

GEOSTATISTICAL SOLUTIONS FOR DOWNSCALING REMOTELY SENSED LAND SURFACE TEMPERATURE

Qunming Wang^{a,*}, Victor Rodriguez-Galiano^b, Peter M. Atkinson^{a,c,d}

^a Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK - wqm11111@126.com

^b Department of Physical Geography, University of Seville, Seville 41004, Spain

^c School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, BT7 1NN, Northern Ireland, UK

^d Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK

KEY WORDS: Image fusion, Downscaling, Geostatistics, Land surface temperature (LST), Landsat thermal imagery

ABSTRACT:

Remotely sensed land surface temperature (LST) downscaling is an important issue in remote sensing. Geostatistical methods have shown their applicability in downscaling multi/hyperspectral images. In this paper, four geostatistical solutions, including regression kriging (RK), downscaling cokriging (DSCK), kriging with external drift (KED) and area-to-point regression kriging (ATPRK), are applied for downscaling remotely sensed LST. Their differences are analyzed theoretically and the performances are compared experimentally using a Landsat 7 ETM+ dataset. They are also compared to the classical TsHARP method.

1. INTRODUCTION

Land surface temperature (LST) is a crucial variable that impacts on many aspects of the human and physical environments. Remotely sensed images provide good sources for estimating LST, which can be retrieved from the thermal infrared (TIR) bands (e.g., in Landsat 5 TM, Landsat 7 ETM+ or Landsat 8 TIRS sensor (TIRS) data). However, the TIR band often has a coarser spatial resolution than the corresponding multispectral bands. For example, the spatial resolutions of TIR bands of Landsat 5 TM, Landsat 7 ETM+ and Landsat 8 TIRS are 120 m, 60 m and 100 m, but the corresponding multispectral bands of the Landsat data are all at a finer spatial resolution of 30 m. In real applications, users always need to obtain more detailed spatial information of LST. Thus, it is great interest to downscale the LST to a finer spatial resolution for more reliable applications.

The availability of finer spatial resolution (30 m) Landsat multispectral bands provides excellent opportunity to be fused with the LST image to produce 30 m LST. A classical method based on TsHARP (Agam et al., 2007) was developed for this issue. However, this method produces block artifacts in predictions. In this paper, four geostatistical solutions, including regression kriging (RK) (Mukherjee et al., 2015), downscaling cokriging (DSCK) (Atkinson et al., 2008; Pardo-Igúzquiza et al., 2006; Rodriguez-Galiano et al., 2012), kriging with external drift (KED) (Sales et al., 2013), and area-to-point regression kriging (ATPRK) (Wang et al., 2015, 2016a,b), were applied for downscaling LST retrieved from Landsat 7 ETM+ TIR band, with the aid of the finer spatial resolution Normalized Difference Vegetation Index (NDVI) image derived from the corresponding multispectral bands. Note that RK and DSCK were already applied for downscaling LST in earlier studies (Mukherjee et al., 2015, Rodriguez-Galiano et al., 2012). It is the first time to adopt KED and ATPRK for the issue in this paper.

2. METHODS

In this paper, the LST image was derived from the TIR band based on the radiative transfer equation method (Rodriguez-Galiano et al., 2012). Let $Z_C(\mathbf{x}_i)$ be the LST value of pixel C centered at \mathbf{x}_i ($i=1, \dots, M$, where E is the number of pixels) in the LST image, and $Y_F(\mathbf{x}_j)$ be the NDVI value of pixel F centered at \mathbf{x}_j ($j=1, \dots, EG^2$, where G is the spatial resolution (zoom) ratio between the LST and NDVI images) in the NDVI image. The notations F and C denote the fine and coarse pixels, respectively. The objective of downscaling is to predict target variables $Z_F(\mathbf{x})$ for all fine pixels. The principles of the four geostatistical solutions are described below.

2.1 ATPRK

The ATPRK method consists of two steps: regression modelling (for overall trend estimation) and area-to-point kriging (ATPK) (for downscaling residual remaining after removal of the trend). Suppose $\hat{Z}_{F-R}(\mathbf{x})$ and $\hat{Z}_{F-ATPK}(\mathbf{x})$ are predictions of the regression and ATPK parts, the ATPRK prediction is

$$\hat{Z}_{F-ATPRK}(\mathbf{x}) = \hat{Z}_{F-R}(\mathbf{x}) + \hat{Z}_{F-ATPK}(\mathbf{x}). \quad (1)$$

For regression prediction, it makes full use of the fine spatial resolution information in the NDVI image, and is calculated as a linear combination of the NDVI image

$$\hat{Z}_{F-R}(\mathbf{x}) = aY_F(\mathbf{x}) + b \quad (2)$$

where a and b are two coefficients. They are estimated from the regression model constructed between the coarse LST image $Z_C(\mathbf{x})$ and the NDVI image upscaled to the coarse spatial resolution, $Y_C(\mathbf{x})$, as shown in (3)

* Corresponding author

$$Z_C(\mathbf{x}) = aY_C(\mathbf{x}) + b + R(\mathbf{x}) \quad (3)$$

where $R(\mathbf{x})$ is the residual term in the regression model. The two coefficients can be estimated by ordinary least squares.

For ATPK prediction, it is calculated as a liner combination of the N coarse residuals surrounding the fine pixel

$$\hat{Z}_{F-ATPK}(\mathbf{x}) = \sum_{i=1}^N \beta_i R(\mathbf{x}_i), \text{ s.t. } \sum_{i=1}^N \beta_i = 1 \quad (4)$$

where β_i is the kriging weight for the i th neighboring coarse pixel. The N weights are calculated by the matrix below

$$\begin{bmatrix} \gamma_{CC}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{CC}(\mathbf{x}_1, \mathbf{x}_N) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{CC}(\mathbf{x}_N, \mathbf{x}_1) & \dots & \gamma_{CC}(\mathbf{x}_N, \mathbf{x}_N) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \\ \theta \end{bmatrix} = \begin{bmatrix} \gamma_{FC}(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \gamma_{FC}(\mathbf{x}, \mathbf{x}_N) \\ 1 \end{bmatrix} \quad (5)$$

In (5), $\gamma_{CC}(\mathbf{x}_i, \mathbf{x}_j)$ is the coarse-to-coarse residual semivariogram between coarse pixels centered at \mathbf{x}_i and \mathbf{x}_j , $\gamma_{FC}(\mathbf{x}, \mathbf{x}_j)$ is the fine-to-coarse residual semivariogram between fine and coarse pixels centered at \mathbf{x} and \mathbf{x}_j , and θ is the Lagrange multiplier.

Suppose \mathbf{s} is the Euclidean distance between the centroids of any two pixels. The semivariograms in (5) are calculated by convolving the fine-to-fine residual semivariogram $\gamma_{FF}(\mathbf{s})$ with the point spread function (PSF) $h_C(\mathbf{s})$

$$\gamma_{FC}(\mathbf{s}) = \gamma_{FF}(\mathbf{s}) * h_C(\mathbf{s}) \quad (6)$$

$$\gamma_{CC}(\mathbf{s}) = \gamma_{FF}(\mathbf{s}) * h_C(\mathbf{s}) * h_C(-\mathbf{s}). \quad (7)$$

The fine-to-fine residual semivariogram $\gamma_{FF}(\mathbf{s})$ is derived by deconvolution of the coarse residual semivariogram of the coarse residual image $R(\mathbf{x})$. Details on deconvolution can be found in Wang et al. (2016a).

2.2 KED

In KED, for each fine pixel, the prediction is calculated as a linear combination of the N surrounding coarse LST pixels

$$\begin{bmatrix} \gamma_{CC}^{ZZ}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{CC}^{ZZ}(\mathbf{x}_1, \mathbf{x}_N) & \gamma_{CF}^{ZY}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{CF}^{ZY}(\mathbf{x}_1, \mathbf{x}_M) & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \gamma_{CC}^{ZZ}(\mathbf{x}_N, \mathbf{x}_1) & \dots & \gamma_{CC}^{ZZ}(\mathbf{x}_N, \mathbf{x}_N) & \gamma_{CF}^{ZY}(\mathbf{x}_N, \mathbf{x}_1) & \dots & \gamma_{CF}^{ZY}(\mathbf{x}_N, \mathbf{x}_M) & 1 & 0 \\ \gamma_{FC}^{YZ}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{FC}^{YZ}(\mathbf{x}_1, \mathbf{x}_N) & \gamma_{FF}^{YY}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{FF}^{YY}(\mathbf{x}_1, \mathbf{x}_M) & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \gamma_{FC}^{YZ}(\mathbf{x}_M, \mathbf{x}_1) & \dots & \gamma_{FC}^{YZ}(\mathbf{x}_M, \mathbf{x}_N) & \gamma_{FF}^{YY}(\mathbf{x}_M, \mathbf{x}_1) & \dots & \gamma_{FF}^{YY}(\mathbf{x}_M, \mathbf{x}_M) & 0 & 1 \\ 1 & \dots & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \\ \beta_1 \\ \vdots \\ \beta_M \\ \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} \gamma_{FC}^{ZZ}(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \gamma_{FC}^{ZZ}(\mathbf{x}, \mathbf{x}_N) \\ \gamma_{FF}^{ZY}(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \gamma_{FF}^{ZY}(\mathbf{x}, \mathbf{x}_M) \\ 1 \\ 0 \end{bmatrix} \quad (11)$$

$$\hat{Z}_{F-KED}(\mathbf{x}) = \sum_{i=1}^N \alpha_i Z_C(\mathbf{x}_i), \text{ s.t. } \sum_{i=1}^N \alpha_i = 1. \quad (8)$$

The kriging weight α_i is estimated by the kriging matrix below

$$\begin{bmatrix} \gamma_{CC}(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma_{CC}(\mathbf{x}_1, \mathbf{x}_N) & 1 & Y_C(\mathbf{x}_1) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \gamma_{CC}(\mathbf{x}_N, \mathbf{x}_1) & \dots & \gamma_{CC}(\mathbf{x}_N, \mathbf{x}_N) & 1 & Y_C(\mathbf{x}_N) \\ 1 & \dots & 1 & 0 & 0 \\ Y_C(\mathbf{x}_1) & \dots & Y_C(\mathbf{x}_N) & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \\ \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} \gamma_{FC}(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \gamma_{FC}(\mathbf{x}, \mathbf{x}_N) \\ 1 \\ Y_F(\mathbf{x}) \end{bmatrix} \quad (9)$$

where $\gamma_{CC}(\mathbf{x}_i, \mathbf{x}_j)$ and $\gamma_{FC}(\mathbf{x}, \mathbf{x}_j)$ have the same meanings as those in (5). $Y_C(\mathbf{x}_i)$ is the NDVI for coarse pixel centered at \mathbf{x}_i . The kriging weights need to be calculated for each fine pixel.

2.3 DSK

In DSK, the coarse LST and fine NDVI are considered jointly. The prediction is a linear combination of the N neighboring coarse LST pixels and M neighboring fine NDVI pixels.

$$\hat{Z}_{F-DSK}(\mathbf{x}) = \sum_{i=1}^N \alpha_i Z_C(\mathbf{x}_i) + \sum_{j=1}^M \beta_j Y_F(\mathbf{x}_j) \quad (10)$$

$$\text{s.t. } \sum_{i=1}^N \alpha_i = 1, \sum_{j=1}^M \beta_j = 0$$

The kriging weights α_i and β_j are estimated by the kriging matrix in (11), where $\gamma_{CC}^{ZZ}(\mathbf{x}_i, \mathbf{x}_j)$ and $\gamma_{FF}^{YY}(\mathbf{x}_i, \mathbf{x}_j)$ are coarse and fine auto-semivariograms estimated from the original coarse LST pixels and fine NDVI pixels, respectively. $\gamma_{CF}^{ZY}(\mathbf{x}_i, \mathbf{x}_j)$ is the cross-semivariograms between coarse LST and fine NDVI pixels. $\gamma_{FF}^{ZY}(\mathbf{x}, \mathbf{x}_j)$ is the cross-semivariogram between fine LST and fine NDVI images, and $\gamma_{FC}^{ZZ}(\mathbf{x}, \mathbf{x}_j)$ is the fine-to-coarse semivariogram between fine and coarse LST pixels. $\gamma_{FF}^{ZY}(\mathbf{x}, \mathbf{x}_j)$ and $\gamma_{FC}^{ZZ}(\mathbf{x}, \mathbf{x}_j)$ need to be estimated by deconvolution and convolution processes.

2.4 RK

RK is a method different to ATPRK, where the coarse residuals from the regression model in (3) are downscaled to the fine spatial resolution by ordinary kriging

$$\hat{Z}_{F-K}(\mathbf{x}) = \sum_{i=1}^N \lambda_i R(\mathbf{x}_i), \text{ s.t. } \sum_{i=1}^N \lambda_i = 1 \quad (12)$$

where λ_i is the kriging weight for the i th neighboring coarse pixel, and the weights are calculated by the matrix in (13)

$$\begin{bmatrix} \gamma(\mathbf{x}_1, \mathbf{x}_1) & \dots & \gamma(\mathbf{x}_1, \mathbf{x}_N) & 1 \\ \vdots & & \vdots & \vdots \\ \gamma(\mathbf{x}_N, \mathbf{x}_1) & \dots & \gamma(\mathbf{x}_N, \mathbf{x}_N) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \\ \theta \end{bmatrix} = \begin{bmatrix} \gamma(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \gamma(\mathbf{x}, \mathbf{x}_N) \\ 1 \end{bmatrix}. \quad (13)$$

In (13), $\gamma(\mathbf{x}_i, \mathbf{x}_j)$ and $\gamma(\mathbf{x}, \mathbf{x}_j)$ are estimated from the coarse residuals directly, where the coarse or fine pixel is treated as a centroid. The RK prediction is a combination of the regression prediction in (2) and the kriging prediction in (12)

$$\hat{Z}_{F-RK}(\mathbf{x}) = \hat{Z}_{F-R}(\mathbf{x}) + \hat{Z}_{F-K}(\mathbf{x}) \quad (14)$$

2.5 Differences between the geostatistical solutions

The centroid-based RK method treats each coarse or fine pixel as a centroid and ignores the spatial support of pixels, and thus, its implementation is simple. Different from RK, the other three methods (DSCK, KED and ATPRK) take into account explicitly the size of support and predict variables from areal supports to points via semivariogram deconvolution to parameterize kriging for prediction (see the semivariograms in (5), (9) and (11)).

The differences between DSCK, KED and ATPRK lie mainly in model complexity and computational cost. Specifically, DSCK is a one-stage method which considers the coarse LST and fine NDVI simultaneously, by including both auto-semivariograms and cross-semivariogram in the single kriging matrix. The kriging weights need to be calculated only once. However, both the auto-semivariogram $\gamma_{FF}^{ZZ}(\mathbf{x}, \mathbf{x}_j)$ and cross-semivariogram $\gamma_{FC}^{ZZ}(\mathbf{x}, \mathbf{x}_j)$ in (11) involve complex deconvolution and convolution, which would sometimes require manual intervention.

KED does not need the cross-semivariogram that relates coarse LST with fine NDVI in DSCK. Alternatively, it includes the fine NDVI in the kriging matrix. The extension of the matrix can result in unstable matrix (also the case for DSCK), especially when the covariate (i.e., fine NDVI) does not vary smoothly in space (Goovaerts, 1997). Furthermore, the KED kriging weights are calculated locally for each fine pixel, which greatly increases the computational cost, especially for large areas.

As a recently developed solution, ATPRK noticeably simplifies the model complexity and reduces the computational cost. The utility of the cross-semivariogram in DSCK (i.e., accounting for the cross-correlation between coarse LST and fine NDVI) is

achieved by the simple regression model in ATPRK instead (see (2) and (3)). ATPRK requires only auto-semivariogram modelling in downscaling residuals. The size of the kriging matrix in ATPRK in (5) is much smaller than that in DSCK in (11). Thus, ATPRK is much easier to automate and more user-friendly than DSCK. Compared to KED, the kriging weights need to be calculated only once for the entire study area. This is because ATPRK explicitly separates trend estimation (i.e., regression modelling) from ATPRK-based spatial prediction of residuals. This makes ATPRK a fast image downscaling approach and free of the risk of instability in the kriging weights calculation.

With respect to the existing TsHARP method, it is also a two-step approach. The first step also involves regression modelling (as in ATPRK and RK). In the second step of TsHARP, however, the coarse residuals are added back to fine regression predictions straightforwardly. The inconsistency in spatial resolutions between the regression prediction and residual can result in blocky artifacts in TsHARP predictions.

3. EXPERIMENTS

The four geostatistical methods and the TsHARP method were tested using a data covering a 15 km by 15 km region in Granada, Spain. The data were acquired on 20 July, 2002. For objective evaluation, the 60 m LST and 30 m NDVI images were degraded to 120 m and 60 m, respectively. The task was to fuse the 120 m LST with the 60 m NDVI images to produce the 60 m LST image. The available 60 m LST image was used to evaluate the accuracy.

Fig. 1 shows the results of the five methods, and Fig. 2 shows the results of a sub-area for convenience of visual comparison. From the results, we can observe that there are noticeable blocky artifacts in the TsHARP result. Due to the unstable kriging matrix, there are also blocky artifacts in the KED result. Compared to RK, DSCK and ATPRK can reproduce more spatial variations (see the restoration of the red pixels).

The quantitative assessment is in Table 1, where five indices are used, including the root mean square error (RMSE), correlation coefficient (CC), universal image quality index (UIQI) (Wang and Bovik, 2012), relative global-dimensional synthesis error (ERGAS), and coherence. Coherence (quantified by the CC) is an index measuring the relation between the coarse LST and the coarse LST obtained by upscaling the LST downscaling predictions. As seen from Table 1, ATPRK can produce more accurate results than the other methods, and can perfectly preserve the original coarse LST data. Although the TsHARP prediction has perfect coherence with the coarse LST, the RMSE and ERGAS are larger, and the CC and UIQI are smaller than that of DSCK, KED and ATPRK. Failing to account for the size of supports, RK leads to the smallest coherence.

4. CONCLUSION

This paper studies the use of the four geostatistical solutions (RK, DSCK, KED and ATPRK) for downscaling LST and compares their performances based on a case study for the Landsat 7 ETM+ dataset. RK fails to account for the change of support in downscaling and produces less accurate result than the other three geostatistical solutions. Their accuracies are generally greater than the classical TsHARP method. Amongst all methods, ATPRK can produce the most accurate result, and can perfectly preserve the original coarse LST data.

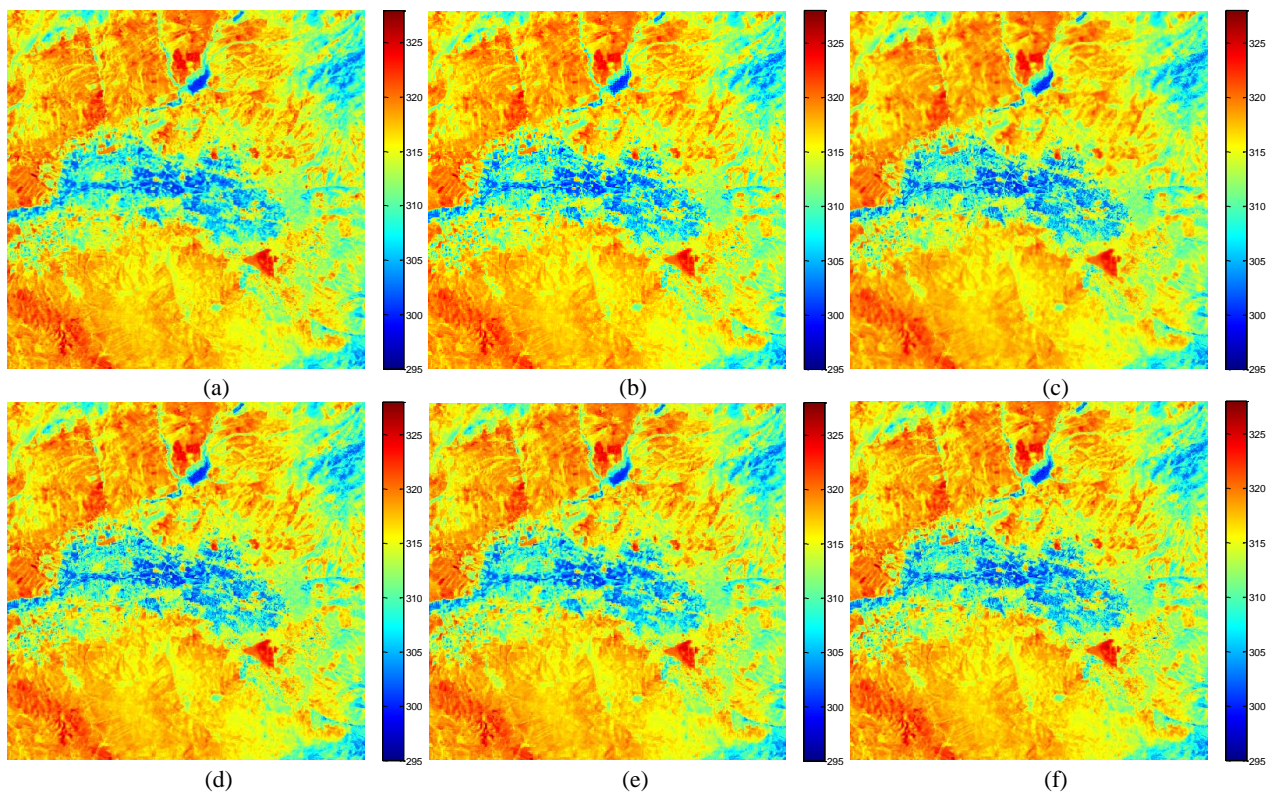
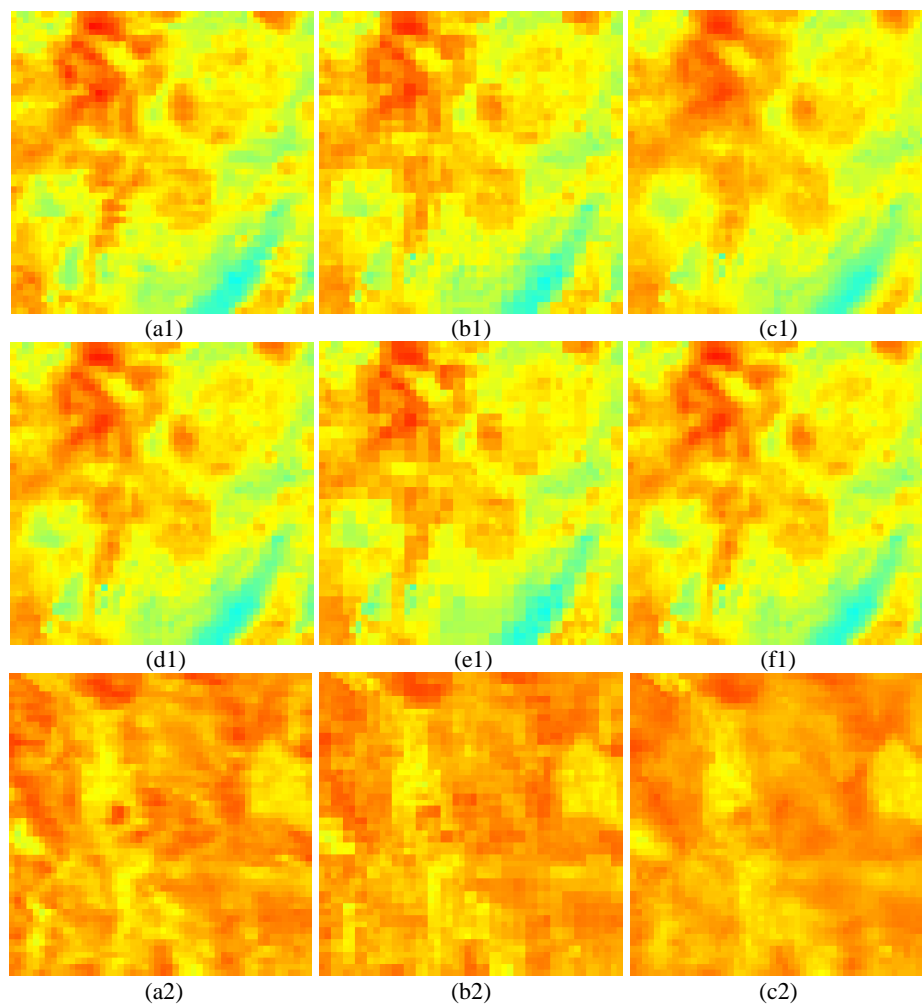


Figure 1. 60 m downscaling results of LST. (a) 60 m reference. (b) TsHARP. (c) RK. (d) DSCK. (e) KED. (f) ATPRK.



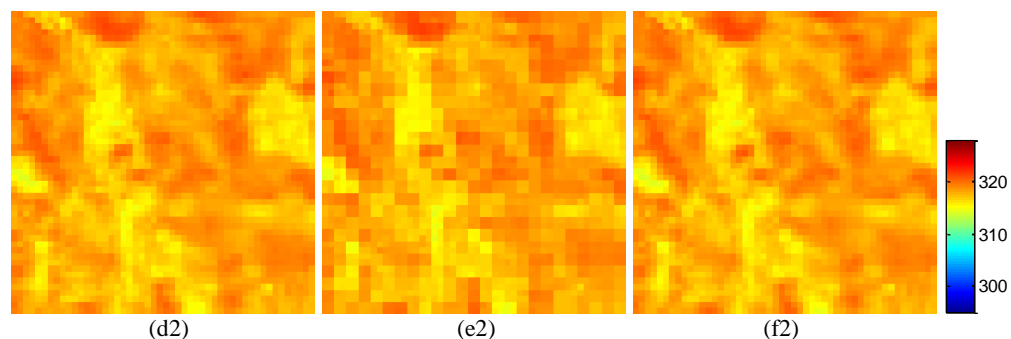


Figure 2. Results of two zoom areas in Fig. 1. (a) 60 m reference. (b) TsHARP. (c) RK. (d) DSCK. (e) KED. (f) ATPRK.

Table 1. Quantitative assessment on the methods for the entire study area

	RMSE	CC	ERGAS	UIQI	Coherence
Ideal	0	1	0	1	1
TsHARP	0.9480	0.9763	0.1505	0.9763	1
RK	1.0925	0.9681	0.1735	0.9680	0.9883
DSCK	0.9046	0.9785	0.1436	0.9785	0.9996
KED	0.9142	0.9778	0.1452	0.9778	0.9951
ATPRK	0.8468	0.9811	0.1345	0.9811	1

REFERENCES

Agam, N., Kustas, W. P., Anderson, M. C., Li, F. Q., Neale, C. M. U., 2007. A vegetation index based technique for spatial sharpening of thermal imagery. *Remote Sensing of Environment*, 107, pp. 545–558.

Atkinson, P. M., Pardo-Igúzquiza, E., Chica-Olmo, M., 2008. Downscaling cokriging for super-resolution mapping of continua in remotely sensed images. *IEEE Transactions on Geoscience and Remote Sensing*, 46, pp. 573–580.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*. Oxford Univ. Press, London, U.K.

Mukherjee, S., Joshi, P. K., Garg, R. D., 2015. Regression-kriging technique to downscale satellite-derived land surface temperature in heterogeneous agricultural landscape. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, 8(3), pp. 1245–1250.

Pardo-Igúzquiza, E., Chica-Olmo, M., & Atkinson, P. M., 2006. Downscaling cokriging for image sharpening. *Remote Sensing of Environment*, 102, pp. 86–98.

Rodriguez-Galianoa, V., Pardo-Igúzquiza, E., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2012. Downscaling Landsat 7 ETM+ thermal imagery using land surface temperature and NDVI images. *International Journal of Applied Earth Observation and Geoinformation*, 18, pp. 515–527.

Sales, M. H. R., Souza, Jr. C. M., & Kyriakidis, P. C., 2013. Fusion of MODIS images using kriging with external drift. *IEEE Transactions on Geoscience and Remote Sensing*, 51, pp. 2250–2259.

Wang, Q., Shi, W., Atkinson, P. M., Zhao, Y., 2015. Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, 166, pp. 191–204.

Wang, Q., Shi, W., Atkinson, P. M., 2016a. Area-to-point regression kriging for pan-sharpening. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, pp. 151–165.

Wang, Q., Shi, W., Atkinson, P. M., Pardo-Igúzquiza, E., 2016b. A new geostatistical solution to remote sensing image downscaling. *IEEE Transactions on Geoscience and Remote Sensing*, 54(1), pp. 386–396.

Wang, Z., Bovik, A. C., 2002. A universal image quality index. *IEEE Signal Processing Letters*, 9, pp. 81–84.