Reconstructing MODIS LST Based on Multitemporal Classification and Robust Regression

Chao Zeng, Huanfeng Shen, *Member, IEEE*, Mingliang Zhong, Liangpei Zhang, *Senior Member, IEEE*, and Penghai Wu

***Abstract*—The Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature (LST) product can offer accurate LST with high temporal and spatial resolution, but the quality is often degraded by cloud. To improve the usability of the MODIS LST, this letter proposes a reconstruction method based on multitemporal data. First, a multitemporal classification is employed to distinguish the different land surface types. The invalid LST values can then be predicted using a robust regression with the multitemporal information from the other LSTs. Finally, postprocessing is proposed to eliminate outliers. Simulated and actual experiments show that the method can accurately reconstruct the missing values.**

***Index Terms*—Classification, Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature (LST), multitemporal, reconstruction, robust regression.**

I. Introduction

LAND surface temperature (LST) is one of the most important parameters in physical processes, at both local and global scales [1], [2]. It combines the results of all surface–atmosphere interactions and energy fluxes between the atmosphere and the ground [3], [4]. LST is therefore widely used in a variety of fields, including hydrology, meteorology, vegetation ecology, environmental monitoring, and so on [5]–[8].

As remote sensing is playing an increasingly important role in Earth science research and environmental problem solving, LST estimated from satellite-based thermal infrared data has become more valued [9]–[14]. The Moderate Resolution Imaging Spectroradiometer (MODIS) sensors aboard the Terra and Aqua satellites are one of the most widely used sources of satellite-derived LST. The method used in MODIS LST retrieval is the generalized split-window algorithm [14]. However, the algorithm can only work well when the data are acquired in sunny weather. When cloud and other atmospheric disturbances obstruct the radiation, the retrieval of LST is greatly affected. As a result, in the MODIS LST products, only the clear-sky pixels have valid values. For the regions with undesirable climate conditions that cause frequent cloud cover, little information can be found in the daily MODIS LST. This greatly limits the subsequent applications in related fields. Research into MODIS LST reconstruction is therefore of great importance.

The main reconstruction methods for MODIS LST are interpolation methods. As only limited spatial information is used, the accuracies of these methods are often not satisfying. Some researchers have attempted to take more related factors into account, such as elevation, and some improved interpolation methods such as cokriging have also been employed [15], [16]. Of all the LST-related factors, elevation is preferred. This is because elevation is a quantifiable parameter and closely related to LST [17], [18]. However, these elevation-based methods have only been proven to be effective in plateaus and mountainous areas, due to the effect of elevation being impaired when the altitude is reduced. It should be noted that the cloudy LST is lower than the cloudless LST for the land surface received solar radiation is hidden by the cloud. Although the above methods did not consider the character of cloud and the restored LST value is only truly under the clear-sky condition, these studies are very meaningful for the application related to the trend of LST [19], [20].

Compared with the elevation-based methods, the land-surface-property-based methods are applicable to wider conditions. Although the physical properties of land surfaces are complex and difficult to quantify, they can be qualitatively separated by classification algorithms. The corresponding classification map of the degraded LST image can then provide auxiliary information for the LST restoration. Based on this idea, a new LST reconstruction method using multitemporal classification and robust regression is proposed in this letter.

II. Methods

For convenience, in this letter, the defective MODIS LST to be reconstructed is defined as the target LST, whereas the other auxiliary LST is referred to as the fill LST. The fill LST must be close in time to the target LST and has complementary information. The reconstruction is based on the hypothesis that the surface temperature of a ground feature linearly changes, and similar changes occur on neighboring

Manuscript received April 21, 2014; revised July 1, 2014; accepted July 29, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 41271376, the National High Technology Research and Development Program (863) under Grant 2013AA12A301, the Wuhan Science and Technology Program (2013072304010825), the Program for Changjiang Scholars and Innovative Research Team in University (IRT1278), and the Fundamental Research Funds for the Central Universities under Grant 2012619020210.

C. Zeng and L. Zhang are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: zengchaozc@hotmail.com; zlp62@public.wh.hb.cn).

H. Shen and M. Zhong are with the Key Laboratory of Geographic Information System of Ministry of Education, School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn; zml890414@163.com).

P. Wu is with the School of Resources and Environmental Engineering, Anhui University, Hefei 230601, China (e-mail: dbqwph2007@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2014.2348651

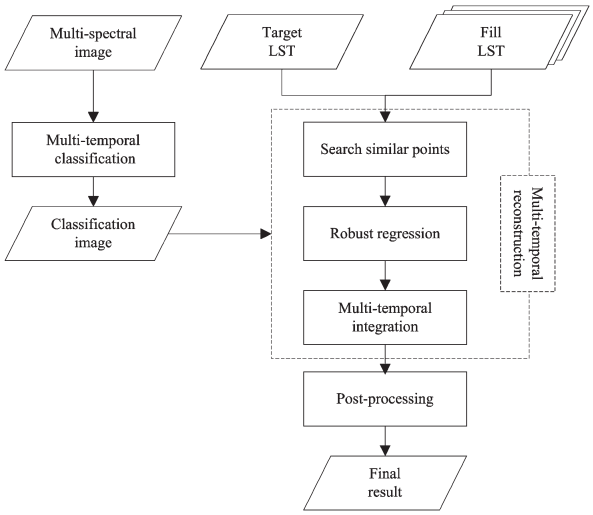

Fig. 1. Flowchart of the proposed method.

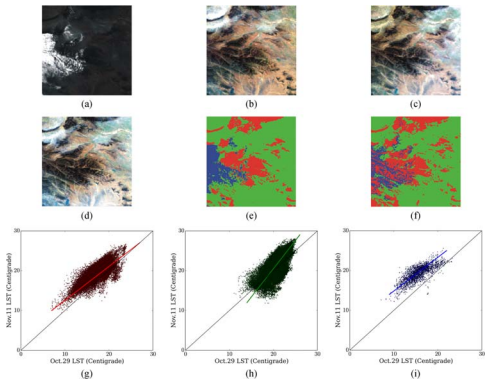

Fig. 2. Subimages of Terra MODIS of 1000-m-pixel resolution from over Henan, China, acquired on (a) October 29, 2001; (b) November 11, 2001; (c) November 14, 2001; and (d) November 18, 2001, respectively, and the classification image by (e) the ordinary method and (f) the improved K-means method. Panels (g), (h), and (i) are the scatterplots between the October 29 and November 11 LST corresponding to the three classes in image (f).

similar features. Therefore, the missing surface temperature value can be reconstructed by the value acquired at another time and the linear relationship calculated from the nearest similar pixels. To separate the various relationship types, an unsupervised multitemporal classification is first implemented. For each invalid pixel, the nearest similar points are searched according to the classification map. A local robust regression is then performed on the similar points to recover the target invalid pixel. In some cases, more than one LST is needed to fill all the invalid pixels; hence, there may be several recovered values. In this situation, a multitemporal integration method is employed to get the synthesized result. Finally, a postprocessing procedure is applied to eliminate the outliers. Fig. 1 presents a flowchart of the proposed reconstruction method. The details will be discussed below.

A. Multitemporal Classification

A classification process is employed to distinguish the features with different linear relationships between the LST images. Since most of the land surface information is missing in the LST product, MODIS Level-1B reflectance data are used in the classification process. However, as multitemporal LSTs are involved in the reconstruction, the land surface may change during the time interval. Furthermore, in most cases, it is difficult to obtain a cloud-free image for the land surface classification. To solve this problem, a multitemporal classification method is applied. The land bands (bands 1–7) are obtained from each MODIS Level-1B reflectance product, resampled to 1-km resolution, and composed as one. The corresponding temporal LSTs are used as mask images, and the contaminated pixels are removed. An improved K-means algorithm is then carried out. First, the initial cluster centers are randomly assigned from the locations with valid values in all the bands. The distances between each pixel and the cluster centers are then calculated. The distance between the kth pixel and the ith cluster center can be expressed as

$$d_{ki} = \sqrt{\frac{1}{n_k}\sum_{j=1}^{n_k}(x_{kj} - c_{ij})^2} \tag{1}$$

where x_{kj} and c_{ij} denote the value of the kth pixel and the ith cluster center in the jth band, and n_k denotes the total number of valid pixels at this location. It should be noted that as the locations of the contaminated pixels are different in the multitemporal images, the value of n_k varies for the different locations. That is the main difference from the original K-means algorithm. Subsequently, the pixels are all assigned to the cluster center they are closest to. The mean values of all the pixels belonging to each type are then calculated and used instead of the former cluster centers. This is repeated until the cluster centers no longer change, and the final classification result can be obtained.

To validate the proposed method, a comparison between the ordinary and improved K-means algorithm was made, and the results are shown in Fig. 2. The study region is a 300 × 300 km area in Henan, China. Fig. 2(a)–(d) shows four MODIS Level-1B images (true color composite R = band 1, G = band 4, B = band 3) acquired on October 29, 2001; November 11, 2001; November 14, 2001; and November 18, 2001, respectively. The two classification methods were tested on the composed data of the four images, and the number of classes was set as three. Fig. 2(e) is the classification result of ordinary K-means, and Fig. 2(f) is the result of the improved K-means. The result of the ordinary K-means is inaccurate as one of the four MODIS Level-1B images is badly contaminated by cloud. However, the improved K-means can overcome this disadvantage and gives a more accurate classification map. As shown in Fig. 2(g)–(i), the linear relationships of the three classes between the two LST images are obvious. It indicates that the multitemporal classification method is effective, and the missing LST value can be reconstructed by the linear relationship between LST images.

B. Multitemporal Reconstruction

With the classification results as a reference, a multitemporal reconstruct process can be implemented to fill the defective LST image. For each missing pixel, its similar pixels are defined as the nearest valid pixel within the same class. To search for the similar pixels, an adaptive window is used on both target and fill LST. To guarantee the robustness of the reconstruction, a

threshold is set for the number of concerned similar pixels as K. The window size is first set at an initial value, and it is enlarged by two pixels if threshold K cannot be met. Considering that too large a window size will take a lot of computation time, a maximum window size can be set. If threshold K still cannot be met when the maximum window size is reached, it means that this reference LST is not suitable for this position; hence, the missing value will be filled using another fill LST acquired in a different day. After the searching procedure, two groups of values from the target LST and the fill LST can be obtained. Let D_T and D_F denote the searched values from the target LST and fill LST, then the linear relationship between them can be expressed as

$$D_T = a \times D_F + b \tag{2}$$

where a and b are the linear regression coefficients.

To avoid the effects of abnormal values, the linear relationship is calculated by a robust regression using iteratively reweighted least squares [21]. Equation (2) can be rewritten in matrix notation as

$$\mathbf{P} = \mathbf{xQ} \tag{3}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are matrices of the values from the target LST and fill LST, and $\mathbf{x}$ is the vector containing the linear regression coefficients. The initial coefficients are calculated using the generalized least squares, i.e.,

$$\mathbf{x}_1 = (\mathbf{Q^T Q})^{-1}\mathbf{Q^T P}. \tag{4}$$

The coefficients are iteratively updated until the corresponding fitting error falls in a desired range. The ith fitting error can be expressed as

$$\mathbf{e}_i = \mathbf{x}_i\mathbf{Q} - \mathbf{P}. \tag{5}$$

The weights of the next iteration $\mathbf{w}_{i+1}$ are then calculated by a Huber function, i.e.,

$$\mathbf{W}_{i+1}(j) = \begin{cases} 1, & |\mathbf{e}_i(j)| \le h \\ h/|\mathbf{e}_i(j)|, & |\mathbf{e}_i(j)| > h \end{cases} \tag{6}$$

where $\mathbf{e}_i(j)$ and $\mathbf{W}_{i+1}(j)$ are the jth values in vector $\mathbf{e}_i$ and $\mathbf{W}_{i+1}$, and h is the Huber parameter. In this letter, the medium value of vector $\mathbf{e}_i$ is used as an adaptive Huber parameter.

The updated coefficients can be obtained by

$$\mathbf{x}_{i+1} = (\mathbf{Q^T W}_{i+1}\mathbf{Q})^{-1}\mathbf{Q^T W}_{i+1}\mathbf{P}. \tag{7}$$

The final coefficients are obtained when the iteration is terminated, and the fill value can be calculated by

$$V_T = a \times V_F + b \tag{8}$$

where V_T is the filled value in the target LST, V_F is the reference value in the fill LST, and a and b are the coefficients.

Threshold K is an important parameter in the reconstruction procedure. With a smaller threshold, the size of the searching window will also be smaller. Generally, as LSTs are highly spatially correlated, the linear relationship will fit better if the similar points are concentrated in a small window. However, if the threshold is not large enough, a few abnormal points can greatly disturb the linear relationship. Therefore, the threshold should be appropriately adjusted to ensure accuracy. According to our experience, in most cases, good results can be obtained when K is set to between 20 and 30.

Under ideal conditions, the reconstruction of the target LST could be completed by only one fill LST. However, due to various factors, it is often difficult to find such a high-quality fill LST. Usually, more than one fill LST is needed to completely fill all the missing values. As a result, there may be several predicted values for a single pixel. Considering the random error, the final value is calculated using a weighted integration. The weight r_i can be calculated by

$$r_i = 1 \Big/ \left(d_i \times \sum_{k=1}^{n} (1/d_k) \right) \tag{9}$$

where n is the total number of all the reconstructed values, and d_i denotes the distance between the ith filled value and the ideal filled value. However, the ideal value cannot be obtained; hence, the calculation of d_i is instead by

$$d_i = V_i - V_0 \tag{10}$$

where V_i denotes the ith reconstructed value, and V_0 denotes a roughly estimated value of the invalid pixel. In this letter, the mean value of all the valid pixels in a 3×3 region is used.

Finally, the final filled value V can be expressed as

$$V = \sum_{k=1}^{n} r_i \times V_i. \tag{11}$$

C. Postprocessing

Almost all the missing values will have been restored after following the above steps, except for some abnormal pixels. As shown in Fig. 3(a), some outliers can be found on the upper-left corner. Here, it can be clearly seen that several pixels are much brighter than the surrounding pixels. This is because the linear relationship is unrealistic when most of the similar pixels are far away from the target pixel. To deal with this problem, an image-based histogram analysis is applied as an outlier detector. The quartiles of the image DN value are calculated, and then, two boundaries are set. The equations for the boundaries are given as

$$B_L = Q_1 - 1.5 \times (Q_3 - Q_1) \tag{12}$$

$$B_H = Q_3 + 1.5 \times (Q_3 - Q_1) \tag{13}$$

where B_L and B_H denote the upper and lower boundaries, and Q_1 and Q_3 are the first and third quartiles.

A pixel with a value out of the boundaries will be treated as an outlier. As the range of values on the target LST is quite large, the original LST is divided into 100×100 pixel blocks, and postprocessing is implemented in each block, respectively. A 3×3 pixel mean value filter is implemented to remove the outliers.

By following the aforementioned steps, the invalid pixels can be reconstructed one by one. The processing is terminated when all the invalid pixels are recovered.

III. Experiments

Here, the performance of the proposed method is tested using Terra MODIS surface temperature data (MOD11A1) and Terra MODIS Level-1B reflectance data (MOD02). The data were obtained from the LADDS website

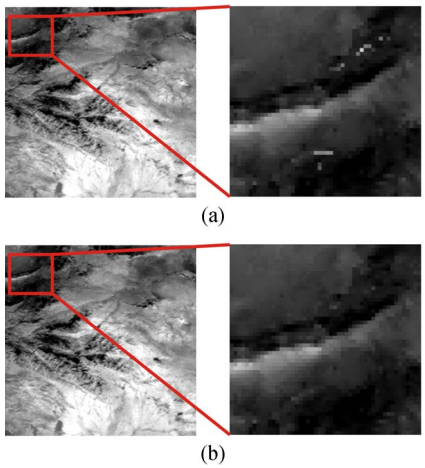

Fig. 3. Reconstruction results (a) before and (b) after postprocessing.

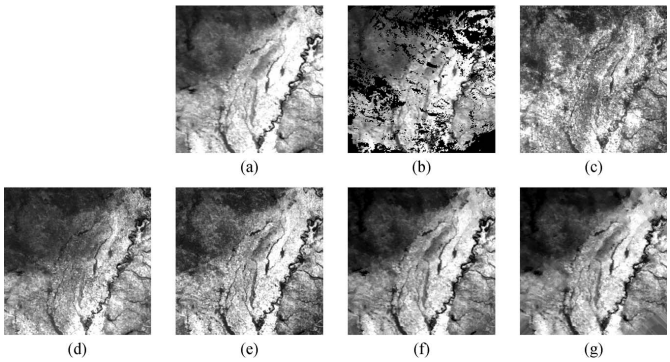

Fig. 4. Tested (a) Terra MODIS LST acquired on October 10, 2010; (b) the simulated defective LST image; and (c)–(e) the fill LSTs acquired on September 4, 2010; September 20, 2010; and October 6, 2010, respectively. Images (f) and (g) are the reconstructed results by MNKI and the proposed method, respectively.

TABLE I
QUANTITATIVE EVALUATION RESULTS OF THE EXPERIMENT SHOWN IN FIG. 4

Method	MNKI	Proposed method
MSE	2.176	0.692
r	0.926	0.977

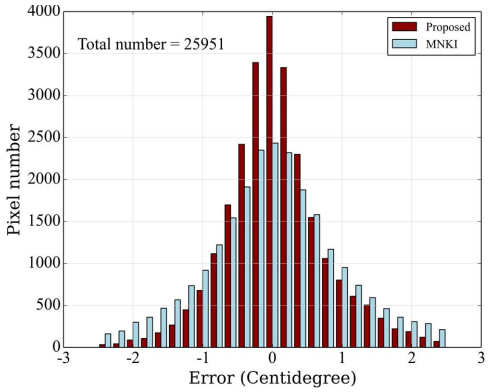

Fig. 5. Errors of reconstructed LST in Fig. 4(g) and (h).

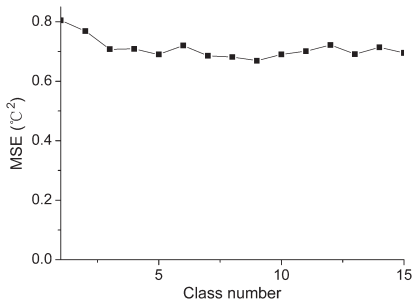

Fig. 6. MSEs when using different class numbers, as tested in Fig. 4(c)–(f).

http://ladsweb.nascom.nasa.gov/. Both the MODIS LST and the MODIS Level-1B data were preprocessed geometrically and radiometrically. The MODIS LST data were reprojected to Universal Transverse Mercator (UTM) mode using the MODIS reprojection tool, and in order to fit with MODIS LST data, the MODIS Level-1B data were also reprojected to UTM mode using ENVI. All data used in the letter were cropped subimages of 300×300 pixels, and the LSTs were transformed to degrees Celsius (coefficient $= 0.02$, deviation value $= -273.15$).

Experiments were performed on both simulated and real MODIS LSTs. To verify the advantage of the proposed method, the moving neighborhood kriging interpolation (MNKI) method was implemented as a comparison [22]. In the multitemporal-data-based experiments, four MODIS LST images were used, i.e., one as a target LST and the other three as fill LSTs. The corresponding MODIS Level-1B data were used in the multitemporal classification process.

Fig. 4(a) shows the tested Terra MODIS LST, acquired on October 10, 2010, over Tennessee, Missouri, and Arkansas, in the United States, and Fig. 4(b) is the simulated defective LST image. Fig. 4(c)–(e) shows the fill LSTs acquired on September 4, 2010; September 20, 2010; and October 6, 2010, respectively. Fig. 4(f) and (g) shows the reconstructed result by MNKI and the proposed method (the class number is set as 5; K is set as 30), respectively. As the most defective pixels are scattered over the experiment image, MNKI and the proposed method can get good results in these regions. However, on the lower-right corner, where there is a rather wide gap, obvious artifacts can be found in Fig. 4(f), whereas Fig. 4(g) shows a great visual result and is consistent with the original LST. The experimental results show that the multitemporal-data-based method can deal with the large defective region much better than the other methods.

In order to quantitatively compare the results of the above methods, the Pearson correlation coefficient (r) and the mean square error (MSE) of all the filled pixels was calculated, and the results are shown in Table I. Compared with the MNKI method, the proposed method shows much higher accuracy in the indexes. Moreover, the errors of the reconstructed LSTs are compared in Fig. 5; the proposed method also shows a significant advantage.

As previously mentioned, the most important parameter that needs to be determined is the class number. The determination of this parameter was tested through a series of comparative experiments, using the images shown in Fig. 4(c)–(e). The MSE values of the reconstructed pixels were calculated using different class numbers from 1 (no classification) to 15, and the results are shown in Fig. 6. The quantitative assessment shows better results when the number of classes is set to more than 3, and the errors are rather obvious when the class number

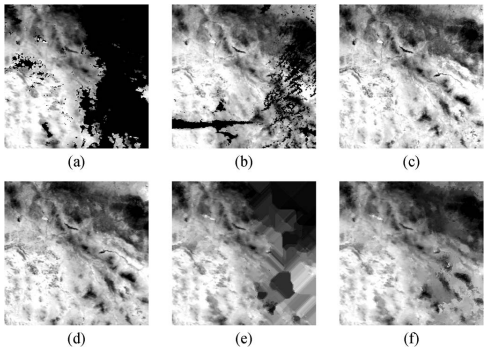

Fig. 7. (a) Target defective LST image acquired on June 23, 2005. (b)–(d) Fill LSTs acquired on May 30, 2005; June 6, 2005; and June 18, 2005, respectively. (e) and (f) Reconstructed result by MNKI and the proposed method, respectively.

is set too small. This is because a too small value of class number cannot distinguish the different land surface types. It is noteworthy that, as the scene types in Fig. 4(a) are rather complex, a greater value of class number is more likely to result in an accurate result. However, considering that more classes will increase the computation time significantly, we suggest that the class number is set at less than 10 in most cases.

To further verify the efficacy of the proposed method, an experiment using actual defective LSTs acquired over Arizona, USA, was implemented. The target LST, as shown in Fig. 7(a), was acquired on June 23, 2005. Fig. 7(b)–(d) shows the fill LSTs acquired on May 30, June 6, and June 18, respectively. The experimental results reconstructed by MNKI and the proposed method (the class number is set as 4; K is set as 30) are shown in Fig. 7(e) and (f), respectively. As more than one third of the values are missing in the target LST, the spatial information is mostly lost in the defective region. The MNKI method gives a rather poor result in this experiment. With the complementary information from the other LSTs, the proposed method provides much better visual connectivity.

IV. Conclusion

This letter has presented a reconstruction method for MODIS LST using multitemporal classification and rough regression. Compared with the existing methods, the proposed method makes use of multitemporal data and the differences in the physical properties of the varied land surfaces. The experimental results demonstrate that the proposed method is effective, and the qualitative classification information is helpful for the LST reconstruction. The subsequent LST applications such as the analysis of urban heat island may benefit from this study.

However, there are still some issues that need to be noted. More geographic factors such as elevation could be involved in further research for high-altitude areas. Furthermore, the number of classes is empirically determined in this letter. Theoretically, the land surface types could be distinguished more clearly when adding to the number of classes, and the linear relationship would be more accurate. However, this also makes it more likely that not enough similar points could be found within an appropriate range, and then, the linear relationship may be inaccurate as the spatial correlation of similar points is weakened. A more rational method will be explored in the future.

References

[1] M. Anderson *et al.*, "A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales," *Remote Sens. Environ.*, vol. 112, no. 12, pp. 4227–4241, Dec. 2008.

[2] W. Kustas and M. Anderson, "Advances in thermal infrared remote sensing for land surface modeling," *Agric. Forest Meteorol.*, vol. 149, no. 12, pp. 2071–2081, Dec. 2009.

[3] H. Mannstein, "Surface energy budget, surface temperature and thermal inertia," in *Remote Sensing Applications in Meteorology and Climatology*. Dodrecht, The Netherlands: Springer-Verlag, 1987, pp. 391–410.

[4] P. Sellers, F. Hall, G. Asrar, D. Strebel, and R. Murphy, "The first ISLSCP Field Experiment (FIFE)," *Bull. Amer. Meteorol. Soc.*, vol. 69, no. 1, pp. 22–27, Jan. 1988.

[5] A. J. Arnfield, "Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island," *Int. J. Climatol.*, vol. 23, no. 1, pp. 1–26, Jan. 2003.

[6] J. Hansen, R. Ruedy, M. Sato, and K. Lo, "Global surface temperature change," *Rev. Geophys.*, vol. 48, no. 4, pp. RG4004-1–RG4004-29, Dec. 2010.

[7] J. A. Voogt and T. R. Oke, "Thermal remote sensing of urban climates," *Remote Sens. Environ.*, vol. 86, no. 3, pp. 370–384, Sep. 2003.

[8] Q. Weng, "Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 4, pp. 335–344, Jul. 2009.

[9] B. Liu, L. Zhang, X. Zhang, B. Zhang, and Q. Tong, "Simulation of EO-1 hyperion data from ALI multispectral data based on the spectral reconstruction approach," *Sensors*, vol. 9, no. 4, pp. 3090–3108, Apr. 2009.

[10] Q. Tong, Y. Xue, and L. Zhang, "Progress in hyperspectral remote sensing science and technology in China over the past three decades," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 70–91, Jan. 2014.

[11] Z. Li *et al.*, "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013.

[12] J. C. Jimenez-Muñoz and J.A. Sobrino, "A generalized single-channel method for retrieving land surface temperature from remote sensing data," *J. Geophys. Res., Atmos.*, vol. 108, no. D22, pp. 4688–4696, Nov. 2003.

[13] Z.-h. Qin, A. Karnieli, and P. Berliner, "A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel–Egypt border region," *Int. J. Remote Sens.*, vol. 22, no. 18, pp. 3719–3746, Jan. 2001.

[14] Z. Wan and J. Dozier, "A generalized split-window algorithm for retrieving land-surface temperature from space," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 892–905, Jul. 1996.

[15] D. Cai, N. Guo, and C. Li, "Interpolation of air temperature based on DEM over Eastern Region of Gansu," *J. Arid Meteorol.*, vol. 1, no. 6, pp. 10–17, Jan. 2009.

[16] X. Li, G. Cheng, and L. Lu, "Comparison study of spatial interpolation methods of air temperature over Qinghai–Xizang Plateau," *Plateau Meteorol.*, vol. 22, no. 6, pp. 565–573, 2003.

[17] L. Ke, Z. Wang, C. Song, and Z. Lu, "Reconstruction of MODIS LST time series and comparison with land surface Temperature (T) among observation stations in the Northeast Qinghai–Tibet Plateau," *Progr. Geogr.*, vol. 30, no. 7, pp. 819–826, Jul. 2011.

[18] M. Neteler, "Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data," *Remote Sens.*, vol. 2, no. 1, pp. 333–351, Jan. 2010.

[19] D. R. Streutker, "A remote sensing study of the urban heat island of Houston, Texas," *Int. J. Remote Sens.*, vol. 23, no. 13, pp. 2595–2608, Jan. 2002.

[20] L. Zhang *et al.*, "Spatio-temporal reconstruction of air temperature maps and their application to estimate rice growing season heat accumulation using multi-temporal MODIS data," *J. Zhejiang Univ. Sci. B, Biomed. Biotechnol.*, vol. 14, no. 2, pp. 144–161, Feb. 2013.

[21] D. P. O'Leary, "Robust regression computation using iteratively reweighted least squares," *SIAM J. Matrix Anal. Appl.*, vol. 11, no. 3, pp. 466–480, Jul. 1990.

[22] T. C. Haas, "Kriging and automated variogram modeling within a moving window," *Atmos. Environ. A, Gen. Topics*, vol. 24, no. 7, pp. 1759–1769, 1990.