Journal Name
Environmental Science: Processes & Impacts
Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Generating daily high spatial land surface temperatures by combining Aster and MODIS land surface temperature products for environmental process monitoring

Mingquan Wu[a,*], Hua Li[a], Wenjiang Huang[b], Zheng Niu[a], Changyao Wang[a]

Received (in XXX, XXX) XthXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

Abstract: There is a shortage of daily high spatial land surface temperature (LST) data for use in high spatial and temporal resolution environmental process monitoring. To address this shortage, this work used the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM), Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM), and the Spatial and Temporal Data Fusion Approach (STDFA) to estimate high spatial and temporal resolution LST by combining Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) LST and Moderate Resolution Imaging Spectroradiometer (MODIS) LST products. The actual ASTER LST products were used to evaluate the precision of the combined LST imageries using the correlation analysis method. This method was tested and validated in study areas located in Gansu Province, China. The results show that all the models can generate daily synthetic LST imagery with a high correlation coefficient (r) of 0.92 between the synthetic imagery and the actual ASTER LST observations. The ESTARFM has the best performance, followed by the STDFA and the STARFM. Those models had better performance in desert areas than in cropland. The STDFA had better noise immunity than the other two models.

Keywords: MODIS; ASTER; Land surface temperature; Multi-sensor fusion;

Environmental impact

The land surface temperature (LST) product is very important for understanding the impact of global climate change and human activities on climate change. However, there is a shortage of LST data for high spatial and temporal environmental process monitoring. To solve this problem, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM), Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM), and the Spatial and Temporal Data Fusion Approach (STDFA) were tested and validated to estimate high spatial and temporal resolution land surface temperature data by combining Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) LST and Moderate Resolution Imaging Spectroradiometer (MODIS) LST products in study areas located in Gansu Province, China. The results show that all the models can accurately generate daily synthetic LST imagery, with a correlation coefficient higher than 0.92. The generated daily synthetic LST images enable high spatial and temporal resolution environmental process monitoring of environmental pollution, the ecological environment, and agriculture.

1. Introduction

Land surface temperature (LST) data on the earth's surface is used for the analysis and simulation of important surface energy balance parameters,[1] and is also an important input parameter in soil moisture, drought, and crop yield estimation models.[2-3] LST is also widely used in regional hydrology ecological and environmental research.[4-5] Because the response of different vegetation types or ecosystems to temperature is an important part of global change research, LST plays a large role in monitoring global climates[6-8] and studying urban heat island effects.[9-11] Thus, the LST product is very important to understanding the impact of global climate change and human activities on climate change.[12]

The main method of estimating LST is an inversion approach using thermal infrared remote sensing images.[13] Currently, global LST products can be estimated from Advanced Spaceborne

Thermal Emission and Reflection Radiometer (ASTER) and Moderate Resolution Imaging Spectroradiometer (MODIS) data.[14-15] The thermal infrared data from Landsat and the Huan Jing Constellation satellite (HJ) can also be used to estimate LST.[16-17] However, due to limitations in satellite data acquisition and the influence of clouds, there is a lack of high spatial and temporal resolution LST data. MODIS LST products (MOD11A1) have a high temporal resolution and have been widely used in global or large-scale area monitoring.[18] However, these products are not suitable for high precision monitoring in small areas due to their lower spatial resolution. The LST products of ASTER, Landsat, and HJ have a high spatial resolution and have been widely used in small study areas, for example, to examine urban heat islands.[19] However, these products are not suitable for time series monitoring due to the long review cycle of the satellites. Therefore, there is a pressing need to combine those two types of data to achieve both high spatial and high temporal resolutions.

Recently, several data fusion approaches have been proposed to blend high spatial resolution data and high temporal resolution data. These approaches generate synthetic high spatial resolution imagery with a high temporal resolution. Gao *et al.*[20] introduced the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) for blending MODIS and Landsat imagery. Several studies applied this model and demonstrated the use of STARFM within a mainly coniferous area, as well as for urban environmental variables extraction, vegetated dry land ecosystem monitoring, public health studies, and the generation of daily land surface temperatures.[21-25] Zhu *et al.*[26] enhanced the STARFM for complex heterogeneous regions. Other researchers proposed methods that fuse Landsat and MODIS data to generate high temporal resolution synthetic Landsat data based on a linear mixed model.[27-29] Wu *et al.*[30] introduced a Spatial and Temporal Data Fusion Approach (STDFA) based on linear mixing theory and applied it to the estimation of a high-resolution Leaf Area Index[31] and crop mapping.[32] In these cases, the methods were originally proposed fusing Landsat and MODIS reflectance images. Thus, there is a need to test the ability of these methods with other sensors and products.

The present work seeks to apply these methods in the fusion of ASTER and MODIS LST products. The objectives of this study are to (1) compare the suitability of applying STARFM, ESTARFM, and STDFA in the fusion ASTER and MODIS LST products; and (2) evaluate the applicability of these three models for different land use types.

2. Study area and data preparations

2.1. Study area

Zhangye country, Gansu province, China was selected as the study area for this research. The latitude and longitude range for the region are from 38° 02′ 32.23″ to 38° 09′ 24.17″ N, and from 100° 00′ 29.63″ to 100° 02′ 45.89″ E, respectively. Zhangye has a continental dry climate with an annual average temperature of 6 ℃. The coldest month is January, while the hottest month is July. The main land forms of Zhangye are plains, cities, and deserts. The plains in the region are mainly devoted to agriculture.

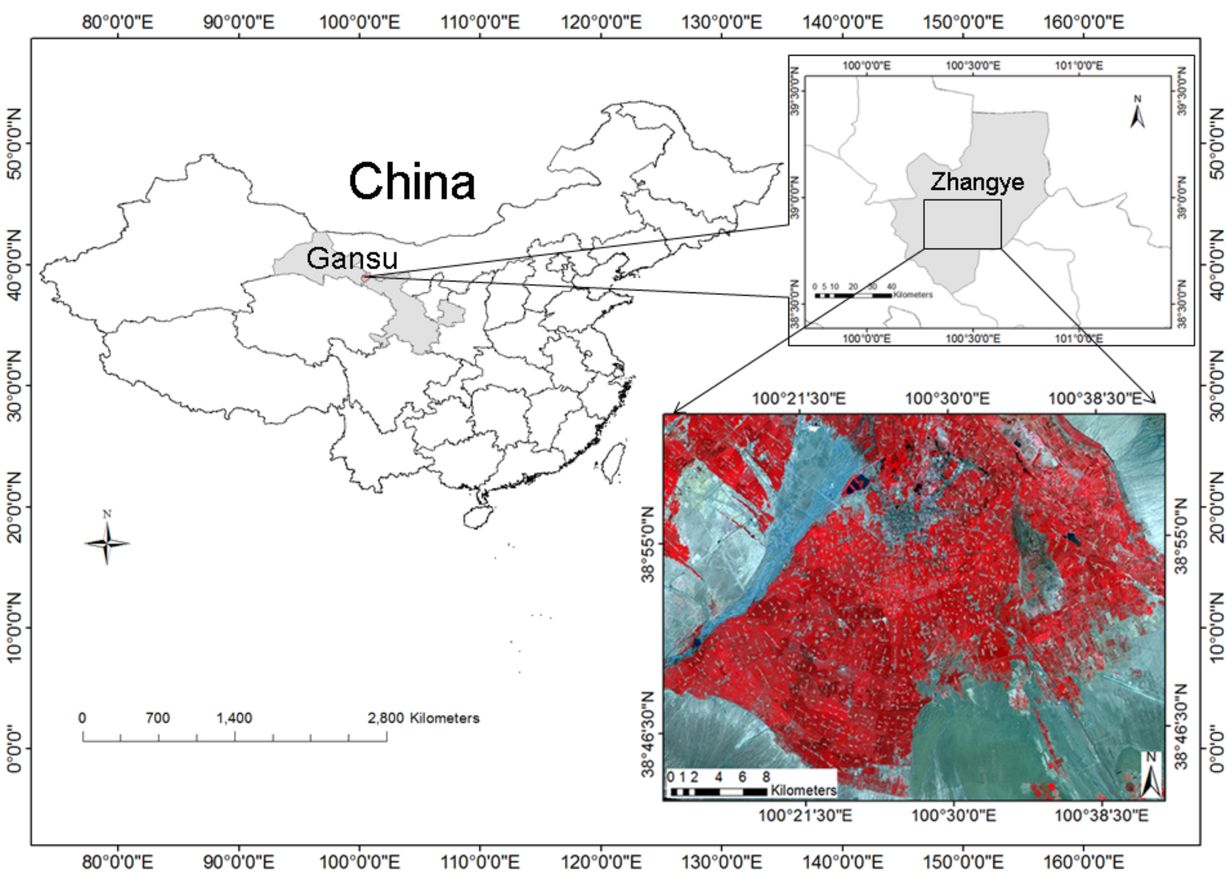

Fig. 1 Location of the study site.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

2.2. Data preparation

Eleven ASTER LST data sets and twenty-nine MODIS LST datasets (MOD11A1) were used in this study (Table 1).

Table 1 Main characteristics of ASTER and MODIS images used in the research

ASTER	Usage	MOD11A1	ASTER	Usage	MOD11A1	ASTER	Usage	MOD11A1
06/15/2012	Reference	06/15/2012			07/31/2012			08/28/2012
		06/18/2012	08/02/2012	Reference and Validation	08/02/2012			08/29/2012
		06/19/2012			08/04/2012	09/03/2012	Reference and Validation	09/03/2012
		06/20/2012	08/11/2012	Reference and Validation	08/11/2012	09/12/2012	Reference and Validation	09/12/2012
		06/21/2012			08/14/2012			09/15/2012
06/24/2012	Reference and Validation	06/24/2012			08/15/2012			09/17/2012
		07/04/2012	08/18/2012	Reference and Validation	08/18/2012	09/19/2012	Reference and Validation	09/19/2012
07/10/2012	Reference and Validation	07/10/2012			08/21/2012			09/21/2012
		07/14/2012			08/22/2012			09/26/2012
		07/19/2012	08/27/2012	Reference and Validation		09/28/2012	Reference	09/28/2012

2.2.1. ASTER data

ASTER is one of five remote sensory devices on board the Terra satellite. ASTER provides earth images every 16days in 14 different bands in a spectral range from visible to thermal infrared light. Its spatial resolution ranges between 15 to 90 meters. ASTER data are widely used to create detailed maps of land surface temperatures, emissivity, reflectance, and elevation.

Eleven ASTER LST datasets, provided by the Cold and Arid Regions Science Data Center at Lanzhou, were used in this study (Table 1). The data were acquired under clear sky conditions and were estimated from ASTER L1B thermal infrared data using the temperature emissivity separation (TES) inversion algorithm. The thermal infrared data were atmospherically corrected using the water vapour scaling (WVS) method in MODTRAN with MODIS atmosphere profile data products (MOD07) acquired on the same day as the ASTER data. The ASTER L1B thermal infrared data were then georeferenced using a second-order polynomial warping approach based on the selection of an appropriate number of ground control points (GCPs). This procedure used ASTER L3 data with the nearest neighbour resampling method and a position error within 0.7 ASTER pixels. The accuracy of this data was evaluated using the ground measured LST data. The results showed that the average deviation of this product was less than 0.5K with an RMSE less than 2K.[33-34]

2.2.2. MODIS LST data

MODIS is a key instrument aboard the Terra and Aqua satellites. MODIS provides earth images every 1 to 2 days in 36 spectral bands, ranging from visible to thermal infrared light. Its spatial resolution ranges from 250 to 1000 meters.

Twenty-nine MODIS LST data sets (MOD11A1, 1 km, collection5) obtained under clear sky conditions were used in this study (Table 1). These MODIS images were re-projected from the native sinusoidal projection to a UTM-WGS84 reference system, and were resized to the selected study areas using the MODIS reprojection tool (MRT). The MODIS data were then georeferenced by a second-order polynomial warping approach based on the selection of an appropriate number of GCPs on 1000m ASTER images, using a nearest neighbour resampling method. The 1000m ASTER images were generated from georeferenced ASTER images with the pixel aggregate resampling method, and the position error was within 0.78 ASTER pixels. The MODIS LST data included some default data. Median filtering methods were used to remove these default values.

3. Approach

3.1 Model introduction

3.1.1 STARFM

In STARFM model, the relationship between the fine resolution image and the coarse resolution image is described as:

$$r(x, y, t_k) = R(x, y, t_k) + \xi(t_k) \qquad (1)$$

where $r(x, y, t_k)$ is the fine resolution reflectance of target pixels (x,y) at time t_k; $R(x, y, t_k)$ is the coarse resolution reflectance; $\xi(t_k)$ is the sensor difference. Suppose the ground coverage type and system errors does not change over prediction date, the STARFM model predicts synthetic high spatial imageries from low spatial imageries as follows: [20]

$$r(x,y,t_k)=R(x,y,t_k)+r(x,y,t_0)-R(x,y,t_0) \quad (2)$$

By introducing additional information from the neighbouring pixels to reduce the influences of land cover change, surface heterogeneity, and solar geometry bidirectional reflectance distribution function (BRDF) changes, a weighted STARFM model can be determined as follows:

$$r(\mathrm{x},\mathrm{y},t_k)=\sum_{i=1}^{w}\sum_{j=1}^{w}\sum_{k=1}^{n}W_{ijk}\left(R(\mathrm{x_i},\mathrm{y_j},\mathrm{t_k})-R(\mathrm{x_i},\mathrm{y_j},t_0)+r(\mathrm{x_i},\mathrm{y_j},\mathrm{t_0})\right) \quad (3)$$

Where k is the number of pixels (x_i,y_j) in the window w; w is the search window size; W_{ijk} is the weight determined by the spectral difference S_{ijk} and temporal difference T_{ijk} between the fine and low resolution data, and D_{ijk} is the distance between the target pixel and the candidate pixel. Those parameters are calculated as follows:

$$W_{ijk}=S_{ijk}\times T_{ijk}\times d_{ijk} \quad (4)$$

$$S_{ijk}=\left|R(\mathrm{x_i},\mathrm{y_j},t_k)-r(\mathrm{x_i},\mathrm{y_j},\mathrm{t_k})\right| \quad (5)$$

$$T_{ijk}=\left|R(\mathrm{x_i},\mathrm{y_j},t_k)-R(\mathrm{x_i},\mathrm{y_j},\mathrm{t_0})\right| \quad (6)$$

$$D_{ijk}=1+d_{ijk}/A \quad (7)$$

$$d_{ijk}=\sqrt{(x_i-x)^2+(y_i-y)^2} \quad (8)$$

where A is the constant used to determine whether the spectral similarity weight or the distance weight is more important.

3.1.2 ESTARFM

The ESTARFM is based on the assumption that the change of reflectance of each endmember is linear over time. Thus the ratio v_k of the change of reflectance for kth endmember to the change of reflectance for a coarse pixel can be described as:

$$(r(t_k)-r(t_0))\big/(R(t_k)-R(t_0))=v_k \quad (9)$$

Similar to the STARFM, a weighted ESTARFM model considers information from neighbouring pixels as follows:[26]

$$r(x,y,t_k)=r(x_i,y_j,t_0)+\sum_{i=1}^{w}\sum_{j=1}^{w}\sum_{k=1}^{n}W_{ijk}\times v_{ijk}\left(R(x_i,y_j,t_k)-R(x_i,y_j,t_0)\right) \quad (10)$$

The weight W_{ijk} of ESTARFM is calculated as follows:

$$W_i=(1/D_i)\Big/\sum_{i=1}^{n}(1/D_i) \quad (11)$$

$$D_i=(1+S_i)\times d_i \quad (12)$$

$$d_i=1+\sqrt{(x-x_i)^2+(y-y_i)^2}/(w/2) \quad (13)$$

$$S_i=\frac{E\left[(r_i-E(r_i))\times(R_i-E(R_i))\right]}{\sqrt{D(r_i)}\times\sqrt{D(R_i)}} \quad (14)$$

3.1.3 STDFA

The STDFA is based on the linear mixing theory. According to linear mixing theory, the reflectance of a coarse-resolution spatial pixel measured by a sensor is its composite value. The response of each coarse spatial resolution pixel is assumed to be a linear combination of the responses of each land cover class contributing to the mixture which was expressed as:[35]

$$R(x,y,t)=\sum_{c=0}^{k}f_c(x,y,c)\times\bar{r}(c,t)+\xi(x,y,t) \quad (15)$$

Constrained: $\sum_{c=0}^{k}f_c(x,y,c)<1$; and $f_c(x,y,c)\geq 0$ for all where $f_c(x,y,c)$ is the fractional cover of class c in coarse pixel (x,y) at time t, which is usually assumed not to change over time, $\bar{r}(c,t)$ is the mean reflectance of fine resolution homogeneous pixels belonging to land cover class c and $\xi(x,y,t)$ is the residual error term. Using the ordinary least squares technique, time series mean reflectance values can be obtained by solving the linear system of Equation (15). Then based on the assumption that the temporal variation properties of each fine resolution pixels in the same class are constant, the STDFA model predicts synthetic high spatial imagery as follows:[30]

$$r(x,y,t_k)=r(x,y,t_0)+\bar{r}(c,t_k)-\bar{r}(c,t_0) \quad (16)$$

where $\bar{r}(c,t_k)$ and $\bar{r}(c,t_0)$ are the mean reflectance of land cover c.

3.2 Application of spatial and temporal fusion models

For the STARFM, ESTARFM, and STDFA methods, two types of data must be input: the reference images and the time series MODIS LST products. These latter must contain MODIS LST data acquired on the same day as the reference image, and at least one MODIS LST data set acquired on the same day as the ASTER LST image that we want to predict. The selection of reference images has an important influence on the fusion accuracy.[36] Usually, the ASTER LST data that are from a date near the predicted date are used as the reference images. However, since the temporal change in LST over time may not be constant, similar dates do not necessarily indicate similarity in LST. Therefore, in this work, the ASTER LST data used for the reference images were selected based on the day with the highest correlation of MODIS data between that day and the day being predicted. Two-day ASTER and MODIS LST images are also needed in the ESTARFM. The early day ASTER and MODIS LST images were acquired on the same day as the reference ASTERLST image. The later day ASTER and MODIS LST images were acquired on a date later than the day being predicted. Two days of ASTER LST images (one day for the reference ASTERLST image, the second after the prediction date) were also needed for classification in the STDFA. For the STARFM and ESTARFM, the MODIS LST images must be resampled to the same resolution of the ASTER LST data by the nearest neighbour resampling method.

3.3 Evolution of spatial and temporal fusion models

Since the objective of the STARFM, ESTARFM, and STDFA methods was to generate synthetic ASTER LST data, nine real ASTER LST data sets were used to validate the algorithm. Obviously, the closer the synthetic ASTER LST imagery is to the actual imagery, the higher the precision. The algorithm was validated using two methods. First, visual interpretation was used to qualitatively evaluate the difference between the synthetic and real ASTER LST imagery. If the synthetic and real ASTER LST imagery can be distinguished visually, the accuracy of the method is not very high. Second, a correlation analysis was used to quantitatively evaluate the similarity between the actual observations and the synthetic imagery. Several indicators, such as the coefficient (r), variance, mean absolute difference(MAD), bias, and root mean square error(RMSE), were used to represent the precision of this model. Higher r and lower variance, MAD, bias, and RMSE indicate a higher accuracy.

4. Results

4.1. Overall accuracy comparison

Using the STARFM, ESTARFM, and STDFA, 29 synthetic ASTER LST images were generated using each method. Nine actual ASTER LST images were used to evaluate the accuracy of those three methods. Figure 2 shows the actual MODIS LST data, the actual ASTER LST data, and three synthetic ASTER LST images generated by those three methods on 10 July 2012. Through visual interpretation, we found that the synthetic ASTER LST data generated by the ESTARFM and STDFA were very similar to the actual ASTER LST data, while the resolution of the synthetic ASTER LST data generated by the STARFM was some what lower the actual ASTER LST data. However, all synthetic ASTER LST images had a higher resolution than the MODIS LST product. Table 2 shows the result of the correlation analysis. From Table 2, it is seen that all methods generate ASTER LST images very similar to the actual ones, with coefficient (r) higher than 0.92 and RMSE lower than 3.4k. The ESTARFM had the best performance, followed by the STDFA and then the STARFM.

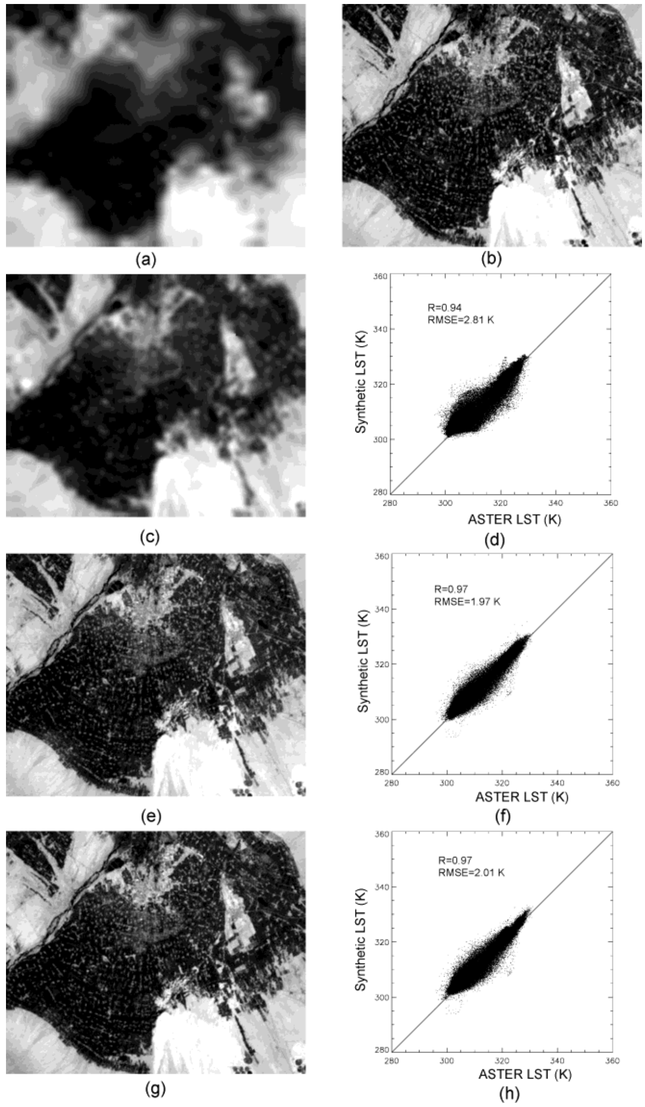

Fig.2A comparison between MODIS LST, ASTER LST, and synthetic LST data generated by the STARFM, ESTARFM, and STDFA acquired on 10 July 2012: (a) MOD11A1 LST data; (b) ASTER LST data; (c) synthetic LST data generated by the STARFM; (d) scatter plot between the ASTER LST and synthetic LST generated by the STARFM; (e) synthetic LST data generated by the ESTARFM; (f) scatter plot between the ASTER LST and synthetic LST generated by the ESTARFM; (g) synthetic LST data generated by the STDFA; (h) scatter plot between the ASTER LST and synthetic LST generated by the STDFA.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Table 2 Comparison between actual LST and synthetic LST images

Day	STARFM					ESTARFM					STDFA				
	R	Var	MAD	RMSE	bias	R	Var	MAD	RMSE	bias	R	Var	MAD	RMSE	bias
06/24/2012	0.94	8.92	2.26	2.99	0.09	0.96	5.87	1.79	2.43	-0.09	0.96	6.42	1.90	2.54	-0.09
07/10/2012	0.94	7.79	2.06	2.81	-0.29	0.97	3.73	1.38	1.97	-0.38	0.97	3.82	1.40	2.01	-0.48
08/02/2012	0.95	6.57	1.85	2.60	0.45	0.97	3.73	1.38	1.97	0.38	0.97	3.62	1.37	1.96	0.48
08/11/2012	0.94	6.32	1.89	2.59	-0.64	0.97	3.48	1.36	2.01	-0.75	0.95	5.06	1.69	2.35	-0.66
08/18/2012	0.92	6.44	1.98	2.71	-0.95	0.95	7.54	2.24	2.93	-1.03	0.94	6.62	2.14	2.81	-1.13
08/28/2012	0.92	11.15	2.56	3.38	-0.49	0.93	9.36	2.29	3.12	-0.58	0.94	7.85	2.19	2.85	-0.53
09/03/2012	0.93	6.74	1.94	2.67	0.63	0.92	9.36	2.29	3.12	0.58	0.93	7.50	2.15	2.79	0.54
09/12/2012	0.93	6.99	1.94	2.67	-0.37	0.96	4.18	1.41	2.09	-0.46	0.95	5.36	1.58	2.38	-0.55
09/19/2012	0.94	5.26	1.68	2.35	0.52	0.96	4.18	1.41	2.10	0.46	0.94	5.25	1.58	2.35	0.54

4.2 Accuracy comparison of each land cover types

To demonstrate the performance of the STARFM, ESTARFM, and STDFA for each land cover type, an accuracy comparison for each land cover was conducted on those nine days. Table 3 showed the correlation analysis results for those three methods with each land cover type. Figure 3 shows scatter plots with the actual and synthetic ASTER LST images in each class, which fit well to the 1:1 lines. From Table 3 and Fig. 3, we find that the ESTARFM and STDFA have the same precision evaluated using all nine days of images in the desert class. The ESTARFM had slightly better performance than the STDFA in the crops and cities classes. Further more, the STARFM had the worst performance in each land cover type.

Table 3 Comparison between actual LST and synthetic LST imageries of different land cover types

	All nine days			Crops			Cities			Desert		
	STARFM	ESTARFM	STDFA	STARFM	ESTARFM	STDFA	STARFM	ESTARFM	STDFA	STARFM	ESTARFM	STDFA
R	0.73	0.96	0.96	0.15	0.62	0.55	0.35	0.80	0.79	0.75	0.96	0.96
Var	50.00	6.01	6.05	65.08	5.33	6.00	76.88	10.10	10.83	23.90	3.33	3.61
MAD	2.20	1.79	1.84	2.15	1.79	1.94	3.21	2.45	2.55	1.74	1.30	1.37
RMSE	7.07	2.46	2.47	8.13	2.31	2.45	8.80	3.22	3.34	4.94	1.83	1.90
bias	0.01	-0.21	-0.21	-1.00	0.00	0.11	0.77	-0.51	-0.58	0.70	0.12	0.06

Environmental Science: Processes & Impacts

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

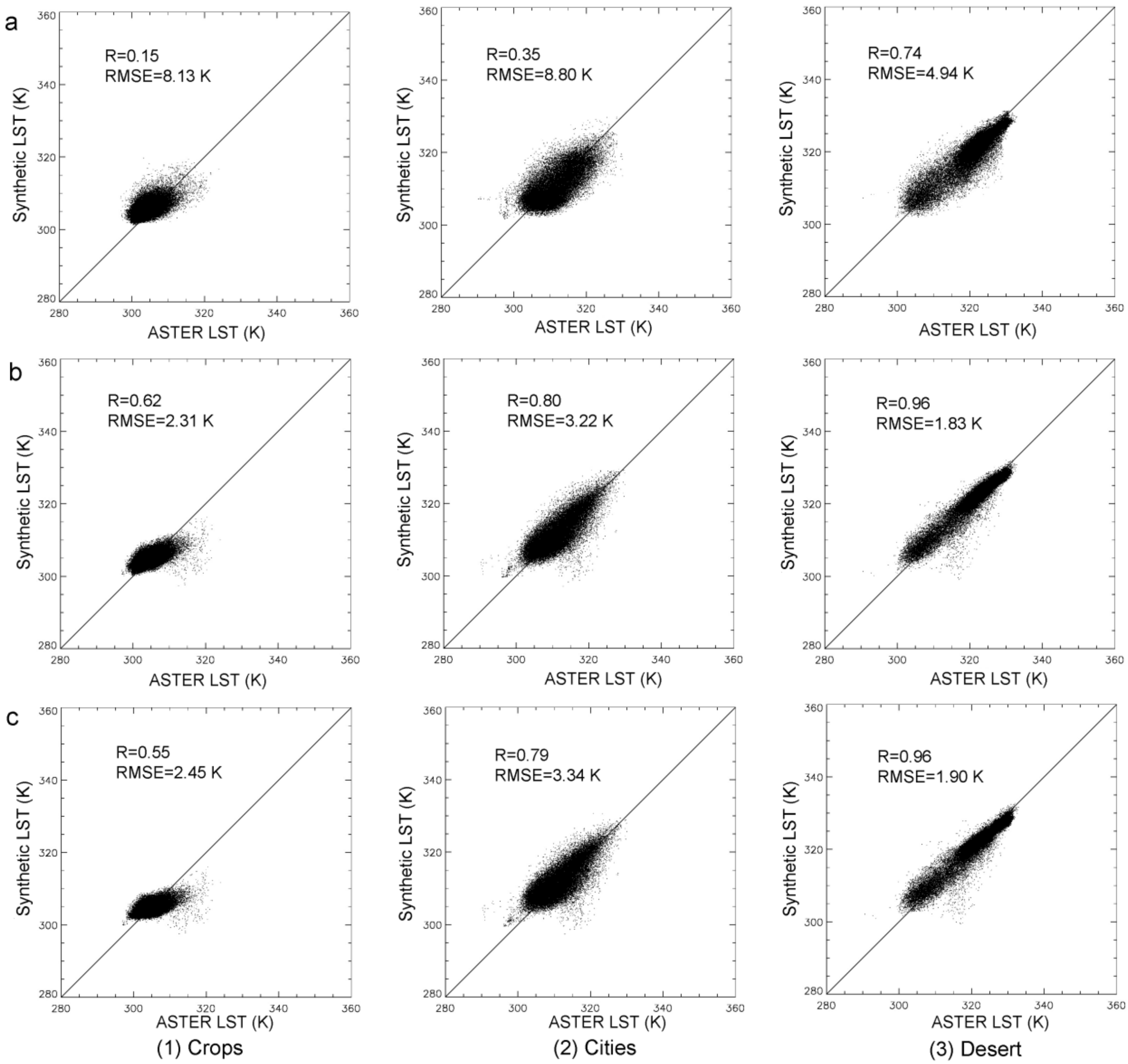

Fig.3 Comparison of ASTER LST and synthetic LST data for different land cover types generated by (a) the STARFM,(b) the ESTARFM, and (c) the STDFA.

5. Discussion

The correlation analysis between the actual observations and the synthetic ASTER LST images showed high correlation with r higher than 0.92. This demonstrated that spatial and temporal data fusion methods, such as the STARFM, ESTARFM, and STDFA, can be used to combine ASTER LST and MODIS LST data to generate daily synthetic ASTER LST data. However, issues remain that should be addressed in the application of these methods.

(1) Influence of noise. ASTER and MODIS image noise will reduce the accuracy of the three models. Table 4 shows the correlation analysis results between the actual observations and the synthetic images generated without removing the default data from the MODIS LST data. Comparing Table2 and Table4, we see that the STDFA had the smallest decrease in correlation coefficient r and had the highest tolerance for noise, while the ESTARFM had the greatest decrease in correlation coefficient r and had the lowest tolerance for noise. This is because the STARFM and ESTARFM use a 3×3 window in their prediction,

while the STDFA uses a 40×40 window. Noise values appearing in the window would therefore have a greater influence on the ESTARFM.

Table 4 Comparison of ASTER LST and synthetic LST data on 10 July 2012 generated by the STARFM, ESTARFM, and STDFA. MODIS LST data is input with noise.

	STARFM	ESTARFM	STDFA
R	0.92	0.77	0.96
Var	10.35	36.90	6.68
MAD	2.17	2.57	2.13
RMSE	3.22	6.09	2.92
bias	-0.20	0.45	1.36

(2) Influence of spatial variability. The spatial and temporal data fusion methods assume that the temporal effects are constant over time. This may be true in classes like deserts and cities, but may be violated in classes like crops. The temporal changes in crops are easily affected by farm management practices such as irrigation,[36] so there would be reduced accuracy for crops. Different methods were used to handle the spatial variability of reflectance. The STARFM and ESTARFM use additional information from a 3×3 window to reduce the influence of spatial variability, while the STDFA uses a 40×40 window. Thus, spatial variability still exists in the 40×40 window of the STDFA, which results in its lower accuracy compared to the ESTARFM. The ESTARFM is more suitable for complex surface areas.

(3) The spatial and temporal data fusion methods use optical images that are easy affected by cloudy weather. For example, during the 106 days from 15 June 2012 to 28 September2012, there have been only 29 MODIS images without clouds. The effective data rate is only 23%. Therefore, the development of an optical and radar data fusion algorithm is an important direction for multi-source remote sensing data.[37-42]

6. Conclusions

The STARFM, ESTARFM, and STDFA methods were compared and validated in the generation of daily ASTER LST data in Zhangye Country, Gansu province, China. The results showed the following:

(1) All three methods can generate synthetic ASTER LST images very similar to the actual ASTER LST imageries, with r values higher than 0.92 and RMSE values lower than 3.4k. The ESTARFM had the best performance, followed by the STDFA and then the STARFM.

(2) All three methods had their best performance with desert land types. The next best performance was for cities, and the worst performance was for crops. The ESTARFM and STDFA have the same precision in the desert class, while the ESTARFM had slightly better performance than the STDFA in the crops and cities classes. The STARFM method had the worst performance for each land cover type.

(3) ASTER and MODIS image noise will reduce the accuracy of these three models. The STDFA had the best tolerance to noise, while the ESTARFM had the worst tolerance to noise.

Acknowledgments

We thank the Cold and Arid Regions Science Data Center at Lanzhou, Gansu, China, for providing data: "Li Hua, Wang Heshun, Du Yongming, Xiao Qing, Liu Qinhuo. Hi WATER: ASTER LST and LSE dataset in 2012 in the middle reaches of the Heihe River Basin. Cold and Arid Regions Science Data Center at Lanzhou, 2015. doi:10.3972/hiwater.220.2015.db". This work was supported by the National Natural Science Foundation of China (41301390), the National Science and Technology Major Project (2014AA06A511) and the Major State Basic Research Development Program of China (2013CB733405, 2010CB950603), and the National Science and Technology Major Project of China. The funders had no role in choosing the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Notes and references

[a]*The State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China. Fax: +861064806258; Tel: +861064806258; E-mail: wumq@radi.ac.cn*
[b]*Laboratory of Digital Earth Sciences, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing,China. Fax: +861082178008; Tel: +861082178009; E-mail:Huangwj@radi.ac.cn*

1. M.C. Anderson, W.P. Kustas, J.M. Norman, C.R. Hain, J. R. Mecikalski, L. Schultz, et al., Mapping daily evapotranspiration at field to continental scales using geostationaryand polar orbiting satellite imagery, *Hydrology and Earth System Sciences* 2011, **15**(1), 223–239.
2. T. Carlson, An overview of the “triangle method” for estimating surface evapotranspiration and soil moisture from satellite imagery. *Sensors*, 2007, **7**(8), 1612–1629.
3. M.S. Moran, Thermal infrared measurement as an indicator of plant ecosystemhealth, *Thermal Remote Sensing in Land Surface Processes*, 2004, 257–282.
4. Z.M. Wan, Z.L. Li, A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Transaction on Geoscience and Remote Sensing* 1997, **35**(4), 980-996.
5. E. Shamir, K.P. Georgakakos, MODIS Land Surface Temperature as an index of surface air temperature for operational snowpack estimation, *Remote Sensing of Environment* 2014, **152**, 83–98.
6. Z.Q. Lv, Q.G. Zhou, Utility of Landsat Image in the Study of Land Cover and Land Surface Temperature Change, *Procedia Environmental Sciences* 2011, **20**, 1287–1292.
7. S.N.A. Buyadi, W.M.N.W. Mohd, A. Misni, Impact of Land Use Changes on the Surface Temperature Distribution of Area Surrounding the National Botanic Garden, Shah Alam, *Procedia - Social and Behavioral Sciences* 2013, **101**, 516-525.
8. J. Jiang, G. Tian, Analysis of the impact of Land use/Land cover change on Land Surface Temperature with Remote Sensing, *Procedia Environmental Sciences* 2010, **2**, 571-575
9. G. Guo, Z. Wu, R. Xiao, Y. Chen, X. Liu, X. Zhang, Impacts of urban biophysical composition on land surface temperature in urban heat island clusters, *Landscape and Urban Planning* 2015, **135**, 1-10.
10. L. Hu, N.A. Brunsell, The impact of temporal aggregation of land surface temperature data for surface urban heat island (SUHI) monitoring, *Remote Sensing of Environment* 2013, **134**, 162-174.
11. M.L. Imhoff, P. Zhang, R.E. Wolfe, & L. Bounoua, Remote sensing of the urbanheat island effect across biomes in the continental USA, *Remote Sensing of Environment* 2010, **114**(3), 504 – 513.
12. P.D. Jones, GLOBAL CHANGE | Climate Record: Surface Temperature Trends, *Reference Module in Earth Systems and Environmental Sciences, from Encyclopedia of Atmospheric Sciences (Second Edition)*, 2015, 110-120.

13. Z. Guo, S.D. Wang, M.M. Cheng, Y. Shu, Assess the effect of different degrees of urbanization on land surface temperature using remote sensing images, *Procedia Environmental Sciences* 2012, **13**, 935-942.
14. S. Duan, Z. Li, B. Tang, H. Wu, R. Tang, Generation of a time-consistent land surface temperature product from MODIS data, *Remote Sensing of Environment* 2014, **140**, 339-349.
15. M. Hais, T. Kučera, The influence of topography on the forest surface temperature retrieved from Landsat TM, ETM + and ASTER thermal channels, *ISPRS Journal of Photogrammetry and Remote Sensing* 2009, **64**, 585-591.
16. S. Zhao, Q. Qin, F. Zhang, et al., Research on Using a Mono-Window Algorithm for Land Surface Temperature Retrieval from Chinese Satellite for Environment and Natural Disaster Monitoring(HJ-1B) Data, *Spectroscopy and Spectral Analysis* 2011, **31**, 1552-1556.
17. F. Li, T.J. Jackson, W.P. Kustas, et al., Deriving land surface temperature from Landsat 5 and 7 during SMEX02/SMACEX, *Remote Sensing of Environment* 2004, **92**, 521-534.
18. N.T. Son, C.F. Chen, C.R. Chen, L.Y. Chang, V.Q. Minh, Monitoring agricultural drought in the Lower Mekong Basin using MODIS NDVI and land surface temperature data, *International Journal of Applied Earth Observation and Geoinformation* 2012, **18**, 417-427.
19. F. Chen, S. Yang, Z. Su, Ba. He, A new single-channel method for estimating land surface temperature based on the image inherent information: The HJ-1B case, *ISPRS Journal of Photogrammetry and Remote Sensing* 2015, **101**, 80-88.
20. F. Gao, J. Masek, M. Schwaller, F. Hall, On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance, *IEEE Transactions on Geoscience and Remote Sensing* 2006, **44**(8), 2207-2218.
21. T. Hilker, M.A.Wulder, N.C. Coops, et al., A new data fusion model for high spatial- and temporal- resolution mapping of forest disturbance based on Landsat and MODIS, *Remote Sensing of Environment* 2009, **113**(8), 1613-1627.
22. H. Liu, Q.H. Weng, 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* **2012**, 117, 57–71.
23. J.J. Walker, K.M. de Beurs, R.H. Wynne, and F. Gao, Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology, *Remote Sensing of Environment* 2012, **117**, 381–393.
24. Q.H. Weng, P. Fua, F. Gao, Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data, *Remote Sensing of Environment* 2014, **145**, 55–67.
25. M. Schmidt, R. Lucas, P. Bunting, et al., Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia, *Remote Sensing of Environment* 2015, **158**, 156–168.
26. X.L. Zhu, J. Chen, F. Gao, et al., An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions, *Remote Sensing of Environment* 2010, **114**, 11, 2610-2623.
27. L. Busetto , M. Meronib, R. Colombo, Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series, *Remote Sensing of Environment* 2008, **112**(1), 118–131.
28. B. Zhukov, D. Oertel, F. Lanzl, and G.Reinhackel, Unmixing-based multisensor multiresolution image fusion, *IEEE Transactions on Geoscience and Remote Sensing* 1999, **37**, 1212-1226.
29. F.Maselli, Definition of spatially variable spectral endmembers by locally calibrated multivariate regression analyses, *Remote Sensing of Environment* 2001,**75**(1), 29-38.
30. M.Q. Wu, Z. Niu, C.Y. Wang, C.Y. Wu, L.Wang, The use of MODIS and Landsat time series data to generate high resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model, *Journal of Applied Remote Sensing* 2012, **6**, 1-8.
31. M.Q. Wu, C.Y. Wu, W.J. Huang, Z. Niu, C.Y. Wang, High-resolution Leaf Area Index estimation from synthetic Landsat data generated by a spatial and temporal data fusion model, *Comput. Electron. Agric.*, 2015, **115**, 1-11.
32. M.Q. Wu, Z. Niu, C.Y. Wang, Mapping paddy fields by using spatial and temporal remote sensing data fusion technology, *Transactions of the CSAE* 2010, **26**(Supp.2), 48-52.
33. H. Li, D. Sun, Y. Yu, H. Wang, Y. Liu, Q. Liu, Y. Du, H. Wang, B. Cao, Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China, *Remote Sensing of Environment* 2014, **142**, 111-121.
34. H. Wang, Q. Xiao, H. Li, Y. Du,Q. Liu, Investigating the Impact of Soil Moisture on Thermal Infrared Emissivity Using ASTER Data. *IEEE Geoscience and Remote Sensing Letters* 2015, **12**(2): 294-298.
35. J.J. Settle, N.A. Drake, Linear mixing and the estimation of groundcover proportion, *International Journal of Remote Sensing* 1993, **14**, 1159-1177.
36. Y. Shi, G. Yang, L. C, et al., Intercomparison of the different fusion methods for generating high spatial-temporal resolution data, *J. Infrared Millim. waves* 2015, **34**, 92-99.
37. D. McAlpin, F.J. Meyer, Multi-sensor data fusion for remote sensing of post-eruptive deformation and depositional features at Redoubt Volcano, *Journal of Volcanology and Geothermal Research* 2013, **259**, 414–423.
38. J. Amorós-López, L. Gómez-Chova, L. Alonso, et al., Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring, *International Journal of Applied Earth Observation and Geoinformation* 2013, **23**, 132–141
39. J. Reiche, J. Verbesselt, D. Hoekman, et al., Fusing Landsat and SAR time series to detect deforestation in the tropics, *Remote Sensing of Environment* 2015, **156**, 276–293
40. M. Liu, X. Liu, J. Li, et al., Evaluating total inorganic nitrogen in coastal waters through fusion ofmulti-temporal RADARSAT-2 and optical imagery using random forest algorithm, *International Journal of Applied Earth Observation and Geoinformation* 2014, **33** ,192–202.
41. Y. Zhang, H. Zhang, H. Lin. Improving the impervious surface estimation with combined use of optical and SAR remote sensing images, *Remote Sensing of Environment* 2014, **141**, 155–167.
42. M. Schmidt, R. Lucas, P. Bunting, et al., Multi-resolution time series imagery for forest disturbance and regrowth monitoring in Queensland, Australia, *Remote Sensing of Environment* 2015, **158**, 156–168