to 120 m and subset to the study area. The retrieval of LSTs from brightness temperature was accomplished by using the single channel algorithm (Jimenez-Munoz & Sobrino, 2003) since its reported computation error was close to 1.5 K for water vapor from 0.5 to 3 g/cm<sup>2</sup>. The input of water vapor was derived from the NCEP Reanalysis data (Kalnay et al., 1996). Land surface emissivity values were first calculated from both the classification based method (Snyder, Wan, Zhang, & Feng, 1998) and the NDVI threshold method (Sobrino et al., 2008). The comparison of the emissivity values computed from the GlobalLand30 with those derived from the NDVI data showed the mean difference of 0.01. According to Jiménez-Muñoz and Sobrino (2006a, 2006b), the 0.01 emissivity difference may produce the LST retrieval error < 0.4 K. Thus, the NDVI threshold method was adopted because of its feasibility

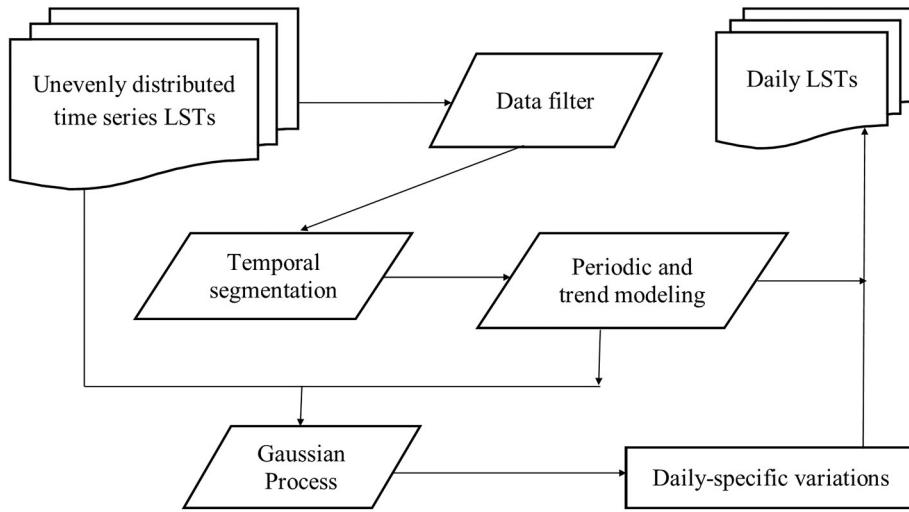
to derive time series emissivity dataset. In addition, surface temperature measurements from 11 weather stations in the years of 2008, 2009, and 2010 were collected and compared with the corresponding satellite-derived LSTs to assess the accuracies of satellite-derived LSTs and LSTs reconstructed by the DELTA algorithm. The accuracy assessment was performed based on a point (weather station temperature record)-pixel (satellite-derived LST data) comparison scheme.

## 3. Methodology

The DELTA algorithm consisted of four modules: (1) data filter, (2) temporal segmentation, (3) periodic and trend modeling, and (4) Gaussian process. Each module was applied on the pixel-by-pixel basis. Fig. 2 shows the flowchart of the DELTA algorithm to generate daily LSTs at 120 m. The data filter module removed outliers and constrained surface reflectance and LSTs within the set range. Temporal segmentation aimed at dividing time series LSTs into relatively stationary segments to reduce the effect of disturbance events on LST reconstructions. In the phase of periodic and trend modeling, the unevenly distributed time series LSTs were fitted by a parametric model consisting of annual temperature cycle (ATC) and trend variations. Then, the residuals between the LST observations and model predictions were analyzed by a non-parametric Bayesian technique, Gaussian process regression (GPR). Finally, LSTs at the daily interval were reconstructed by adding up the variations of ATC and trend component, and daily-specific anomalies inferred by the GPR.

### 3.1. Data filter

The data filter ensured that only valid surface reflectance and LST measurements over time were utilized. First, the filter screened out the pixels which had the values of surface reflectance beyond the range from 0 to 10,000 (the reflectance data downloaded was first scaled by a factor of 10,000) and LSTs outside the range from 250 to



**Fig. 2.** The flowchart of the algorithm DELTA consisting of four modules: data filter; temporal segmentation; periodic and trend modeling; and Gaussian process. The final daily LSTs reflected annual temperature cycle, inter-annual trend, and daily-specific variations.

340 K to exclude extreme outliers. In addition, quality flags from both the metadata and the *Tmask* were used to ensure that only clear-sky pixels were selected and input into the DELTA.

### 3.2. Temporal segmentation

The temporal segmentation procedure was based on the consideration that disturbance events, such as deforestation, desertification, and urbanization, can induce non-stationary LST variations over time. The division of the whole time series into a sequence of discrete segments was conducted because different land covers and surface conditions may generate different temporal characteristics in LSTs. The treatment of individual data segments thus eased the LST reconstruction without impacting the non-stationary variations.

The segmentation method was adapted from the Continuous Change Detection and Classification (Zhu & Woodcock, 2014b) algorithm and was applied only to band 2, band 3, band 4, band 5, and band 7 to reduce the computation time. First, it identified land cover changes by using the time series additive model (Eq. (1)).

$$P(i, d) = a_i + b_{1i} \cos(\omega d) + b_{2i} \sin(\omega d) + c_i d \quad (1)$$

where  $P(i, d)$  is the value predicted for surface reflectance at Julian date  $d$ ,  $\omega$  is the angular frequency,  $a_i$ ,  $b_i$  and  $c_i$  are the coefficients for the mean value, intra- and inter-annual changes. The change was flagged if the normalized difference exceeded the pre-define thresholds as

showed in Eq. (2) in three consecutive days.

$$\frac{|P(i, d) - P(i, d)|}{3 \times RMSE} > 1 \quad (2)$$

Given that land cover change may happen in the first and last observations during the process of the model initialization, another three conditions (Eq. (3)) were listed below to identify abnormal observations for the first and last observations of the model initialization and the observations between them.

$$\frac{1}{k} \sum_{i=1}^k \frac{|c_i(d)|}{3 \times RMSE_i} > 1 \text{ or } \frac{1}{k} \sum_{i=1}^k \frac{|P(i, d_1) - P(i, d_1)|}{3 \times RMSE_i} > 1 \text{ or } \frac{1}{k} \sum_{i=1}^k \frac{|P(i, d_n) - P(i, d_n)|}{3 \times RMSE_i} > 1 \quad (3)$$

With  $d_1$  and  $d_n$  the Julian dates for the first and last observation during model initialization, and  $T_{model}$  is the total time used for the model initialization.

### 3.3. Periodic and trend modeling

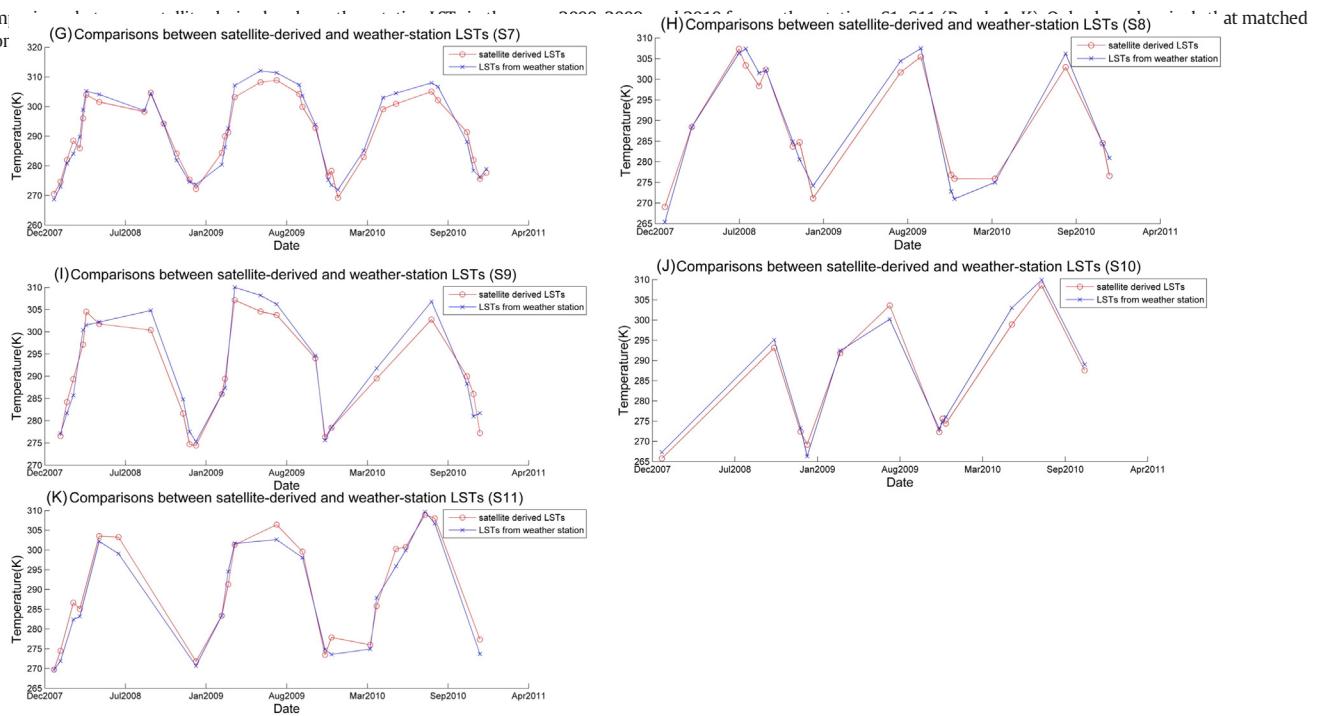
In this phase, the parametric model (Eq. (1)) consisted of two sinusoidal functions, a mean value, and a linear trend. The two sinusoidal functions could characterize the annual temperature cycle (ATC), which was part of the fluctuation attributed to the Earth's changing position over the course of the year (Bechtel, 2012; Thomson, 1995; Weng & Fu, 2014a). ATC parameters were effective only if satellite

**Table 1**  
Comparisons between satellite-derived and weather-station LSTs.

Station name	Correlation coefficient	Mean error (K)	Mean absolute error (K)	Land cover type
S1	0.94	-0.9	2.4	Impervious surface
S2	0.94	-0.8	2.7	Forest
S3	0.89	0.4	2.4	Cropland
S4	0.98	-2.5	2.8	Impervious surface
S5	0.96	-0.1	1.8	Grassland
S6	0.92	-1.0	2.3	Impervious surface
S7	0.98	-0.6	2.5	Impervious surface
S8	0.80	-0.3	2.5	Impervious surface
S9	0.95	-0.8	2.4	Impervious surface
S10	0.87	-0.6	1.8	Forest
S11	0.92	1.4	2.1	Impervious surface

Note: S1–S11 are weather stations in the study area. The distributions of the 11 weather stations can be seen in Fig. 1. The Pearson correlation coefficients are statistically significant under the two-tailed tests. The land cover type column records the land cover of each weather station.

**Fig. 3.** Com  
the locator



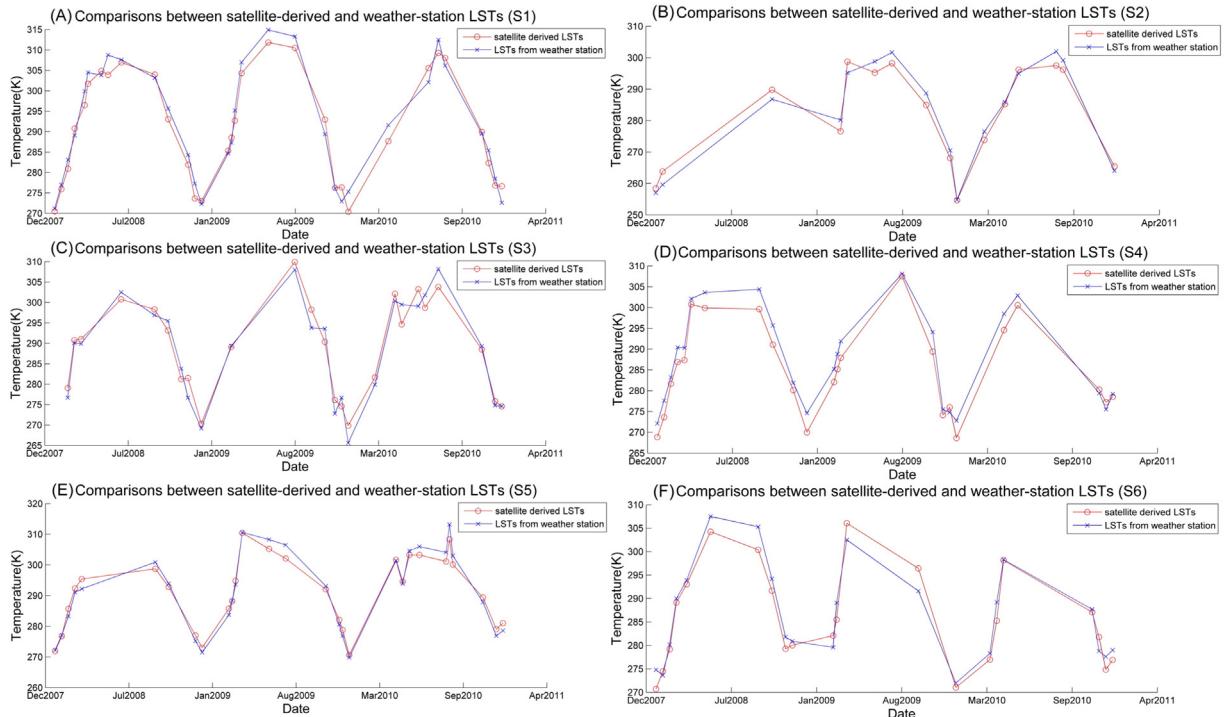
**Fig. 3 (continued).**

observations were cyclic-stationary. The mean value and the linear trend functions included in the parametric model can capture the overall and the inter-annual LST changes over the study period. Although previous studies (Huang et al., 2013; Weng et al., 2014; Wu et al., 2015) have proven that the data fusion algorithms can predict LSTs at high spatial and temporal resolutions, inter-year LST changes have so far been neglected. The inter-year LST change may become apparently obvious when long-term LST (e.g., 20 years) observations were analyzed (Julien & Sobrino, 2012). Given the factors above, the periodic and trend

modeling procedure was applied only to within individual homogeneous segments to avoid the impacts of land cover changes.

#### 3.4. Gaussian process regression

The Gaussian process (GP) module was included because the ATC and inter-annual model may not capture all the LST variations. Weng and Fu (2014a) employed an ATC model to characterize the landscape thermal patterns and revealed the overall mean root mean square



error (RMSE) of 7.4 K and the overall mean RMSE of 4.3 K by using annual, semi-annual, and the trend models to fit LSTs (Fu & Weng, 2015). Since the model residuals (differences between the model predictions and satellite observations) may be regarded to relate to daily-specific variations associated with weathers, GPR was employed to make inference of the residuals at the daily basis.

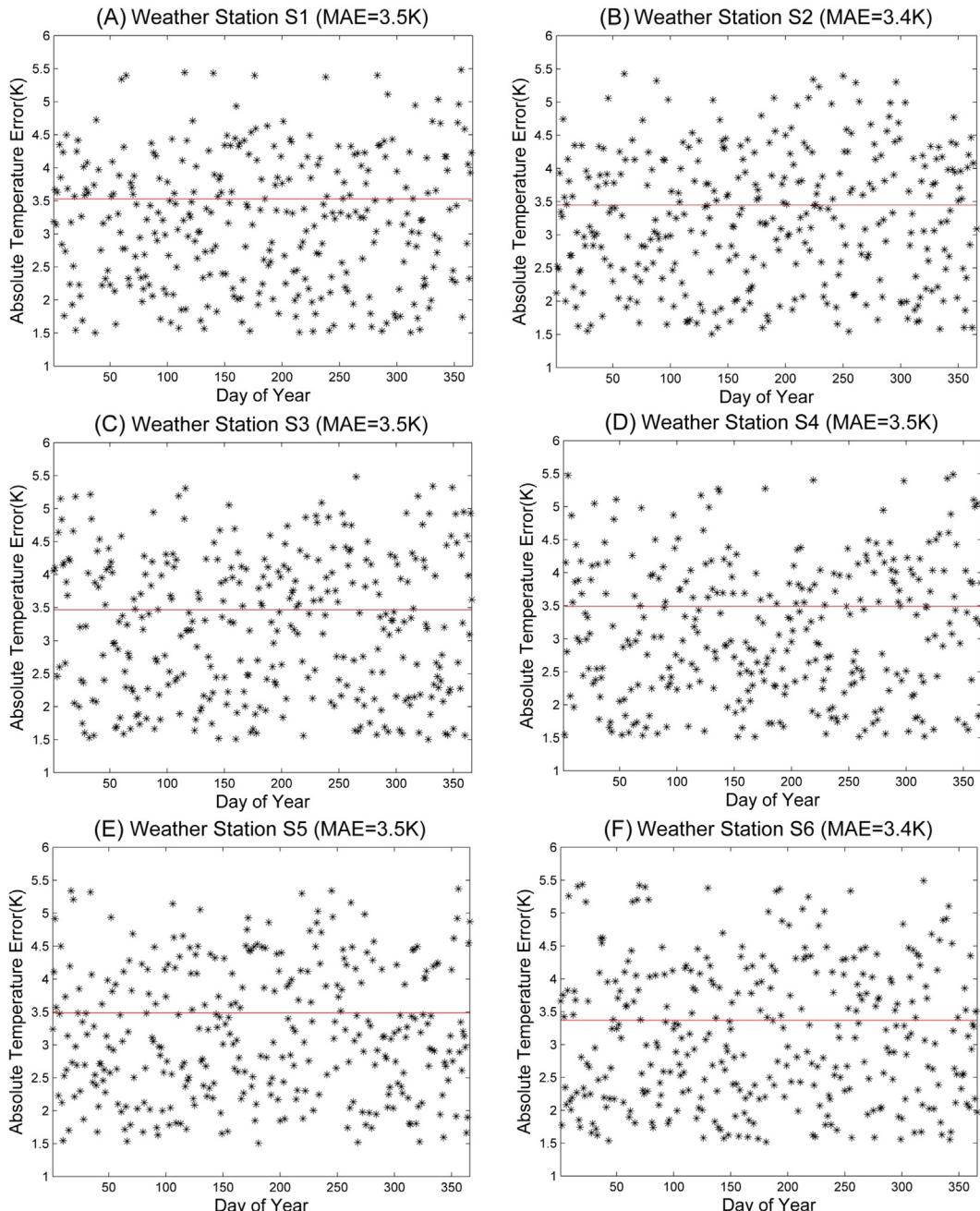
The GPR can be interpreted as a distribution over and inference occurring in the space of function from the function-space view (Rasmussen & Williams, 2006). It has been received much attention in the field of machine learning and can provide the Bayesian approach to establish the relationship between the input  $x$  (Julian dates) and the output variable  $f(x)$  (model residual computed as the difference between LST observation and the model prediction

from section 3.3) with kernels in the form of (Eq. (4)):

$$f(x) = \sum_{i=1}^n \alpha_i k(x_i, x) \quad (4)$$

where  $\alpha$  is the kernel used to assign weight,  $k$  is the kernel function evaluating the covariance (similarity) between the input variables (Julian dates)  $x_i$  ( $i = 1, \dots, n$ ) and the test data  $x$  (Julian dates). For the current study, based on a trial and error test, the kernel covariance function of “Materniso” was adopted (Eq. (5)).

$$k(r) = \sigma_s^2 \frac{2^{1-v}}{\pi(v)} \left( \sqrt{2v} \frac{r}{l} \right)^v K_v \left( \sqrt{2v} \frac{r}{l} \right) + \delta * \sigma_n^2 \quad (5)$$



**Fig. 4.** Daily temperature difference (black stars) between values estimated from DELTA and from weather stations S1–S11 (Panels A–K) for year 2008. The red line is the mean absolute error (MAE). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where  $K_v$  is a modified Bessel function,  $\mathcal{T}$  is the gamma function,  $r$  is the distance between  $x$  and  $x_i$ ,  $v, l$  are the positive parameters,  $\sigma_s$  is the signal standard deviation,  $\sigma_n$  is the noise standard deviation, and  $\delta$  is the Kronecker's symbol. The kernel was parameterized collectively by a hyperparameter  $\theta = [\sigma_s, v, l, \sigma_n]$ . Now suppose that the observed variable was formed by noisy observations of the true underlying function  $y = f(x) + \varepsilon$  and the noise was additive independently identically Gaussian distributed with zero mean and variance  $\sigma_n^2$ , under the prior assumption, the joint distribution of the observed target values and the function values at the test locations were given by the Eq. (6) ( $x_*$  denotes the test data (Julian dates)).

$$\begin{pmatrix} y \\ f_* \end{pmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} k(x, x) + \sigma_n^2 I & k(x, x_*) \\ k(x_*, x) & k(x_*, x_*) \end{bmatrix} \right) \quad (6)$$

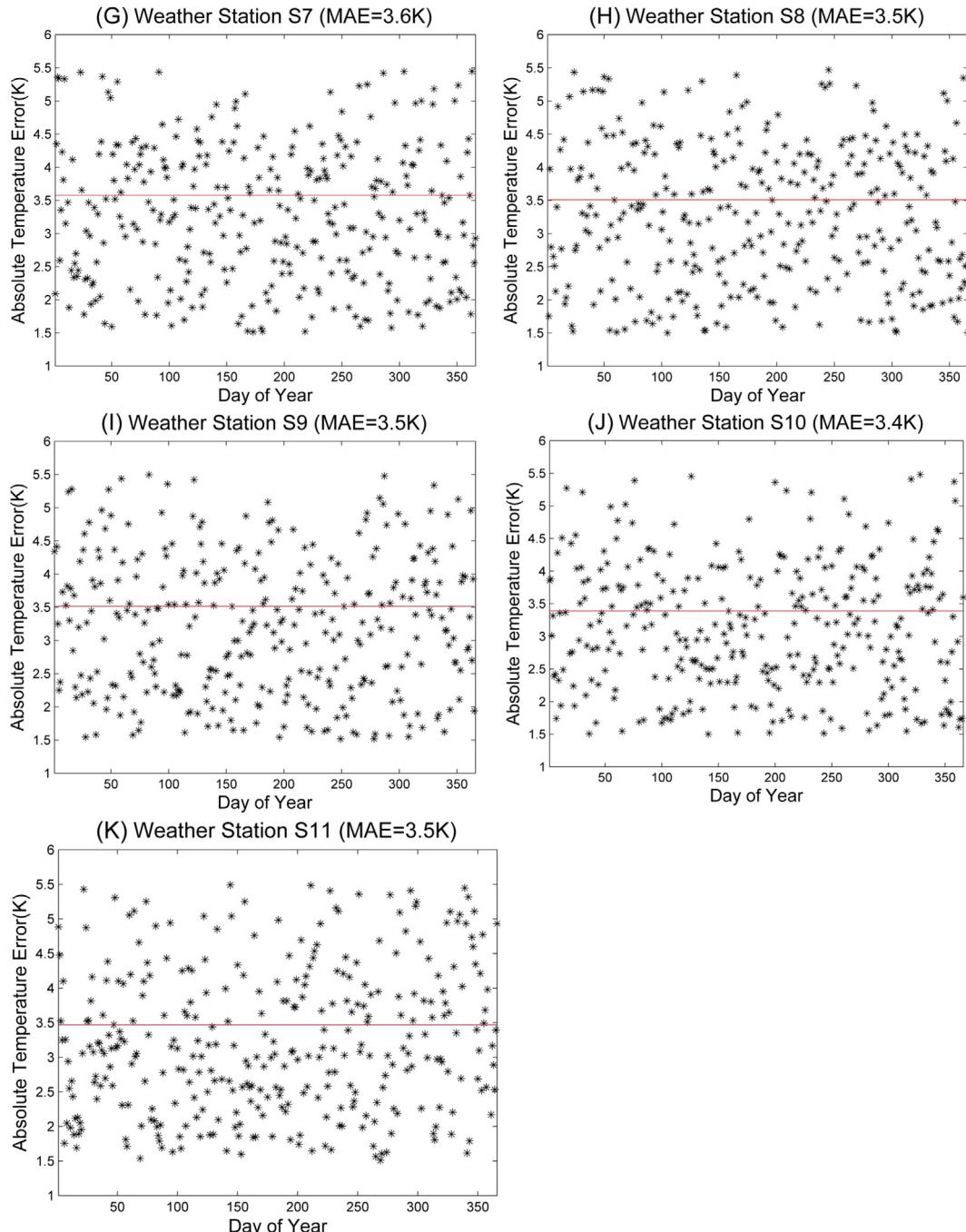


Fig. 4 (continued).

For the prediction purposes, the GPR was employed by computing the posterior distribution over the unknown values  $f_*$  with the hyperparameters  $\theta$  typically selected by maximizing Type-II Maximum Likelihood, using the marginal likelihood of the observations. Therefore, the predictive mean  $\bar{f}_*$  (predicted LST residuals) and variance  $v(\bar{f}_*)$  can be estimated using the Eq. (7).

$$\bar{f}_* = k_*^T (k + \delta_n^2 I)^{-1} y, \quad v(\bar{f}_*) = k(x_*, x_*) k_*^T (k + \delta_n^2 I)^{-1} k_* \quad (7)$$

with

$$k_* = (k(x_*, x_1), \dots, k(x_*, x_n))^T$$

Based on the four modules, daily LSTs at 120 m resolution were reconstructed by adding up the ATC, trend, and daily-specific variations. Given that the segmentation procedure may produce segments with <6 pixel dates, these pixel dates were merged with other longer segments to ease the computation.

#### 4. Results

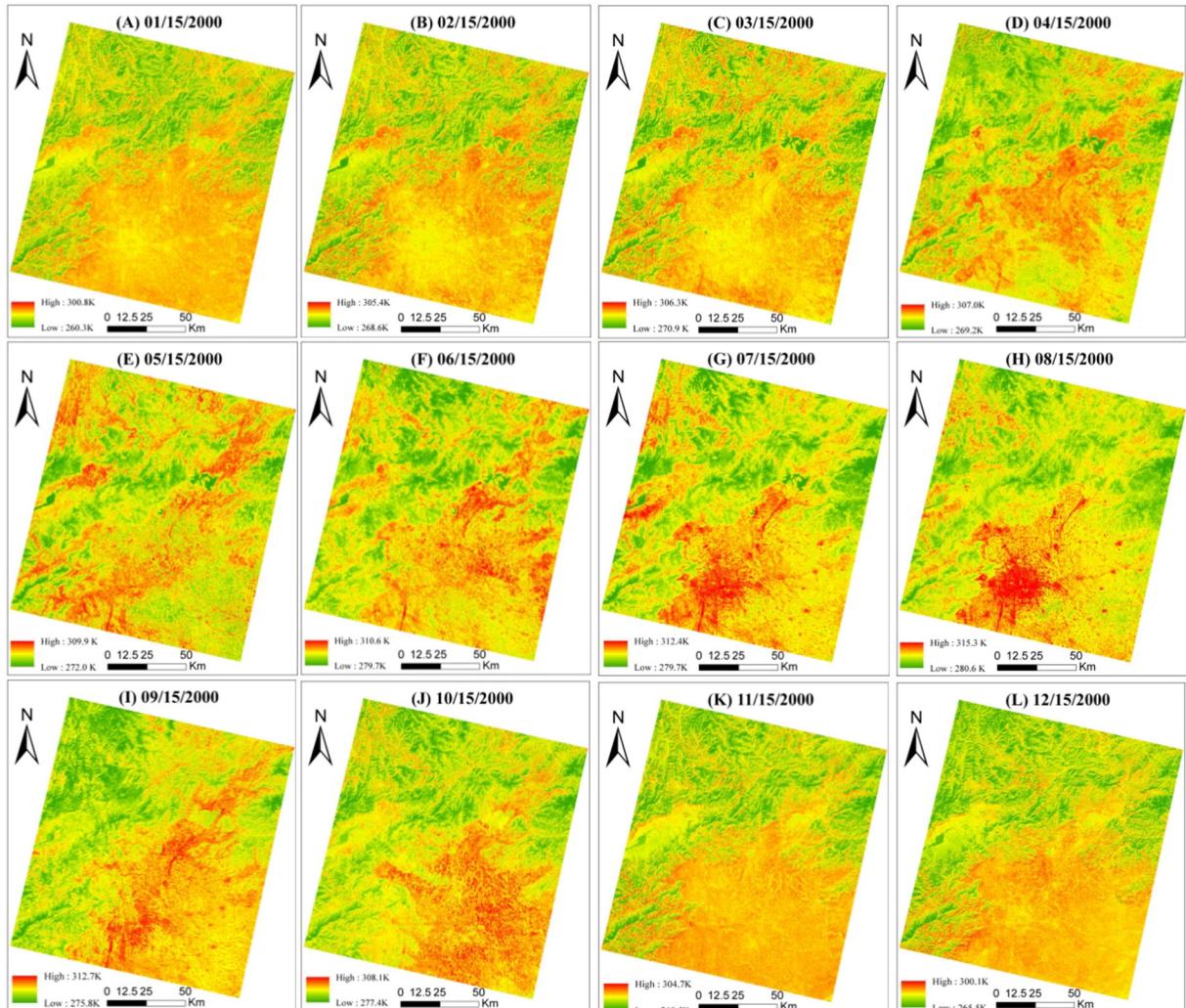
Accuracy assessment was performed to compare satellite derived LSTs with *in situ* LSTs from 11 weather stations (Fig. 1S1–S11). Since the ground measurements were recorded hourly, the *in situ* LSTs did not match the satellite overpassing times (approximately 10:30 am local time). This study utilized linear interpolation method to estimate reference LSTs at the satellite acquisition time based on ground measurements at 10 am and 11 am.

The accuracy of the DELTA algorithm to reconstruct daily LSTs was further assessed against the LST measurements from weather stations. The daily LSTs in 2008 were generated and compared with the field measurements. The correlation coefficient (CC), mean error (ME), and mean absolute error (MAE) were computed as accuracy measures.

##### 4.1. Accuracy assessment for satellite-derived LSTs

Satellite derived LSTs that matched the locations of weather stations were extracted and compared with the corresponding field LST measurements from 2008 to 2011. Only clear-sky pixels screened by both

the quality flags and *Tmask* were selected for the comparison. Validation results show that the Pearson correlation coefficient ranged from 0.80 to 0.98 (Table 1), indicating that the good agreement between the satellite-derived LSTs and *in situ* LSTs was achieved. All correlations were statistically significant (a two-tailed test) with *p-value* < 0.01. Weak correlation (coefficient: <0.9) was observed for the weather stations S3, S8, and S10. Further analysis showed that the biophysical settings of these weather stations were distinct from the land covers of corresponding pixels that the stations were located. For instance, Station S3 was located in a small plot of cropland, while the land cover type of the corresponding pixel was impervious surface. Panels A–K in Fig. 3 further show that satellite-derived LSTs exhibited similar variation patterns over the time. However, the validation also indicates that in general LSTs provided by satellite sensors were less than ground measurements. In this study, satellite LSTs were retrieved from the TIR sensors by the single channel algorithm with input of water vapor from the NCEP Reanalysis dataset. The dependence of water vapor for the single-channel algorithm resulted from the fact that water vapor is the mainly absorber in the thermal infrared region. The water vapor data used for LST computation came from the linear interpolation between the two time points that were close to the satellite overpassing time. According to the pattern of diurnal variations of water vapor (Dai, Wang, Ware, & Van Hove, 2002), this study may have used a smaller water vapor value. Consequently, it is reasonable that satellite-derived LSTs showed smaller values than field measurements. In addition, Table 1 shows that the mean absolute error between



**Fig. 5.** LST maps (A–L) for the year 2000. Only LST images of 15th of each month in the year 2000 are listed. The corresponding statistics by land cover for each date are shown in Table 2.

**Table 2**

Average LST by land cover in Beijing in 2000.

Date	Cropland	Forest	Grassland	Shrubland	Water	Wetland	Barren	Impervious
01/15	275.8	269.5	273.1	273.7	270.1	271.5	273.5	274.7
02/15	279.4	274.3	277.6	278.1	272.4	273.3	277.5	278.0
03/15	286.7	282.2	285.4	286.4	278.8	278.5	286.2	285.9
04/15	294.1	290.4	292.3	295.0	286.7	286.9	291.4	294.4
05/15	300.2	299.7	303.2	304.7	295.9	295.9	305.4	303.4
06/15	305.7	301.8	305.5	307.8	301.1	300.7	308.4	308.6
07/15	305.2	303.9	306.3	306.8	303.2	302.9	308.0	309.5
08/15	300.9	299.4	301.2	301.0	299.9	300.2	303.1	305.0
09/15	297.6	294.1	295.8	297.3	293.9	295.4	296.5	299.4
10/15	291.1	285.1	288.2	289.4	286.0	287.8	290.1	290.7
11/15	280.9	276.9	279.1	278.5	277.6	279.4	279.1	281.0
12/15	275.6	270.6	272.9	273.0	270.7	273.6	274.2	275.7
Annual average	291.1	287.3	290.5	291.0	286.4	287.1	291.7	292.2

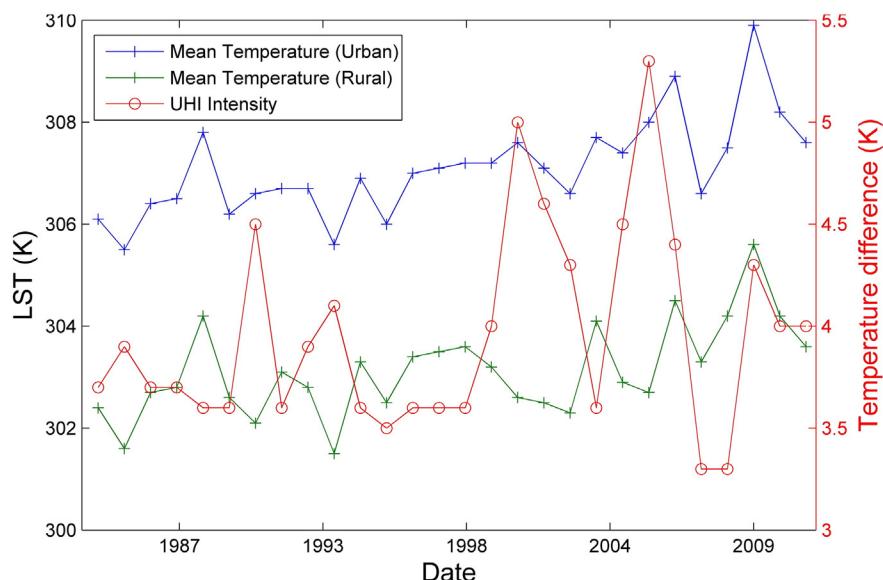
Note: Temperatures maps (A–L) from Fig. 5 were used for performing the statistics by land cover delineated by GlobalLand30. Unit in Kelvin.

satellite-derived and weather-station LSTs is at least 1.8 K, larger than the error of 1.5 K reported by Jimenez-Munoz and Sobrino (2003). The difference between the two studies may be attributed to different methods in collecting field temperatures. For the current study, temperature measurements were recorded hourly by using thermometers positioned on the ground surface. The *in situ* LSTs by Jimenez-Munoz and Sobrino (2003) were obtained by combining soil and vegetation temperatures based on the fractions of bare soil and vegetation inside the field of view of the field radiometer. The field measurements by Jimenez-Munoz and Sobrino (2003) were more accurate than LSTs from the weather stations especially for mixed pixels. In addition, thermal conditions of the study area were different from the test sites of Jimenez-Munoz and Sobrino (2003), so were different image acquisition dates and times. However, the absolute mean error of 1.8–2.8 K (Table 1) among the 11 weather stations suggests that the accuracy of satellite-derived LSTs was in the acceptable range.

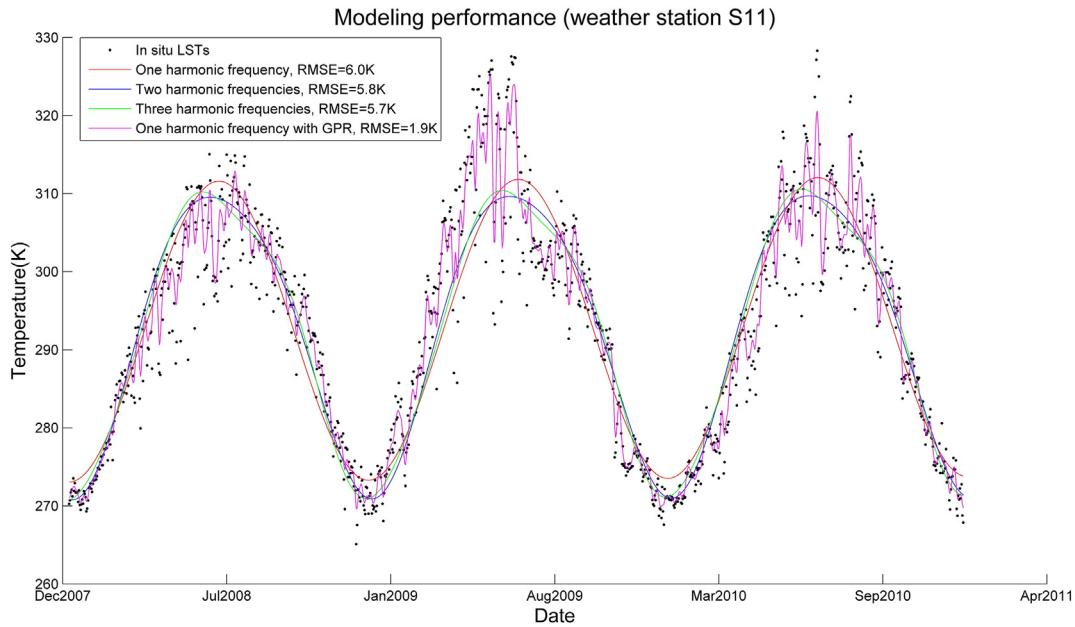
#### 4.2. Accuracy assessment for LSTs reconstructed by DELTA

To assess the effectiveness of DELTA to generate LSTs, daily *in situ* LSTs collected from the 11 weather stations (S1–S11 in Fig. 1) were compared against the reconstructed daily LSTs in year 2008. The original

pixels used to reconstruct the consistent LST datasets were excluded before the comparison. Fig. 4 shows that the mean absolute error (MAE) ranged from 3.4 to 3.5 K (~1% of the average LST). Since satellite-derived LSTs had an average retrieval bias of 2.3 K, it is expected that reconstructed daily LSTs possessed a larger error. The relatively even MAE among 11 weather stations, as evident in Fig. 4, indicates that the DELTA algorithm performed well. This result should be mainly attributed to the inclusion of GPR to estimate daily-specific LST variations associated with weather conditions. Without GPR, the DELTA algorithm could only predict the harmonic and inter-year variations within individual stationary segments. For example, the modeling performance for the annual temperature cycle, inter-annual trend, and the daily specific variations reduced the LST prediction error to 1.9 K (Fig. 7). Although the Bayesian model was powerful in learning the daily specific variations, the GPR technique was still influenced by the Landsat revisiting time and poor atmospheric conditions. The prediction errors arose since these factors limited the images that could be used to provide the input of weather-related LST variations. As a result, it is impossible for the GPR to predict all the daily-specific LST variations accurately. Overall, based on the *in situ* LSTs from the 11 weather stations, the mean error of reconstructed daily LSTs at the Landsat resolution was 3.5 K.



**Fig. 6.** Mean LSTs for the regions within the Third Ring (referring to as the urban area) and areas that are non-mountainous and outside of the Fifth Ring (referring to as the rural area). The UHI intensity was calculated as the difference in LST between the urban and rural areas.



**Fig. 7.** Modeling performance for the parametric model (Eq. (1)) with different number of harmonic frequencies and the GPR. The root mean square error (RMSE) is calculated between *in situ* LSTs and the model predictions. The *in situ* LSTs were collected from the weather station S11 as shown in Fig. 1. The combined use of one harmonic frequency with the GPR reduced the RMSE to 1.9 K.

#### 4.3. UHI dynamics in Beijing, 1984–2011

One of interesting applications of consistent long-term LSTs generated by the DELTA algorithm is to characterize and quantify the spatio-temporal dynamics of SUHI. Here LST maps of the 15th day of each month in the year 2000 were produced to reveal changes in the spatial and temporal patterns (Fig. 5 A–L). Table 2 shows LST statistics by land cover/use type. The result shows an annual mean LST value of 292.2 K for impervious surface, 291.7 K for barren land, 291.1 K for cropland, 291.0 K for shrubland, 290.5 K for grassland, 287.3 K for forest, 287.1 K for wetland, and 286.4 K for water. These LST statistics are in agreement with previous studies (Weng & Fu, 2014a, 2014b, Fu & Weng, 2015). In addition, the result also revealed that all the land covers had similar annual temperature cycles, *i.e.*, LSTs increased from January to June/July to reach their annual maxima and then decreased until December. Since the ATC was explicitly included in the DELTA algorithm, it is reasonable that LST statistics derived from the reconstructed LST maps displayed seasonality.

Furthermore, LST maps were generated for each August 15th from 1984 to 2011 to quantify changes in the SUHI intensity. The UHI effect refers to the increased temperatures of the dense urban areas with respect to their rural or suburban surroundings. However, there was no consensus about what constitute the representative “urban” and “rural” surfaces (Oke, Johnson, Steyn, & Watson, 1991). In this study, urban and rural areas were identified based on the land cover map. In addition, an appropriate rural reference should be from the same type of topographical setting without major elevation change (Weng, Rajasekar, & Hu, 2011). Therefore, it was assumed that the region inside the Third Ring was the urban area, while the non-mountain and non-urban region outside of the Fifth Ring constituted the rural area (Fig. 1). The SUHI intensity was calculated as the temperature difference between the two regions. Fig. 6 shows mean LSTs in both urban and rural regions and the SUHI intensity from 1984 to 2011. Despite of temporal fluctuations, the mean temperatures in the two regions exhibited an increase trend. The SUHI intensity ranged from 3.3 K (in the years 2007 and 2008) to 5.3 K (in the year 2005). Furthermore, the mean SUHI intensity was 4.2 K from 1998 to 2011, 0.4 K higher than the mean value from 1984 to 1997, suggesting an intensified SUHI in Beijing. Further analysis are warranted to correlate

the temporal SUHI evolution with the urban expansion history to explain the underlying the mechanisms.

## 5. Discussion

### 5.1. Issues related to the DELTA

The DELTA algorithm to reconstruct LSTs was mainly dependent on its ability to model ATC, trend, and daily-specific variations. It first divided time series LSTs into relatively stationary segments in which the DELTA algorithm was applied to remove the impact of land cover changes on LST modeling. As such, accurate determination of dates of land cover changes can improve the prediction. It should be noted, however, that identification of land cover changes was limited by available clear-sky pixels and the 16-day revisit time (8-day if Landsat TM and ETM + were both available). The temporal segmentation module has been reported to provide land cover change dates within 32 days later than real dates (Zhu & Woodcock, 2014b), which suggested that the DELTA cannot model LST variations for the time periods that had undergone land cover changes but were not captured by the satellite sensors.

The DELTA algorithm modeled only one annual frequency to characterize the temperature cycles over time. This assumption was reasonable in that the time series LSTs was divided into segments corresponding to different land covers. An interesting point worthy to discuss is whether increasing the number of harmonic frequencies can help characterize more complete LST variations over time. For the current study, a preliminary modeling procedure was tested to characterize LST variations by setting the number of harmonic frequencies as a free variable. The experimental results from all weather stations suggested that an increase in the number of harmonic functions did not improve significantly the modeling performance. Here *in situ* LSTs from weather station S11 were shown to demonstrate this point. Fig. 7 shows daily LSTs from station S11 and its modeling performance with the use of different number of harmonic functions. The increase of harmonic functions to three frequencies reduced the RMSE only by 0.3 K. With all the image pixels

used for modeling, the overall mean RMSE was about ~5 K even when three frequencies were used in the modeling procedure. Since that the utilization of more frequencies did not help improve the prediction, the GPR was employed to simulate daily-specific variations (also known as the model residuals between satellite-derived LSTs and model predictions). Fig. 7 confirms that the combined use of one harmonic frequency with the GPR can improve the modeling accuracy (in this case, RMSE yielded 1.9 K). Furthermore, the number of images available can affect the performance of ATC modeling. Fu and Weng (2015) showed that the frequency bias of the ATC can reach up to 5 days when the number of images decreased from 250 to 60 in a 10-year period. In this study, the ATC modeling performance depended both on temporal segmentation and the number of images available.

The performance of DELTA was influenced by GPR to infer daily-specific variations from clear-sky LSTs. The effectiveness of GPR depended upon its ability to learn the covariance structure through the input of clear-sky LSTs over the time. The DELTA algorithm assumed that LST variations which were not captured by the ATC and trend components were controlled by daily-specific weather conditions and followed a multivariate Gaussian distribution. This assumption enabled the DELTA to model weather-related LST variations through a Bayesian machine learning technique. The only factor that can affect the performance of GPR was the number of image pixels available. For instance, there were some extremely high temperature observations (>310 K) in Fig. 7. If image pixels did not include high temperature variations, it would not be possible for GPR to learn the weather patterns well and thus not be able to reconstruct LSTs accurately. In addition, GPR did not take into account of the effects of surface conditions, such as wind speed and direction and soil moisture. The DELTA algorithm assumed similar weather and surface conditions in the generation of consistent LSTs. Future researches are warranted to include these factors in the DELTA algorithm.

## 5.2. Issues related to accuracy assessment

Despite the overall reliable accuracies of satellite-derived and DELTA-reconstructed LSTs when compared with field measurements from the weather stations, it is worth noting that the use of ground temperature measurements to validate LSTs derived from satellite sensors may be problematic. LST is defined as a measurement requiring the corrections of both atmosphere and surface radiative properties that influence the emission and reflection of radiation detected by the satellite sensors (Voogt & Oke, 2003). As such, substantial differences between *in situ* and satellite-derived LSTs may arise since remote sensing measurements reflect variability on the Earth's surface properties and processes that occur at different spatial scales from those captured by field measurements (Friedl, 2002). The satellite detected thermal (or the Top of Atmosphere, TOA) signal for a given wavelength contains radiance emitted by surface features, down-welling radiance, and up-welling atmospheric radiance. The single-channel algorithm used only water vapor parameter to simulate the down-welling radiance and the up-welling atmospheric radiance and to estimate surface-emitted radiance. The simplification for the radiative transfer equation neglected the fact that these radiances also depend on other parameters, such as atmospheric temperature and surface pressure (Jimenez-Munoz & Sobrino, 2003). Even with accurate calibration of atmospheric conditions, the surface-emitted radiance should be corrected by spectral emissivity to account for the roughness of the land surface, the amount and vigor of vegetation cover, the thermal property and moisture content of the soil (Friedl, 2002) and be corrected by the elevation and slope factors. As a result, LST retrieval bias may arise from atmospheric correction, noise of the sensor, land surface emissivity, aerosols and other gaseous absorbers, angular effects, wavelength uncertainty,

full-width at half-maximum of the sensor, and the band-pass effects (Jiménez-Muñoz & Sobrino, 2006a, 2006b), as well as surface conditions, such as elevation and slope factors and soil moisture variability.

In addition, the IFOV (instantaneous field of view) of remote sensing imagery generally did not match the source area where the field measurements were collected and used for validation. For the current study, such validation data were typically collected from homogeneous surfaces, while remote sensing measurements possessed fixed IFOV and display composite signatures for most urban pixels (Lu & Weng, 2004). Surface conditions, such as land cover type and soil moisture content near the weather stations, can be different from those pixels in which the weather stations were located. In this study, a point (weather station surface temperature records) to pixel (satellite-derived LSTs) comparison scheme was utilized to perform the accuracy assessment of the DELTA algorithm. Although the result showed the average bias of 3.5 K for the reconstructed LSTs was achieved in 2008, the spatial variations of the reconstructed LSTs were not assessed. Possible improvements can be made to collect field surface temperatures in more locations (e.g., a network of random sampled locations) to validate the spatial variations of reconstructed LSTs.

## 6. Conclusions

This study devised an algorithm to reconstruct consistent, daily LSTs at Landsat resolution based solely on Landsat imagery, *i.e.*, the DELTA algorithm. The DELTA algorithm possessed some advantages compared with the algorithms by blending Landsat and MODIS imagery to generate LSTs (Gao et al., 2006; Huang et al., 2013; Weng et al., 2014; Wu et al., 2015). First, DELTA did not require the input of MODIS images. The requirement of MODIS data input limited the prediction of LSTs before 2000 when MODIS data were not available. Second, the utilization of temporal segmentation (Zhu & Woodcock, 2014b) in DELTA reduced the impact of land cover changes on LST prediction over time. If land cover changes occurred at some points, the corresponding LSTs would be non-stationary leading to the changes of ATC over the time. The assumption to use only one annual frequency to characterize the intra-annual thermal variations was only reasonable to be applied within the stationary (homogenous) segments. Based on *in situ* measurements from the weather stations, results showed that retrieval errors of satellite-derived LSTs ranged from 1.8 to 2.8 K with the average error of 2.3 K. The DELTA algorithm was able to predict daily LSTs at the average bias of 3.5 K. The LST maps of August 15th from 1984 to 2011 showed that the mean LST in both urban and rural regions increased and the UHI intensity in Beijing ranged from 3.3 K to 5.3 K. However, it is suggested that the DELTA algorithm was still influenced by the Landsat revisit frequency and poor atmospheric conditions. A possible way to improve the algorithm is to resort to weather-station LSTs for inferring daily-specific LST variations associated with different weather patterns. In addition, given that DELTA algorithm required intensive computation, parallel distribution of the algorithm among powerful computer clusters is necessary.

## Acknowledgements

The authors are grateful of comments and suggestions provided by anonymous reviewers and the editor, which helped to improve this manuscript. We further thank Beijing Meteorological Administration for provision of surface temperature data from 2008 to 2010. Weng acknowledges a visiting chair professorship awarded to him by South China Normal University during his sabbatical leave.

## 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.
- Bechtel, B., 2012. Robustness of annual cycle parameters to characterize the urban thermal landscapes. *IEEE Geoscience and Remote Sensing Letters* 9, 876–880.
- Bulgin, C.E., Sembhi, H., Ghent, D., Remedios, J.J., Merchant, C.J., 2014. Cloud-clearing techniques over land for land-surface temperature retrieval from the advanced along-track scanning radiometer. *International Journal of Remote Sensing* 35, 3594–3615.
- Carlson, T., 2007. An overview of the “triangle method” for estimating surface evapotranspiration and soil moisture from satellite imagery. *Sensors* 7, 1612–1629.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., Mills, J., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 103, 7–27.
- Dai, A., Wang, J., Ware, R.H., Van Hove, T., 2002. Diurnal variation in water vapor over North America and its implications for sampling errors in radiosonde humidity. *Journal of Geophysical Research-Atmospheres* 107, 11–14.
- Essa, W., van der Kwast, J., Verbeiren, B., Batelaan, O., 2013. Downscaling of thermal images over urban areas using the land surface temperature-impermeable percentage relationship. *International Journal of Applied Earth Observation and Geoinformation* 23, 95–108.
- Friedl, M.A., 2002. Forward and inverse modeling of land surface energy balance using surface temperature measurements. *Remote Sensing of Environment* 79, 344–354.
- Fu, P., Weng, Q., 2015. Temporal dynamics of land surface temperature from Landsat TIR time series images. *IEEE Geoscience and Remote Sensing Letters* 12 (11). <http://dx.doi.org/10.1109/LGRS.2015.2455019>.
- 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, 2207–2218.
- Gong, A., Li, J., Wang, X.-D., Chen, Y.-H., Hu, H.-L., 2006. Study on temporal and spatial distribution characteristics of the urban heat island in Beijing. *Geography and Geo-information Science* 21, 15–18.
- Hilker, T., Wulder, M.A., Coops, N.C., Seitz, N., White, J.C., Gao, F., Masek, J.G., Stenhouse, G., 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, 1988–1999.
- Holzman, M.E., Rivas, R., Piccolo, M.C., 2014. Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index. *International Journal of Applied Earth Observation and Geoinformation* 28, 181–192.
- 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* 5 (10), 1011–1015.
- Imhoff, M.L., Zhang, P., Wolfe, R.E., Bounoua, L., 2010. Remote sensing of the urban heat island effect across biomes in the continental USA. *Remote Sensing of Environment* 114, 504–513.
- Inamdar, A.K., French, A., Hook, S., Vaughan, G., Luckett, W., 2008. Land surface temperature retrieval at high spatial and temporal resolutions over the southwestern United States. *Journal of Geophysical Research-Atmospheres* 113, D07107. <http://dx.doi.org/10.1029/2007JD009048>.
- Ji, C., Liu, W., Xuan, C., 2006. Impact of urban growth on the heat island in Beijing. *Chinese Journal of Geophysics* 49, 69–77.
- Jiménez-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-Atmospheres* 108 (D22).
- Jiménez-Muñoz, J.C., Sobrino, J.A., 2006a. Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. *International Journal of Remote Sensing* 27 (5), 999–1014.
- Jin, M.L., Dickinson, R.E., 2002. New observational evidence for global warming from satellite. *Geophysical Research Letters* 29 (10). <http://dx.doi.org/10.1029/2001GL013833>.
- Jin, M.L., 2000. Interpolation of surface radiative temperature measured from polar orbiting satellites to a diurnal cycle 2. Cloudy-pixel treatment. *Journal of Geophysical Research-Atmospheres* 105, 4061–4076.
- Jin, M.L., Dickinson, R.E., Zhang, D.L., 2005. The footprint of urban areas on global climate as characterized by MODIS. *Journal of Climate* 18, 1551–1565.
- Julien, Y., Sobrino, J.A., 2012. Correcting AVHRR long term data record V3 estimated LST from orbital drift effects. *Remote Sensing of Environment* 123, 207–219.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society* 77, 437–471.
- Jiménez-Muñoz, J.C., Sobrino, J.A., 2006b. Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. *International Journal of Remote Sensing* 27 (5), 999–1014.
- Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2014. Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the northeastern USA. *Remote Sensing of Environment* 150, 132–139.
- 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, 429–440.
- Lin, X., Yu, S., 2005. Interdecadal changes of temperature in the Beijing region and its heat island effect. *Chinese Journal of Geophysics* 48, 39–45.
- 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.
- Liu, J.G., Moore, J.M., 1998. Pixel block intensity modulation: Adding spatial detail to TM band 6 thermal imagery. *International Journal of Remote Sensing* 19, 2477–2491.
- Liu, L., Breitner, S., Pan, X., Franck, U., Leitte, A., Wiedensohler, A., von Klot, S., Wichmann, H.E., Peters, A., Schneider, A., 2011. Associations between air temperature and cardio-respiratory mortality in the urban area of Beijing, China: A time-series analysis. *Environmental Health* 10, 51.
- Lu, L., Venus, V., Skidmore, A., Wang, T., Luo, G., 2011. Estimating land-surface temperature under clouds using MSG/SEVIRI observations. *International Journal of Applied Earth Observation and Geoinformation* 13, 265–276.
- Lu, D., Weng, Q., 2004. Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ imagery. *Photogrammetric Engineering and Remote Sensing* 70 (9), 1053–1062.
- Mallick, K., Jarvis, A.J., Boegh, E., Fisher, J.B., Drewry, D.T., Tu, K.P., Hook, S.J., Hulley, G., Ardö, J., Beringer, J., Arain, A., Niyogi, D., 2014. A surface temperature initiated closure (STIC) for surface energy balance fluxes. *Remote Sensing of Environment* 141, 243–261.
- Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.K., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geoscience and Remote Sensing Letters* 3, 68–72.
- McMillin, L.M., 1975. Estimation of sea surface temperatures from two infrared window measurements with different absorption. *Journal of Geophysical Research* 80, 5113–5117.
- Mu, F.-Y., Zhang, Z.-X., Chi, Y.-B., Liu, B., Zhou, Q.-B., Wang, C., Tan, W., 2007. Dynamic monitoring of built-up area in Beijing during 1973–2005 based on multi-original remote sensed images. *Journal of Remote Sensing* 11, 257.
- 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, 547–556.
- Oke, T.R., Johnson, G.T., Steyn, D.G., Watson, I.D., 1991. Simulation of surface urban heat islands under ‘ideal’ conditions at night part 2: Diagnosis of causation. *Boundary-Layer Meteorology* 56 (4), 339–358.
- Quan, J., Chen, Y., Zhan, W., Wang, J., Voogt, J., Wang, M., 2014. Multi-temporal trajectory of the urban heat island centroid in Beijing, China based on a Gaussian volume model. *Remote Sensing of Environment* 149, 33–46.
- Rasmussen, C.E., Williams, C.K.I., 2006. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.
- Schott, J.R., Hook, S.J., Barsi, J.A., Markham, B.L., Miller, J., Padula, F.P., Raqueno, N.G., 2012. Thermal infrared radiometric calibration of the entire Landsat 4, 5, and 7 archive (1982–2010). *Remote Sensing of Environment* 122, 41–49.
- Shamir, E., Georgakakos, K.P., 2014. MODIS land surface temperature as an index of surface air temperature for operational snowpack estimation. *Remote Sensing of Environment* 152, 83–98.
- Sobrino, J.A., Jimenez-Munoz, J.C., Soria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., Martincz, P., 2008. Land surface emissivity retrieval from different VNIR and TIR sensors. *IEEE Transactions on Geoscience and Remote Sensing* 46, 316–327.
- Song, Y., Zhang, S., 2003. The study on heat island effect in Beijing during last 40 years. *Chinese Journal of Eco-Agriculture* 11 (4), 126–129.
- Stathopoulou, M., Cartalis, C., 2009. Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation. *Remote Sensing of Environment* 113, 2592–2605.
- Streutker, D.R., 2003. Satellite-measured growth of the urban heat island of Houston, Texas. *Remote Sensing of Environment* 85, 282–289.
- Sun, D.L., Pinker, R.T., Kafatos, M., 2006. Diurnal temperature range over the United States: A satellite view. *Geophysical Research Letters* 33 (5). <http://dx.doi.org/10.1029/2005GL024780>.
- Snyder, W.C., Wan, Z., Zhang, Y., Feng, Y.Z., 1998. Classification-based emissivity for land surface temperature measurement from space. *International Journal of Remote Sensing* 19 (14), 2753–2774.
- Thomson, D.J., 1995. The seasons, global temperature, and precession. *Science* 268, 59–68.
- Tomlinson, C.J., Chapman, L., Thorne, J.E., Baker, C.J., 2012. Derivation of Birmingham's summer surface urban heat island from MODIS satellite images. *International Journal of Climatology* 32, 214–224.
- Voogt, J.A., Oke, T.R., 2003. Thermal remote sensing of urban climates. *Remote Sensing of Environment* 86, 370–384.
- Weng, Q., 2014. *Global urban monitoring and assessment through Earth observation*. CRC Press/Taylor and Francis, Boca Raton, FL, p. 440.
- Weng, Q., Fu, P., 2014a. 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 Sensing of Environment* 140, 267–278.
- Weng, Q., Fu, P., 2014b. Modeling diurnal land temperature cycles over Los Angeles using downsampled GOES imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 97, 78–88.
- 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.
- Weng, Q., 2009. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 335–344.
- Weng, Q., Rajasekar, U., Hu, X., 2011. Modeling urban heat islands with multi-temporal ASTER images. *IEEE Transactions on Geoscience and Remote Sensing* 49 (10), 4080–4089.

- Wu, P., Shen, H., Zhang, L., Götsche, F.-M., 2015. Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature. *Remote Sensing of Environment* 156, 169–181.
- Zakšek, K., Oštir, K., 2012. Downscaling land surface temperature for urban heat island diurnal cycle analysis. *Remote Sensing of Environment* 117, 114–124.
- Zhu, W., Lü, A., Jia, S., 2013. Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment* 130, 62–73.
- Zhu, Z., Woodcock, C.E., 2014a. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment* 152, 217–234.
- Zhu, Z., Woodcock, C.E., 2014b. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment* 144, 152–171.