Remote Sensing of Environment 124 (2012) 49–60

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images

Xiaolin Zhu [a,b], Desheng Liu [b,c,*], Jin Chen [a]

[a] State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
[b] Department of Geography, The Ohio State University, Columbus, OH 43210, USA
[c] Department of Statistics, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 18 December 2011
Received in revised form 17 April 2012
Accepted 22 April 2012
Available online xxxx

Keywords:
Landsat ETM+
SLC-off
Gap filling
Geostatistical
Kriging

ABSTRACT

Since the failure of scan-line corrector (SLC) of the Landsat 7 Enhanced Thermal Mapper Plus (ETM+) sensor, a number of methods have been developed to fill the un-scanned gaps in ETM+ images. Unfortunately, the quality of the images filled by most of these existing methods is still not satisfactory, particularly in heterogeneous regions. Recently, a Neighborhood Similar Pixel Interpolator (NSPI) was developed that can accurately fill gaps in SLC-off images even in heterogeneous regions. However, the NSPI method is a type of deterministic interpolation approach that sets its weight parameters empirically and cannot provide statistical uncertainty of prediction. This study proposes a new gap-filling method called Geostatistical Neighborhood Similar Pixel Interpolator (GNSPI) by improving the NSPI method using geostatistical theory. The simulation study shows that: compared with previous geostatistical methods, the image filled by GNSPI has fewer striping effects; compared with NSPI, GNSPI is less empirical in its weight parameters and can provide uncertainty of prediction. More importantly, it can generate more accurate results than NSPI, especially when there is a long time interval between the input auxiliary image and the target SLC-off image.

1. Introduction

Since its first launch in 1972, Landsat satellite series has been continuously collecting images of the Earth from space (Goward et al., 2006). This longest archive of global mid-resolution, highly calibrated, multispectral data is a unique resource for applications at various scales, such as land use/cover mapping (Fuller et al., 1994), change detection (Byrne et al., 1980; Liu & Cai, 2011), ecosystem dynamics monitoring (Fisher et al., 2006; Masek et al., 2008), as well as biogeochemical parameter estimation (Cohen & Goward, 2004). Currently, the Thematic Mapper (TM) sensor on board Landsat 5 has substantially exceeded its planned design life (Bédard et al., 2008) and has unfortunately been halted since November 2011 (USGS, 2012). The Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor was the successor of TM. However, on May 31st, 2003, the scan-line corrector (SLC) of ETM+ failed permanently, which caused roughly 22% of the pixels to be un-scanned in any ETM+ images (referred to as SLC-off images) (Arvidson et al., 2006). More seriously, the width of these wedge-shaped data gaps increases the further east or west from the nadir path, resulting in large missing areas toward the edges of the scene, which leaves a major problem for the applications of ETM+ images. Therefore, filling gaps in SLC-off ETM+ images is very crucial and necessary to ensure the continuity of this longest satellite observation. In addition, even if Landsat 5 could be recovered before the launch of LDCM (Williams et al., 2006), the filled ETM+ images can be a valuable supplement of Landsat 5 TM to increase the availability of data in the frequently cloudy regions (Ju & Roy, 2008).

After the failure of SLC, many methods have been developed to fill the data gaps. These methods can be generally divided into two categories: deterministic interpolation and geostatistical interpolation. The difference between these two categories is that deterministic interpolation methods predict un-scanned pixels based on some empirically specified mathematical formulas, while geostatistical interpolation methods incorporate some kind of randomness in the prediction based on the statistical properties of the image. The first deterministic interpolation approach was proposed by a joint United States Geological Survey/National Aeronautics and Space Administration (USGS/NASA) Landsat team, referred to as local linear histogram-matching method (USGS, 2004). The linear histogram-matching method uses a moving window to get the localized linear transform function and then converts the pixel value of input auxiliary scenes to fill un-scanned gaps. This method requires that input scenes have minimal clouds, snow cover, low temporal variability and minimum date separation (USGS, 2004), which is not practical in some cloudy regions. As an alternative, MODIS images are used to predict the SLC-off ETM+ images by a semi-physical fusion approach (Roy et al., 2008). In addition, Maxwell (2004) introduced an object-based segmentation model to fill the gaps by using same-scene pixels outside the gaps, and this approach was further evaluated for land cover classification (Bédard et al., 2008). Then, this object-based segmentation approach was further improved by using multi-scale segmentation

* Corresponding author at: 1036 Derby Hall, 154 North Oval Mall, Columbus, OH 43210-1361, USA. Tel.: +1 614 247 2775; fax: +1 614 292 6213.
E-mail address: liu.738@osu.edu (D. Liu).

0034-4257/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2012.04.019

(Maxwell et al., 2007). Although the above methods are demonstrated to be effective to fill the missing pixels in SLC-off images in some case studies, they cannot ensure the accuracy of predictions in heterogeneous landscapes, and they have failed to predict small or narrow objects (Maxwell, 2004; Maxwell et al., 2007; Roy et al., 2008; USGS, 2004). To improve the accuracy of gap filling in heterogeneous landscape areas, Chen et al. (2011) recently proposed a Neighborhood Similar Pixel Interpolator (NSPI) approach to interpolate the gaps. NSPI makes use of the neighboring pixels with similar spectral characteristics to predict the value of un-scanned pixels, and it integrates spatial and temporal information of input images. Because of these two properties, NSPI can get more accurate results for small and narrow landscape features than the previous deterministic approaches (Chen et al., 2011). However, NSPI still has some limitations. As a deterministic approach, NSPI cannot produce the uncertainty of predictions, and the impact of changes between input and target images on the final results is not clear (Chen et al., 2011). In addition, the weight parameters in NSPI are empirically determined by spatial and spectral distances, which may affect their robustness in application for various situations.

For geostatistical methods, ordinary kriging and co-kriging techniques were employed to fill the data gaps (Pringle et al., 2009; Zhang et al., 2007). These geostatistical methods can provide the uncertainty of predictions, which is a significant difference with the deterministic methods. The fundamental idea of geostatistics is Tobler's first law of geography, i.e., near spatial data values are more related to each other than distant data values (Tobler, 1970). Assuming that pixel values in satellite images are geostatistical data generated from an intrinsically stationary process (i.e., a spatial random process satisfying two conditions: 1) a constant mean and 2) the variance of the differences at any two locations only depends on the displacement between the two locations), ordinary kriging uses the target SLC-off image itself to predict its un-scanned pixels and co-kriging adds the secondary image into prediction (Pringle et al., 2009; Zhang et al., 2007). Case studies demonstrated that the geostatistical approaches improved the accuracy of filled results to some extent compared with the USGS's local linear histogram-matching method (Pringle et al., 2009; Zhang et al., 2007). However, there are still some visual striping patterns in the filled results caused by smoothing effects. Specifically, linear features are affected seriously, such as where roads are filled as other landscapes, and where rivers are filled as land surface (Pringle et al., 2009; Zhang et al., 2007). In other words, these geostatistical methods cannot fill the SLC-off images well at the pixel-level, which limits the quantitative applications of these filled images. The reason for this weakness of geostatistical methods may be that the intrinsic stationarity assumption is not always valid especially for heterogeneous landscapes (Pringle et al., 2009).

Considering that NSPI and geostatistical interpolation methods have their respective merits, this study develops a new geostatistical approach for filling gaps in SLC-off ETM+ images by integrating the theory of geostatistics with the idea of the NSPI method. We refer to this new approach as Geostatistical Neighborhood Similar Pixel Interpolator (GNSPI). It is expected that GNSPI will interpolate the gaps more accurately than other methods. The steps of GNSPI will be described first and then a simulation study will be implemented to evaluate the performance of GNSPI.

2. Methodology

2.1. Data for algorithm test

In order to compare the proposed GNSPI with the NSPI method, this study used the same data set in the NSPI paper (Chen et al., 2011). The 15 km × 15 km (500 × 500 Landsat pixels) area is located in eastern Maryland around 39.10°N and 76.14°W. Fig. 1(a) shows the Landsat 5 TM image acquired on April 29, 2010, which was further simulated as SLC-off image with un-scanned gaps (Fig. 1(b)). These simulated gaps were generated using the gap mask from an actual SLC-off ETM+ image so that they represent the real situation in practice. The simulated gaps are composited by 250,000 un-scanned pixels. This simulated SLC-off image (Fig. 1(b)) is used as the target SLC-off image, which will be filled using the auxiliary images. To keep the consistency with NSPI, only the results of green, red and near infrared (NIR) bands (bands 2, 3 and 4) were reported to evaluate the performance of the proposed method. Fig. 2 shows all the auxiliary images in this study, including Landsat 5 TM and Landsat 7 ETM+ images. TM auxiliary images are acquired on February 8, 2010 (Fig. 2(a)), May 15, 2010 (Fig. 2(b)), and September 4, 2010 (Fig. 2(c)), respectively. ETM+ auxiliary images are acquired on March 20, 2010 (Fig. 2(d)), April 5, 2010 (Fig. 2(e)), and June 8, 2010 (Fig. 2(f)), respectively. From Figs. 1 and 2, we can see that it is a very heterogeneous area with relatively small patches.

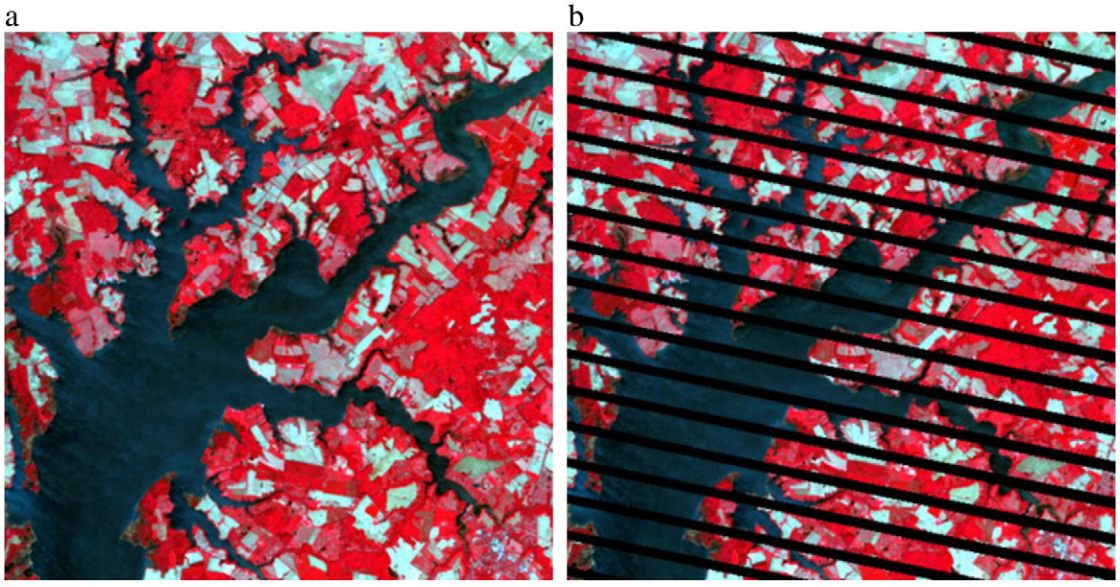

Fig. 1. NIR-red–green composites of Landsat TM images for the simulation study: (a) TM image acquired on April 29, 2010; (b) simulated SLC-off image based on (a). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

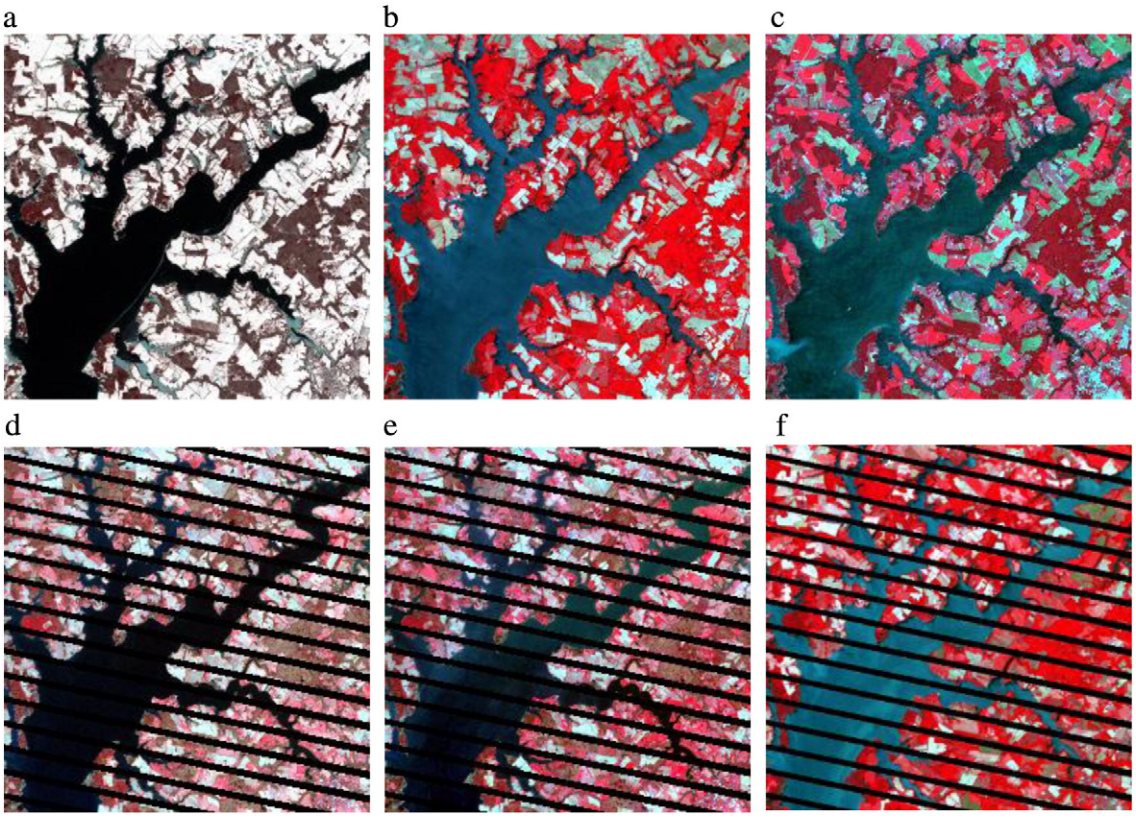

Fig. 2. The auxiliary images used in the case study: (a)–(c) TM images acquired on February 8, May 15 and September 4, 2010, respectively; (d)–(f) ETM+ images acquired on March 20, April 5, and June 8, 2010, respectively.

In particular, the farm land in the area experiences rapid reflectance changes. This area is an ideal region for testing the gap filling algorithms, since most existing methods have some difficulties in accurately filling gaps for areas with a heterogeneous landscape and rapid changes (Maxwell, 2004; Maxwell et al., 2007; Pringle et al., 2009; Roy et al., 2008; USGS, 2004; Zhang et al., 2007).

2.2. Algorithm development

Like NSPI, GNSPI also needs Landsat images acquired at other dates as input images to fill the target SLC-off image (Chen et al., 2011). The input images should have values at the corresponding locations of gaps in the target image. Besides Landsat 5 TM images, Landsat 7 SLC-off ETM+ images can also be used as input images because the gaps are not completely overlapped among scenes acquired on different dates. Landsat 5 TM images will be primarily considered as input images if they are available. However, the Landsat 7 ETM+ images are also very important input data for GNSPI if Landsat 5 ceases to function. Considering that there are some differences for the implementation of GNSPI when using these two types of input data, the steps of GNSPI using TM input images will be introduced first and then the different steps for using ETM+ input images will be clarified.

2.2.1. Using TM images as input data

Fig. 3 shows the flowchart of GNSPI using TM images as input data. The criterion of input image selection is that the image is the most similar to the target image in seasonality, has smallest spectral changes compared with the target image, and should be cloud-free. Generally, the TM input image acquired closer in time to the target image has a higher similarity to the target image. For the convenience of this paper, the acquisition date of this input image is denoted as t_1, while the date of the target SLC-off image is denoted as t_2. Here, the TM image acquired on May 15 (Fig. 2(b)) is chosen as the input image because it is temporally close to the target image acquired on April 29 (Fig. 1(b)). The input image must be geometrically rectified to match the target image. To eliminate the different illumination conditions, both the target SLC-off image and input image are calibrated to top-of-atmosphere reflectance with a range from 0 to 1. It is not necessary to do atmospheric correction and relative radiometric normalization for these images because the effects of

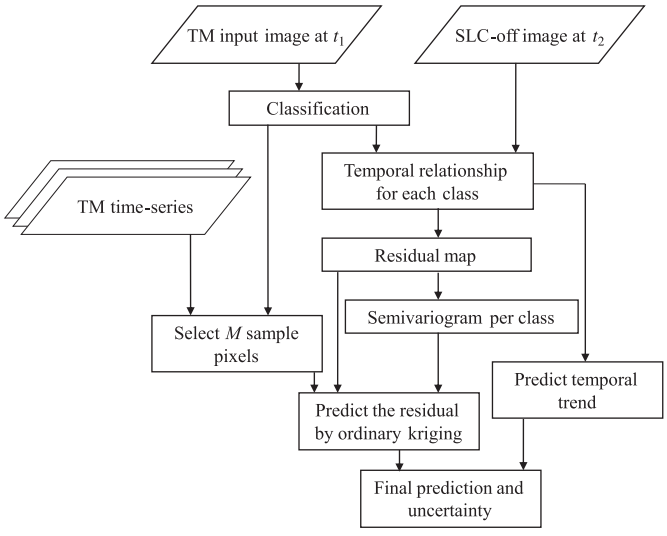

Fig. 3. The flowchart of GNSPI using TM input images.

atmosphere on the images and the small difference of the spectral properties between TM and ETM+ sensors (Teillet et al., 2001) will be taken account into the GNSPI algorithm. However, it should be noted that the difference in illumination conditions between images may not be completely eliminated for mountainous areas even after the radiometric calibration. In this case, the topographic correction process (Riano et al., 2003) is recommended before implementing gap filling.

As Fig. 3 shows, there are 6 main steps for the proposed GNSPI method. These steps correspond to the strategy of kriging technique in practice (Papritz & Stein, 1999). Specifically, the data need to be detrended first using physical or empirical models (steps 1 and 2); second, the semivariogram of the residuals from the detrending model is estimated (step 3); third, the residual of the location without observation is predicted using ordinary kriging (steps 4 and 5); last, the final prediction of the location without observation is obtained by adding back the trend (step 6). All steps of GNSPI will be discussed in detail below.

2.2.1.1. Step 1: classifying the input image at t_1. As in NSPI, it is reasonable to assume that pixels with high spectral similarity should have similar temporal changing pattern (Chen et al., 2011). In order to model the temporal changing trend for each group of spectrally similar pixels in the next step, the input image at t_1 is classified based on the spectral similarity of pixels. Considering that there would be no significant land cover changes during the short time interval between the input image and the target image, the classification map of input image at t_1 can also represent the situation at t_2. Here, to ensure the automation of GNSPI algorithm, we use an unsupervised classifier, ISODATA, to classify the input image at t_1. ISODATA can automatically merge and split classes according to the spectral similarity between pixels to get the optimal classification results (Ball & Hall, 1965), and has been widely used for remotely sensed image classification (Call et al., 2003; Tarabalka et al., 2009). Fig. 4 shows the classification result of TM image of May 15, which includes 3 spectral classes. Based on visual interpretation, the three spectral classes generally correspond to water, vegetation and bare ground.

2.2.1.2. Step 2: modeling temporal relationship for each class. From t_1 to t_2, the reflectance value would experience changes. These changes can be caused by different atmospheric situations and the intrinsic difference of spectral properties between two sensors, but the most significant changes would be caused by land surface dynamics, such as the phenological changes of vegetation and variations in soil moisture. It should be noted that different classes might have different temporal changing patterns. For example, the spectral characteristics of vegetation change dramatically during the growing season compared with that of bare ground. Therefore, the temporal relationship between the input image at t_1 and the target image at t_2 is modeled for each class, respectively by using the classification map from the last step. This temporal relationship can be used to detrend the target image at t_2. A linear function is used to model the temporal relationship for each class (Zhu et al., 2010):

$$R(s,t_2,b) = a_k \times R(s,t_1,b) + b_k, \text{ if pixel } s \text{ belongs to class } k \quad (1)$$

where $R(s, t_1, b)$ and $R(s, t_2, b)$ are reflectance values of pixels at location s in band b for images at t_1 and t_2, respectively; a_k and b_k are coefficients of this linear model for class k. The coefficients can be estimated by linearly regressing the reflectance values of all the same-class pixels from the image at t_1 and the image at t_2, and the locations of these pixels should be outside the gaps of the target image. The regressed parameters for the temporal relationship of each band and each class is listed in Table 1. It is clear that different classes have different temporal relationships, indicating the necessity to model the temporal relationship for each class rather than for the whole image.

Then, the temporal relationship for each class is applied to predict the reflectance values of all the pixels in the target SLC-off image, including the pixels outside or inside the gaps:

$$\hat{R}(s,t_2,b) = \hat{a}_k \times R(s,t_1,b) + \hat{b}_k, \text{ if pixel } s \text{ belongs to class } k \quad (2)$$

where $\hat{a}_k$ and $\hat{b}_k$ are coefficients estimated by linear regression, $\hat{R}(s,t_2,b)$ is the prediction of reflectance for pixel s at t_2. For the pixel s outside the gaps, we can calculate its prediction residual, $\varepsilon(s, b)$, because it has observed values at t_2:

$$\varepsilon(s,b) = R(s,t_2,b) - \hat{R}(s,t_2,b), \text{ if pixel } s \text{ is outside the gaps.} \quad (3)$$

2.2.1.3. Step 3: modeling semivariogram for each class. After the residuals of all pixels outside the gaps are calculated by Eq. (3), we can get a residual image as shown in Fig. 5. From Fig. 5, it is clear that the residual is spatially continuous, especially for the pixels of the same class. It is assumed that the residual of each class is an intrinsically stationary spatial process, i.e., the differences of residuals between two same-class pixels separated by a given distance have a constant mean and a constant variance. Then a semivariogram can be used to model the spatial dependence of residuals for each class (Matheron, 1971). For each class, 1000 pixels outside the gaps in the residual image are randomly selected as samples for estimating the experimental semivariogram. Here, the lag distance is limited from 0 to 40 pixels, because the pixels finally used to predict the un-scanned pixel are located within this distance. For each spatial

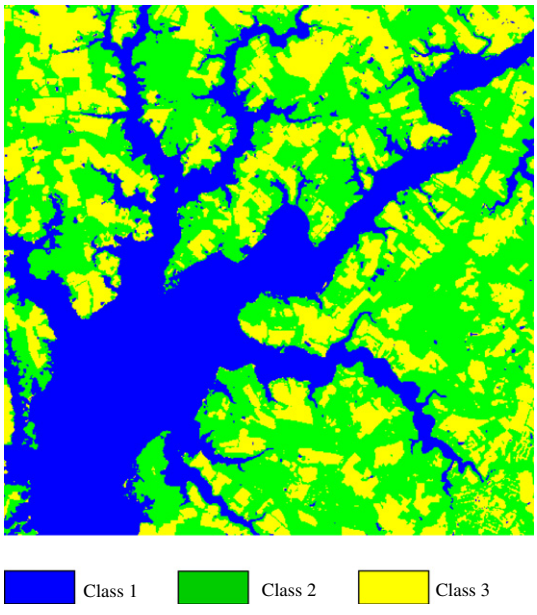

Fig. 4. Classification map of TM image acquired on May 15, 2010.

Table 1
The regressed parameters for the temporal relationship of each band and each class*.

Band	Class 1		Class 2		Class 3	
	a	*b*	*a*	*b*	*a*	*b*
Green	0.3185	0.0389	0.4486	0.0457	0.6679	0.0161
Red	0.3256	0.0289	0.5077	0.0369	0.7758	0.0040
NIR	0.8419	−0.0145	0.5497	0.0835	0.8866	−0.0130

* P-values for all the regression model are less than 0.0001.

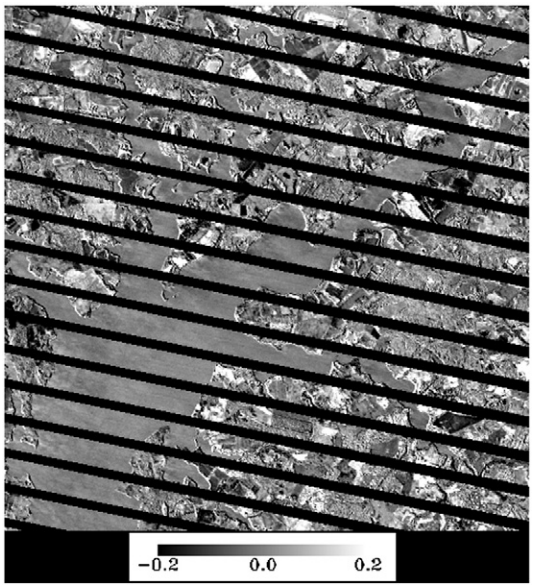

Fig. 5. The residuals of prediction from the temporal relationship for NIR band.

lag h, i.e., the separation distance between two locations, the experimental semivariogram $\gamma(h)$ can be calculated as:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [\varepsilon(s_i, b) - \varepsilon(s_i + h, b)]^2, \quad (4)$$

where $N(h)$ is the number of pairs of samples apart with each other at the distance h. Fig. 6 plots the experimental semivariograms versus spatial lags of green band for class 1. It shows a gradually increasing trend, indicating that residuals of pixels spatially close together are more alike than those farther apart, and the semivariograms would be more stable when the lags reach a certain distance.

In order to use the characteristics of spatial dependence in geostatistical techniques, the experimental semivariograms need to be fitted by a mathematical model, subject to the constraint of positive definiteness. An exponential model is frequently used to fit the experimental semivariograms (Goovaerts, 1997). The parametric function of exponential model is:

$$\gamma(h) = \begin{cases} a + \left(\sigma^2 - a\right)\left(1 - e^{-3h/r}\right), & \text{for} \quad h > 0 \\ 0, & \text{for} \quad h = 0 \end{cases}, \quad (5)$$

where a, σ^2, and r are parameters which represent nugget effect, sill, and range, respectively. These parameters can be estimated through

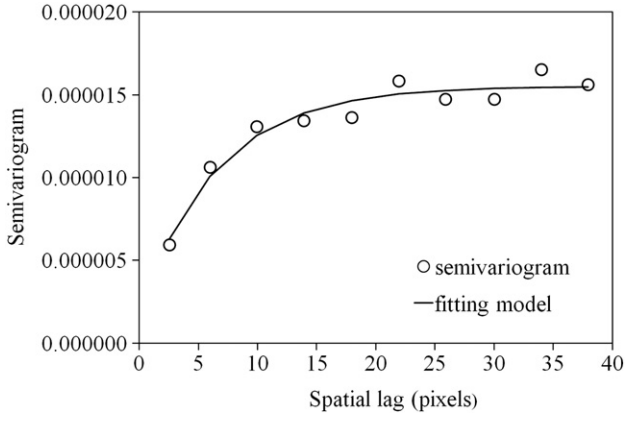

Fig. 6. The experimental semivariogram and the fitted model for green band of class 1.

fitting the experimental semivariograms by a nonlinear weighted least squares method (Cressie, 1985). For the example shown in Fig. 6, the parameters a, σ^2, and r of the fitted model are 1.84×10^{-6}, 1.55×10^{-5}, and 19.51, respectively. The fitted curve is very close to the experimental semivariograms, indicating the effectiveness of the parametric model and fitting method. Table 2 summarizes the parameters of the fitted semivariograms for each band and each class. It shows that the residuals of different classes have different spatial dependence characteristics, so it is reasonable and necessary to model the semivariograms for each class, respectively.

2.2.1.4. Step 4: selection of sample pixels. As in NSPI, GNSPI assumes that spatially adjacent pixels with high spectral similarity (referred to as similar pixels) have similar temporal changing patterns. Eq. (1) describes the temporal change for each class in the whole scene. However, the temporal change has some spatial variability even for pixels from the same class. For example, farm land planted with the same crop would show some different spectral changes due to different fertilizing and irrigating strategies at different locations. Therefore, when the global relationship Eq. (1) is used to predict the reflectance value at t_2, it will cause some local errors, which is calculated by Eq. (3). Based on the assumption that neighboring-similar pixels have similar temporal changing patterns, the residuals of these similar pixels should also show the similarity. For an un-scanned pixel (target pixel), its trend can be first estimated by the global temporal relationship in Eq. (1), but the residual of this prediction is unknown. In contrast, the similar pixels of this target pixel outside the gaps have residual values. Accordingly, it is reasonable to use the residuals of these similar pixels to predict the residual of the un-scanned target pixel. Therefore, selecting sample pixels for prediction is a key point in this approach.

Fig. 7 shows the schematic diagram for selecting sample pixels. The selection process can be implemented using a TM time-series. The number of images included in this time-series depends on the data availability. The strategy for selecting these images is that they are temporally separated in one year so that they can capture the temporal changing characteristics as much as possible. We recommend one image per season if it is possible to compose this time-series, which could accurately identify pixels with the same temporally changing pattern. This time-series includes the input image at t_1. Except the input image at t_1, all other auxiliary images in this time-series are only used to select sample pixels and are not involved in predicting the value of the target pixel. If there are no other available cloud-free images except the input image, the selection process of sample pixels can be implemented only using the input image at t_1. In this study, the time-series includes 3 TM images acquired on February 8, 2010, May 15, 2010, and September 4, 2010 (Fig. 2(a)–(c)).

For convenience, an example by using time-series of two images is used to explain the selecting procedure. These two images are ordered as the 1st image and the 2nd image based on their temporal closeness to the target SLC-off image. These two images include the

Table 2
The parameters of the fitted semivariogram models for each band and each class.

	Band	Sill	Nugget	Range
Class 1	Green	1.55×10^{-5}	1.84×10^{-6}	19.51
	Red	3.13×10^{-5}	0	52.76
	NIR	9.94×10^{-5}	5.06×10^{-5}	14.16
Class 2	Green	1.39×10^{-4}	4.44×10^{-5}	31.70
	Red	4.07×10^{-4}	8.90×10^{-5}	37.49
	NIR	7.12×10^{-4}	7.94×10^{-5}	24.25
Class 3	Green	3.00×10^{-5}	1.51×10^{-5}	17.42
	Red	7.33×10^{-5}	0	43.28
	NIR	7.09×10^{-4}	2.70×10^{-4}	17.83

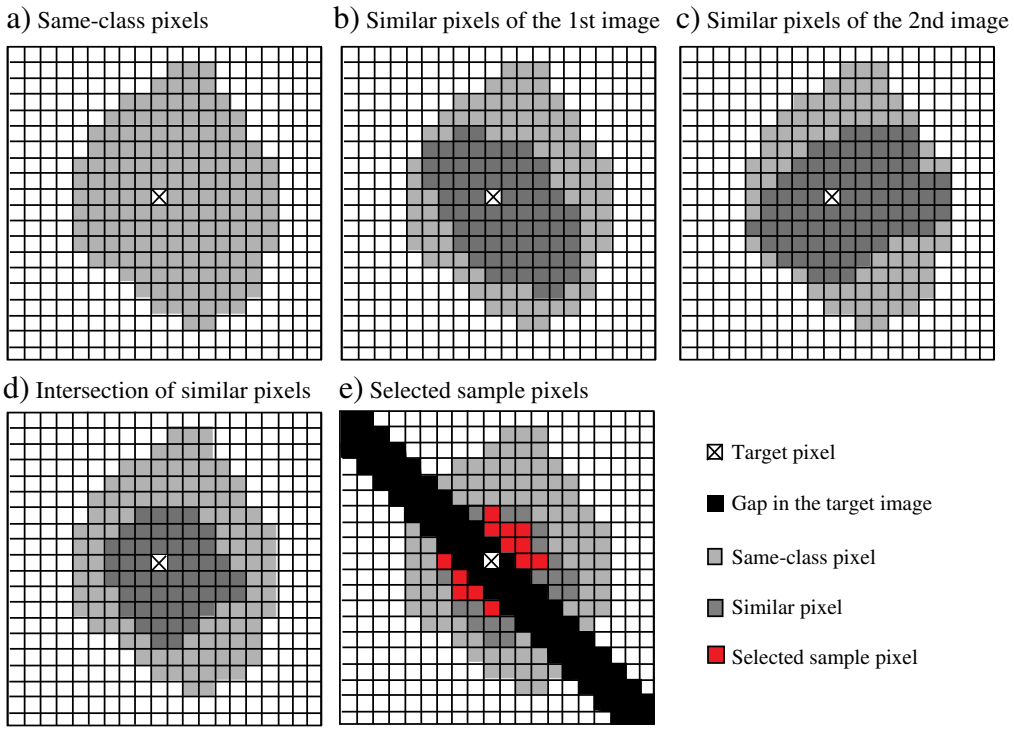

Fig. 7. Schematic diagram of the selection of sample pixels when using TM images.

one selected as t_1 input image which has been classified in step 1. First, from the classification map (Fig. 4), all the pixels in a moving window with the same class as the central target unscanned pixel are selected (the light gray pixels in Fig. 7(a)). Second, because the same-class pixels would also show some locally varying characteristics, it is necessary to refine the same-class pixels to ensure that the selected pixels are more similar to the target pixel. From the 1st image, similar pixels are selected from the same-class pixels according to spectral similarity. The spectral similarity is defined as root mean square difference (RMSD) between each same-class pixel s_i and the target unscanned pixel s_t (Chen et al., 2011):

$$\mathrm{RMSD}_i = \sqrt{\frac{\sum_{b=1}^{n}(R(s_i,1\mathrm{st},b)-R(s_t,1\mathrm{st},b))^2}{n}} \quad (6)$$

where n is the number of bands. A large RMSD represents a large spectral difference. Then, a threshold approach is used to identify similar pixels (Chen et al., 2011). If RMSD of the s_i same-class pixel satisfies Eq. (7), it will be selected as a similar pixel:

$$\mathrm{RMSD}_i \le \left[\sum_{b=1}^{n}\sigma(b,1\mathrm{st})\times 2/m\right]/n \quad (7)$$

where $\sigma(b,1\mathrm{st})$ is the standard deviation of the 1st image for band b, and m is the number of classes in the classification map (Fig. 4). The similar pixels selected from the 1st image are shown in Fig. 7(b). Using the same procedure, similar pixels of the 2nd image are also selected as shown in Fig. 7(c). It is possible that the locations of similar pixels selected from the two images might be different because the spectral similarity between pixels would change over time due to the difference of temporal changing patterns. For example, a pixel of vegetation before green-up would show similar spectral characteristics as a pixel of bare ground but will be very different from the bare ground after it turns to green. If more images are included in the time-series, similar pixels will be selected from these images, respectively using the same approach above. Third, the intersection of the similar pixels from each image is selected (Fig. 7(d)), because the intersection from different dates can provide a more accurate selection of pixels with the same temporal changing pattern as the target pixel. Fourth, there may be a large number of pixels located in the intersection, but a large number of sample pixels used to predict the target pixel will decrease the computing efficiency. Therefore, we only select M sample pixels from the intersection for prediction. Considering that spatially closer pixels may be more similar than spatially distant ones (Zhu et al., 2010; Chen et al., 2011), M pixels outside the gaps of target SLC-off image, which are spatially closest to the target pixel, are selected from the intersection area as shown in Fig. 7(e). This study sets M as 20, which is recommended by NSPI (Chen et al., 2011) and by geostatistical interpolation (Webster & Oliver, 2001). If the pixels in the intersection were fewer than the sample size M, all the pixels in the intersection will be used as sample pixels, regardless of the total number of pixels in the intersection. The size of the moving window depends on the homogeneity of the image. If the area is homogeneous, it is easy to select enough sample pixels within a smaller moving window. Otherwise, a larger moving window is needed. In our case study, the size of the moving window is set as 25×25 pixels, by which 99% of un-scanned pixels can have enough sample pixels for prediction.

2.2.1.5. Step 5: predict the residual of the target pixel by ordinary kriging. As explained above, the residual of the prediction by the temporal relationship (Eq. 2) is unknown for the un-scanned target pixel, while the sample pixels have known residual values. In addition, the residuals show spatial dependence (Fig. 5). Therefore, it is reasonable to use these sample pixels to predict the residual of the target pixel by the geostatistical method. Under the assumption of intrinsic stationarity for the residual image, ordinary kriging predictor is used to estimate the residual value of the target pixel. The prediction

of the target pixel $\varepsilon^*(s_t, b)$ by ordinary kriging is a weighted average of the residual values of M sample pixels:

$$\varepsilon^*(s_t, b) = \sum_{i=1}^{M} w_i \varepsilon(s_i, b), \tag{8}$$

where w_i is the ordinary kriging weight of the ith sample pixel. The weights are determined based on the locations of sample pixels using the theoretical semivariogram obtained from step 3 and two statistical optimality criteria: unbiasedness and minimum mean-squared prediction error. For unbiasedness, the kriging weights are forced to sum to 1. For the second criterion, the Lagrange multiplier method is used to solve this constrained optimization problem to get the weights. In matrix notation, weights can be calculated as:

$$\mathbf{w} = \mathbf{C}^{-1}\mathbf{c}, \tag{9}$$

where,

$$\mathbf{w}^{\mathrm{T}} = [w_1, \cdots, w_M, v],$$
$$\mathbf{C} = \begin{bmatrix} C(s_1,s_1) & \cdots & C(s_1,s_M) & 1 \\ \vdots & \cdots & \vdots & \vdots \\ C(s_M,s_1) & \ddots & C(s_M,s_M) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix},$$
$$\mathbf{c}^{\mathrm{T}} = [C(s_1,s_t), \cdots, C(s_M,s_t), 1],$$

where, v is the Lagrange multiplier, and $C(s_i, s_j)$ represents the covariance between the pixel at location s_i and the pixel at location s_j. The covariance can be computed based on the fitted exponential semivariogram model:

$$C\left(s_i, s_j\right) = \sigma^2 - \gamma\left(\left|s_i - s_j\right|\right), \tag{10}$$

where $|s_i - s_j|$ is the spatial distance between the pixel at location s_i and the pixel at location s_j. After the weights are calculated, the residual of the target pixel can be predicted by Eq. (8), and the variance of the ordinary kriging prediction is:

$$\sigma_{ok}{}^2(s_t, b) = \sigma^2 - \mathbf{c}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}. \tag{11}$$

2.2.1.6. Step 6: calculate the final prediction and uncertainty. The temporal trend of the target pixel $\hat{R}(s_t, t_2, b)$ can be predicted by Eq. (2) using the observation at t_1. The reflectance value of the target pixel $R^*(s_t, t_2, b)$ can be finally predicted by adding the residual from ordinary kringing estimator to the temporal trend:

$$R^*(s_t, t_2, b) = \hat{R}(s_t, t_2, b) + \varepsilon^*(s_t, b). \tag{12}$$

From the variance of ordinary kriging, the half-interval of 95% confidence for the final prediction is:

$$I(s_t, b) = 1.96 \times \sqrt{\sigma_{ok}{}^2(s_t, b)}. \tag{13}$$

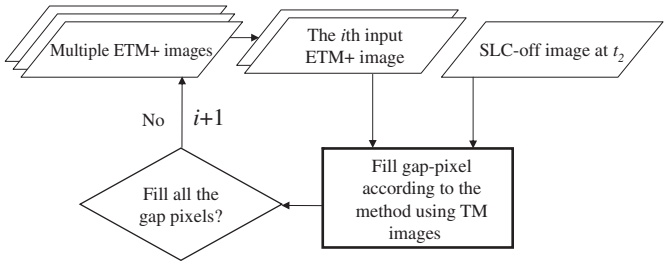

Fig. 8. The flowchart of GNSPI using ETM+ input images.

The relative half-interval is used to represent the uncertainty of the final prediction:

$$U(s_t, b) = 100\% \times I(s_t, b)/R^*(s_t, t_2, b). \tag{14}$$

2.2.2. Using ETM+ images as input data

Fig. 8 shows the flowchart of GNSPI when ETM+ images are used as input data. GNSPI requires that the pixel at the location of the un-scanned target pixel has a value in the input image. Fortunately, the gaps in the ETM+ images do not overlap completely among scenes. Therefore, it is possible to use multiple ETM+ images to fill the gaps by GNSPI. First, we need to select multiple ETM+ images which cover the same area of the target image. The number of input images depends on the degree of gap overlap. In this study, three ETM+ images acquired on March 20, 2010 (Fig. 2(d)), April 5, 2010 (Fig. 2(e)), and June 8, 2010 (Fig. 2(f)) are used as input data for filling gaps of the simulated SLC-off image (Fig. 1(b)). Second, all the input ETM+ images are sorted by acquisition date. From these ETM+ images, the ETM+ image temporally closest to the target image will be preferentially used as input data (Chen et al., 2011). Third, if the un-scanned pixels in the target image are scanned in the first ETM+ input image, they will be first filled according to the method described in Section 2.2.1. Then, the second temporally closest ETM+ image will be used to fill the rest of the un-scanned pixels until all the un-scanned pixels are filled. The only different step of using ETM+ as input images is the selection of sample pixels.

Since the input ETM+ images also have gaps, to guarantee that enough sample pixels outside the gaps can be selected, the size of moving window is enlarged to 31×31 pixels in our case study. Fig. 9 shows the process of selecting sample pixels when using ETM + images as input data. Here, we also use a time-series composed of two ETM+ images as an example to explain the selection process. First, the images are ordered by the temporal closeness to the target image. The temporally closest image will be selected as the input image at t_1, which will be classified to get the classification map. From this map, the pixels outside the gaps with the same class as the target pixel are selected (Fig. 9(a)). Second, similar pixels in the 1st and 2nd images are selected by the threshold method (Chen et al., 2011) (Fig. 9(b) and (c)). Third, the intersection of similar pixels selected from different images is identified. It should be noted that both images in the time-series have gaps and the gaps are partly overlapped. If we only select the intersection of similar pixels outside the gaps in each image, the number of pixels in the intersection would be very limited. Therefore, a specific approach is applied to find out this intersection area: besides the similar pixels simultaneously selected from all the images, the similar pixel selected from the 1st image, which is located within gaps of the 2nd image, is also kept in the intersection because we cannot determine whether it is or not a similar pixel in the 2nd image. Last, M spatially closest similar pixels are selected as final sample pixels for predicting the value of the target pixel (Fig. 9(e)).

3. Testing experiment

3.1. Comparison of prediction accuracy

The simulated SLC-off image (Fig. 1(b)) was used to test the proposed GNSPI method. The images in Fig. 2 were used as auxiliary images to fill the simulated gaps. The actual image of the simulated SLC-off image (Fig. 1(a)) can help us to assess the prediction accuracy quantitatively.

For the case of using TM images as input images, the one temporally closest to the simulated SLC-off image, i.e., the image acquired on May 15 (Fig. 2(b)), will be selected as an input image at t_1. However, considering the difficulty of collecting temporally close

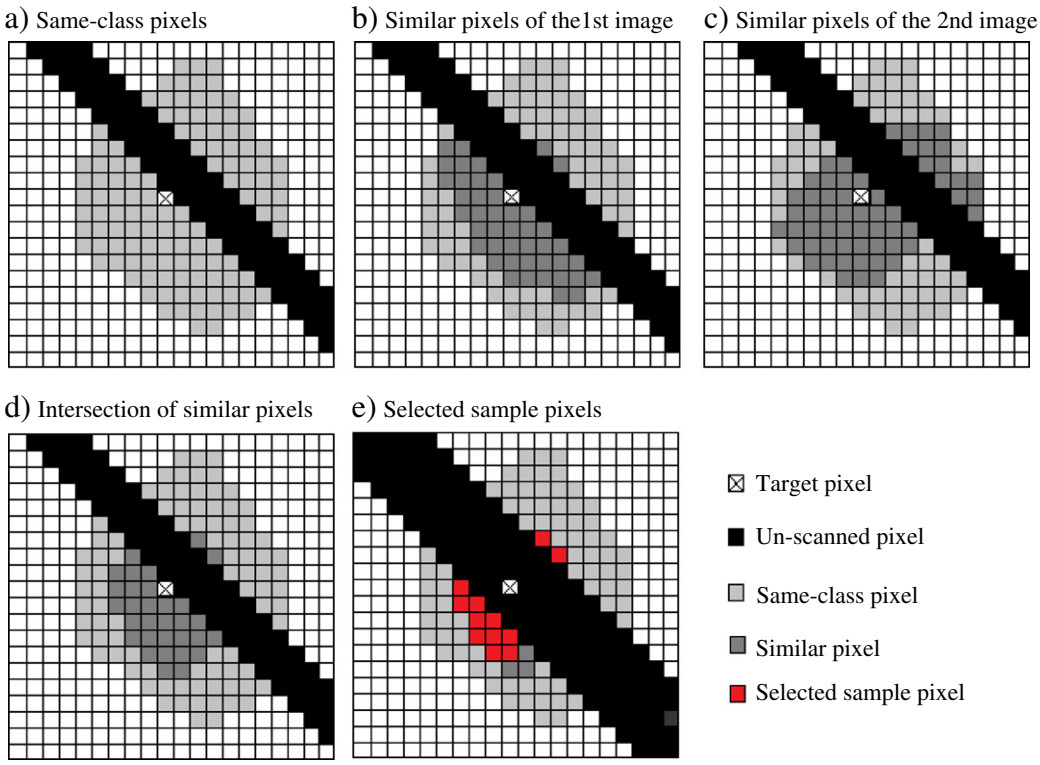

Fig. 9. Schematic diagram of the selection of sample pixels when using ETM+ input images.

images in frequently cloudy regions (Ju & Roy, 2008), it is necessary to investigate the performance of GNSPI when using a temporally further input image. Therefore, the TM image acquired on February 8 (Fig. 2(a)), which has significant spectral changes with the target SLC-off image, was also selected as an input image for filling gaps in the simulated SLC-off image. Accordingly, we can obtain two filled images using Fig. 2(a) and (b) as input images, respectively.

For the case of using ETM+ images as input images, three ETM+ images in Fig. 2 were used to fill the gaps in the simulated SLC-off image. Considering the degree of gap overlap, these three ETM+ images were enough for filling all the un-scanned pixels in the simulated SLC-off image.

To better evaluate the strength and limitations of GNSPI, we compared GNSPI with other two methods: NSPI developed by Chen et al. (2011) and previous geostatistical methods developed by Zhang et al. (2007). For NSPI, input images are also needed, and either TM or ETM+ images can be used as input images. Therefore, NSPI used the same input images as GNSPI to fill the simulated SLC-off image and got three filled results: two by using temporally closer and further TM input images, respectively and one by using multiple ETM+ input images. For previous geostatistical methods, ordinary kriging technique does not need any other input images to fill the data gaps while co-kriging technique incorporates the secondary image to fill the data gaps (Zhang, et al., 2007). However, ordinary kriging technique was shown to be comparable with co-kriging (Zhang et al., 2007). Therefore, we only selected ordinary kriging technique as previous geostatistical methods for comparison, which is referred to as PGM in the following. PGM does not require other input images, so only one filled result by PGM for the simulated SLC-off image can be obtained. In addition, as NSPI has been demonstrated to have better performance than some of the existing methods (Chen et al., 2011), such as USGS's local linear histogram matching method, we did not include these existing methods for comparison.

The accuracy of the above filled results will be quantitatively assessed by comparing them with the actual image. Root mean square error (RMSE) is a commonly used index for evaluating the accuracy (Chen et al., 2011), which can be calculated as:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} [R^*(s_i, t_2, b) - R_0(s_i, t_2, b)]^2}{N}}, \quad (15)$$

where N is the total number of un-scanned pixels, $R^*(s_i, t_2, b)$ and $R_0(s_i, t_2, b)$ are predicted and actual values of the ith un-scanned pixel, respectively. A smaller RMSE indicates a more accurate prediction.

3.2. Comparison of filled image for classification

The filled SLC-off image will be used for various applications. Land cover classification is one important application of Landsat images (Fuller et al., 1994). To investigate the effectiveness of filled Landsat SLC-off images in the application community, all the filled images in this study were classified to get the land cover maps, and then these land cover maps were compared with the one generated from the actual image. In detail, we use a supervised classifier, Support Vector Machines (SVM), to classify these filled images and the actual image, because SVM has been widely used for remotely sensed image classification and found to be competitive with the other best classifiers (Huang et al., 2002). Training samples for each land cover type were selected from the actual image by visual interpretation. There are four land cover classes, i.e. water, natural vegetation, crops and bare ground, and the size of training samples for each class is 634, 775, 765, and 601 pixels, respectively. Then, all the images were classified by SVM method using the same set of training samples, which can eliminate the effects of different training samples on the final classification result.

The error matrix method was applied to quantitatively assess the accuracy of classification (Congalton, 1991). Because we focus on the classification of gap filling result, only the pixels within the gaps were taken into account for building the error matrix. The classification result from the actual image was treated as reference data. We admit that

classification of the actual image is not exactly the ground truth, but it can be used to evaluate whether the filled images can get similar classification results as the actual image when they use the same classifier and the same training samples, which can explain the effectiveness of these filled images in the classification application. Two commonly used statistical indices, overall accuracy (*oa*) and kappa coefficient (*kappa*) calculated from the error matrix (Congalton, 1991), were reported to evaluate the classification accuracy. *oa* is one of the most popular indices used to assess the classification accuracy. It is computed by summing the correct proportions in the error matrix. *kappa* is another popular index calculated as the proportion of agreement between classification result and reference data after chance agreement is removed (Congalton, 1991). Higher values of *oa* and *kappa* indicate better classification results.

4. Results

4.1. Filled results of the simulated SLC-off image

Fig. 10 shows all the filled results of the simulated SLC-off image by different methods and different input images. For simplification, the result filled by NSPI using the input TM image of February 8 is

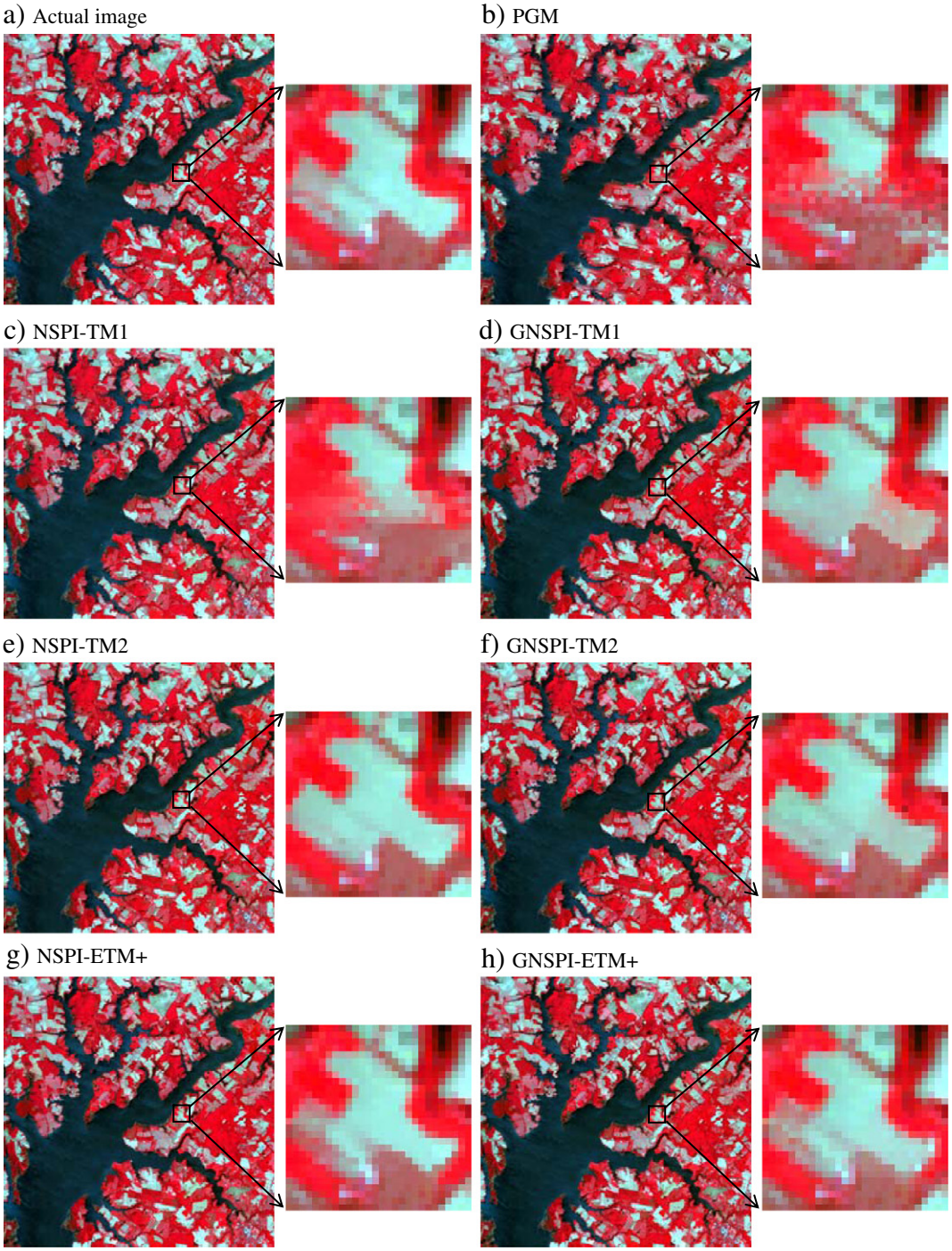

Fig. 10. Results of gap-fill for the simulated SLC-off image. (a) the actual image; (b) the filled image by PGM; (c), (e) and (g) are the filled images by NSPI using TM image of February 8, TM image of May 15 and ETM+ images as input images, respectively; (d), (f) and (h) are the filled images by GNSPI using TM image of February 8, TM image of May 15 and ETM+ images as input images, respectively; the smaller images on the right side of each filled image are the zoomed-in regions marked by a black square.

referred to as NSPI-TM1, while the result using the TM input image of May 15 is referred to as NSPI-TM2 and the result using ETM+ is NSPI-ETM+. Correspondingly, the result filled by GNSPI using the TM image of February 8 as input image is referred to as GNSPI-TM1, while the result using the TM image on May 15 as input image is referred to as GNSPI-TM2 and the result using ETM+ is GNSPI-ETM+. By visual comparison of all filled images shown in Fig. 10, it is clear that the image filled by the previous geostatistical method (PGM) shows the most obvious striping effects, indicating that there are large prediction errors by PGM. For the comparison between NSPI and GNSPI, the images filled by both methods are generally visually close to the actual image. However, from the zoomed area of these filled images, we can see that GNSPI fills the gaps more accurately than NSPI when both methods use a temporally further TM input image. To better assess the similarity between the filled results and the actual image, Fig. 11 shows the scatter plots of all the filled results and the actual image for NIR band. From Fig. 11, we can see that data points of all results are almost equally distributed at the two sides of 1:1 line, indicating small biases of all results. However, the data points in the scatter plot of the NSPI and GNSPI predictions fall closer to the 1:1 line than the PGM prediction. The R^2 value of the PGM prediction is 0.7100, which is the lowest one among all the filled results. For comparison between NSPI and GNSPI, their scatter plots and R^2 values are very close except when using a temporally further TM image as the input image. We can see that when using TM image of February 8 as input image, GNSPI can get a higher R^2 value 0.9368 than NSPI's 0.8481, indicating that the image filled by GNSPI is closer to the actual image than NSPI.

Table 3 reports the RMSE values for all the gap-filling results. Generally, the accuracy of the PGM prediction is the lowest, because the RMSE values of all the three bands are the highest compared with NSPI and GNSPI. For both NSPI and GNSPI, the results filled by temporally closer input images can get higher accuracy. By using the same input image, GNSPI can get smaller RMSE values than NSPI, especially when using a temporally further input TM image. From Table 3, we can see that GNSPI can reduce the RMSE value of NSPI from 0.0121 to 0.0072 for green band, from 0.0173 to 0.0108 for red band, and from 0.0398 to 0.0256 for NIR band when both methods use TM image of February 8 as input data. The relative reduction of RMSE value is more than 35% for each band, indicating that GNSPI can greatly improve the accuracy of gap filling when a temporally distant input image is used.

Table 3
RMSE of gap-filling results.

Band	PGM	NSPI-TM1	GNSPI-TM1	NSPI-TM2	GNSPI-TM2	NSPI-ETM+	GNSPI-ETM+
Green	0.0151	0.0121	0.0072	0.0063	0.0057	0.0071	0.0064
Red	0.0244	0.0173	0.0108	0.0095	0.0086	0.0113	0.0102
NIR	0.0550	0.0398	0.0256	0.0209	0.0190	0.0235	0.0224

As mentioned in the Methodology section, GNSPI can also provide the uncertainty of the prediction by Eq. (13). Table 4 shows the mean uncertainty of each filled results by PGM and GNSPI. Comparing the mean uncertainty of PGM and GNSPI, we can see that GNSPI can greatly reduce the prediction uncertainty. In addition, it is apparent that the uncertainty of prediction by temporally further input images would be larger than that by temporally closer input images. Therefore, using the temporally closer TM images for filling the gaps can reduce the prediction uncertainty. Although the multiple ETM+ input images are also temporally close to the target image, the uncertainty of GNSPI-ETM+ is larger than that of GNSPI-TM2. The reason is likely that the sample pixels selected from ETM+ input image are generally spatially farther from the target pixel than the similar pixels selected from the TM input image.

4.2. Classification of filled results

Table 5 shows the accuracy assessment of the classifications of all filled images using the classification of the actual image as reference data. Higher values of *oa* and *kappa* indicate a stronger agreement between the classifications of the filled image and the actual image. Classification of PGM result has the lowest values of *oa* and *kappa*. The *oa* value of PGM is only 0.6572, which means that more than 34% of pixels are classified as different land cover types from the actual

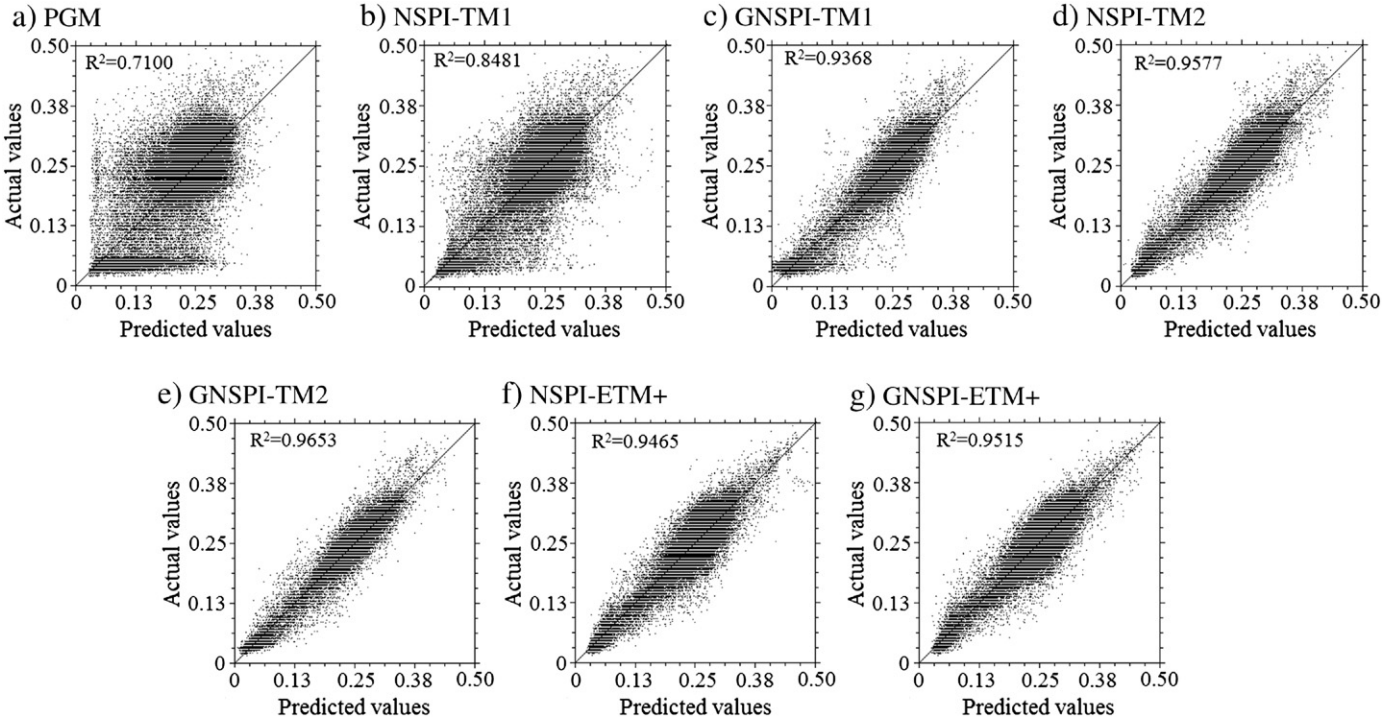

Fig. 11. Scatter plots of the actual and the predicted values of all un-scanned pixels for NIR band. Panels (a)–(g) are the scatter plots of Fig. 10(b)–(h), respectively.

Table 4
The mean uncertainty (%) of filled results by PGM and GNSPI.

Band	PGM	GNSPI-TM1	GNSPI-TM2	GNSPI-ETM+
Green	34.25	15.81	11.30	13.54
Red	61.36	28.74	15.78	28.92
NIR	78.63	29.52	21.13	19.10

image. Therefore, at least in this case study, the filled image by PGM is not good enough for the application of land cover mapping. Generally, classifications of images filled by NSPI and GNSPI have similar accuracy except when using temporally further input images. When using closer input images, both the classification of images filled by NSPI and GNSPI can get an *oa* value around 0.9, which would be good enough for the applications in practice. However, when using temporally further input images, the accuracies of the classification of the images filled by NSPI and GNSPI are both lower than those using a temporally closer input image. However, the reduction of accuracy for GNSPI is smaller than NSPI, indicating that GNSPI is more robust than NSPI in producing a filled image. Therefore, GNSPI would satisfy further applications like land cover mapping even using temporally distant input images.

5. Discussion and conclusions

Landsat 7 ETM+ images have been widely used in studies at regional and global scales, but the SLC failure greatly reduces their utility. In addition, the failure is permanent after subsequent efforts to recover the SLC were not successful, so the necessary and feasible way to address this problem is to fill the gaps in the SLC-off images. Although there are many methods proposed to fill the gaps, the quality of the filled images in heterogeneous regions is still not satisfactory for further applications, with the exception of images filled using the recently developed NSPI method. However, NSPI predicts the value of un-scanned pixels empirically and cannot provide prediction uncertainty. This paper proposed a new method, GNSPI, to fill the gaps in SLC-off images, which is based on the geostatistical theory and NSPI. Through the simulation study, GNSPI was competitive compared with the previous geostatistical methods and NSPI. For the purpose of simplicity and consistency with NSPI, only results of bands 2, 3, and 4 were reported, but the other bands show similar results as these three bands in our simulation study. GNSPI can largely improve the accuracy of gap filling compared with the previous geostatistical method. Although GNSPI is only slightly better than NSPI when using temporally closer input images, GNSPI greatly reduces RMSE values of filled results compared with NSPI when using temporally further input images. The robustness of GNSPI when using temporally further input images is very important because it is difficult to acquire a temporally closer input image in cloudy regions (Ju & Roy, 2008). However, in some frequently cloudy regions, it is also difficult to acquire cloud-free input images even though they are temporally distant from the target image. For this case, clouds in the input image should be removed by cloud removal methods (Melgani, 2006; Zhu et al., 2012) before it is used to fill gaps.

Compared with the previous geostatistical methods, GNSPI has two major strengths. First, GNSPI makes better use of the temporal information from the input image. The information of the input image is used to predict the values of all pixels outside the gaps by using the temporal relationship of Eq. (1). It removes the temporal trend of the target image and obtains the residual image. Ordinary kriging technique requires the image to be an intrinsically stationary spatial process, that is, with constant mean. The previous geostatistical methods apply an ordinary kriging technique directly on the original reflectance image (Pringle et al., 2009; Zhang et al., 2007). However, for the original reflectance image, the intrinsic stationarity assumption is more difficult to satisfy because it is very common that some very low reflectance objects like water and also very bright objects like bare ground coexist in the same image. On the contrary, for the residual image after removing the temporal trend, it better satisfies the assumption of an intrinsically stationary spatial process, so it is more reasonable to apply ordinary kriging on the residual image than the original reflectance image. In addition, GNSPI emphasizes the temporal information of the pixel in the input image at the same location of the gap pixel, which can help to keep the spatial details of the filled image. The previous geostatistical methods can also apply co-kriging technique to use the information of the secondary image. However, the previous geostatistical methods select samples from the secondary image for the predicting process rather than emphasizing the sample at the same location of the un-scanned pixel. This may be the reason why there is no significant difference between filled results of ordinary kriging and co-kriging (Zhang et al., 2007). Second, GNSPI considers the difference between different classes. GNSPI models the spatial dependence for each class rather than for the whole image, which can better satisfy the intrinsic stationarity assumption. Pringle et al. (2009) admitted that it is not plausible to assume that the whole image satisfies the intrinsic stationarity assumption. In addition, GNSPI uses the sample pixels, which have high spectral similarity with the target pixel, to predict the value of the target pixel. Conversely, the previous geostatistical methods selected the spatially closest samples rather than the samples with the highest spectral similarity, which would cause the striping effects on the filled image. For instance, a target pixel in a narrow river uses its surrounding pixels from bare ground as samples, which will cause large prediction errors. In all, through better use of temporal information and consideration of the difference between classes, GNSPI can significantly improve the accuracy of gap filling compared with the previous geostatistical methods.

Compared with NSPI, GNSPI has three main advantages. First, while both GNSPI and NSPI use weighted average interpolator to predict the un-scanned pixel, the weights are generated in different ways. Weights in GNSPI are calculated from geostatistical theory and depend on the spatial dependence of the image. Weights in NSPI are empirically determined and do not depend on the image itself. Therefore, GNSPI should be more robust in various landscapes than NSPI. Second, GNSPI selects better sample pixels in the neighborhood of un-scanned target pixel than NSPI. GNSPI uses a time-series as the auxiliary data to refine the similar pixels, and selects sample pixels with more similar spectral characteristics and more similar temporal change pattern to the target pixel. NSPI just uses the input image itself to select samples. In the case of long temporal intervals between the input and target images, the similar pixels selected by NSPI would be incorrect. Considering that the probability of acquiring one cloud-free observation of a Landsat pixel per season in most regions is high (Ju & Roy, 2008), it is reasonable and feasible to get help from these cloud-free observations to refine the similar pixels. Last but not least, GNSPI applies a geostatistical approach which can provide the uncertainty of prediction, while NSPI is still a type of deterministic approach which cannot produce statistical uncertainty for each prediction (Chen et al., 2011). The prediction uncertainty is important and helpful for further applications of the filled images. It can help users to estimate the potential errors of using filled images as input data sources in their studies. Although the improvement of GNSPI from NSPI is not that significant compared with the improvement of GNSPI from previous geostatistical methods, it has a number of strengths and provides a new perspective to address the SLC-off problem.

Table 5
Standard accuracy assessment of land cover classification from all filled results.

Index	PGM	NSPI-TM1	GNSPI-TM1	NSPI-TM2	GNSPI-TM2	NSPI-ETM+	GNSPI-ETM+
oa	0.6572	0.8248	0.8719	0.8936	0.9057	0.8811	0.8761
kappa	0.5302	0.7595	0.8235	0.8537	0.8703	0.8366	0.8300

It should be noted that GNSPI has some predetermined parameters. Basically, the sample size and the number of classes are two parameters that are most sensitive in GNSPI. Therefore, we mainly investigated the effects of sample size and the number of classes. For sample size *M*, we tested GNSPI by using different sample sizes. The results showed that the accuracy of gap filling is increasingly improved with sample size but it stabilizes when sample size is larger than 20, which is consistent with the results in NSPI (Chen et al., 2011). Considering that larger sample size will cost more computing time, we also recommend 20 as the appropriate sample size for filling gaps using GNSPI in practice. For the classification of input image, the minimum and maximum numbers of classes need to be specified so that ISODATA can automatically determine the optimal number of classes by merging or splitting the classes. Obviously, the number of land cover classes should not be too small or too large: a small number of classes cannot capture the difference in temporal relationships for different land cover types; a large number of classes would make the land cover map more fragmented, causing the sample pixels distant from the target pixel. Therefore, the sensitivity of GNSPI to the number of classes was tested. The result showed that the accuracy of filled images is comparable if they use the number of classes close to the optimal one from ISODATA classifier, indicating the relative robustness of GNSPI to the number of classes.

One disadvantage of GNSPI is its relatively slow computing speed. In the case study, GNSPI needed almost twice the time of NSPI to fill all the un-scanned pixels. However, considering the advantages of GNSPI, i.e., the most accurate gap-filling algorithm and providing uncertainty measures, the relatively slow computing speed is acceptable and worthwhile. There would be three options to improve the computing efficiency. First, fewer sample pixels will reduce the computing time. However, a small sample size will reduce the prediction accuracy. Zhang et al. (2007) only used 8 pixels spatially nearest to the target pixel for prediction, which was considered not enough by Pringle et al. (2009). Webster and Oliver (2001) recommend at least 20 samples for kriging interpolation, which is followed in our case study. Therefore, it is necessary to find a balance between prediction accuracy and computing time when deciding an appropriate sample size. Second, a hybrid method is worth considering in the future, which has been explored by Pringle et al. (2009) and demonstrated to be a feasible way to increase the computing efficiency. This hybrid method could combine GNSPI with several high-speed methods, and could automatically select one of these methods according to the specific situations. For example, for homogeneous regions, the local linear histogram matching method (USGS, 2004) can be applied to predict the un-scanned pixels; NSPI can be chosen if the input image is temporally close to the target image. Last, the computing speed can be speeded up using high-performance computers or parallel computing.

Acknowledgments

This study was supported in part by the International S&T Cooperation Program (2010DFB10030) and by the National Science Foundation of USA (1010314), and by the State Key Laboratory of Earth Surface Processes and Resource Ecology. We thank Jungkuan Liu, Shiguo Jiang, Shanshan Cai and Christine Biermann for helpful comments on the manuscript.

References

Arvidson, T., Goward, S., Gasch, J., & Williams, D. (2006). Landsat-7 long-term acquisition plan: Development and validation. *Photogrammetric Engineering and Remote Sensing*, *72*, 1137–1146.

Ball, G. H., & Hall, D. J. (1965). *ISODATA, a novel method of data analysis and classification.* Tech. Rep. Stanford, CA: Stanford University.

Bédard, F., Reichert, G., Dobbins, R., & Trépanier, I. (2008). Evaluation of segment-based gap-filled Landsat ETM+ SLC-off satellite data for land cover classification in southern Saskatchewan, Canada. *International Journal of Remote Sensing*, *29*, 2041–2054.

Byrne, G. F., Crapper, P. F., & Mayo, K. K. (1980). Monitoring land-cover change by principal component analysis of multi-temporal Landsat data. *Remote Sensing of Environment*, *10*, 175–184.

Call, K. A., Hardy, J. T., & Wallin, D. O. (2003). Coral reef habitat discrimination using multivariate spectral analysis and satellite remote sensing. *International Journal of Remote Sensing*, *24*, 2627–2639.

Chen, J., Zhu, X., Vogelmann, J. E., Gao, F., & Jin, S. M. (2011). A simple and effective method for filling gaps in Landsat ETM plus SLC-off images. *Remote Sensing of Environment*, *115*, 1053–1064.

Cohen, W. B., & Goward, S. N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience*, *54*, 535–545.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46.

Cressie, N. (1985). Fitting variogram models by weighted least squares. *Journal of the International Association for Mathematical Geology*, *17*, 563–586.

Fisher, J. I., Mustard, J. F., & Vadeboncoeur, M. A. (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment*, *100*, 265–279.

Fuller, R. M., Groom, G. B., & Jones, A. R. (1994). The land-cover map of Great-Britain — An automated classification of Landsat Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, *60*, 553–562.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation.* New York: Oxford University Press.

Goward, S. N., Arvidson, T. J., Faundeen, J., Williams, D. L., Irons, J., & Franks, S. (2006). Historical record of Landsat global coverage: Mission operations, NSLRSDA, and international cooperator stations. *Photogrammetric Engineering and Remote Sensing*, *72*(10), 1155–1169.

Huang, C., Davis, L. C., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*, 725–749.

Ju, J. C., & Roy, D. P. (2008). The availability of cloud-free Landsat ETM plus data over the conterminous United States and globally. *Remote Sensing of Environment*, *112*, 1196–1211.

Liu, D., & Cai, S. (2011). A spatial-temporal modeling approach to reconstructing land-cover change trajectories from multi-temporal satellite imagery. *Annals of the Association of American Geographers*, http://dx.doi.org/10.1080/00045608.2011.596357

Masek, J. G., Huang, C. Q., Wolfe, R., Cohen, W., Hall, F., Kutler, J., et al. (2008). North American forest disturbance mapped from a decadal Landsat record. *Remote Sensing of Environment*, *112*, 2914–2926.

Matheron, G. (1971). *The theory of regionalized variables and its applications [(Les Cahiers du center de Morphologie Mathematique de Fontainebleau 5)].* Fontainebleu: CMMF.

Maxwell, S. K. (2004). Filling landsat ETM+ SLC-off gaps using a segmentation model approach. *Photogrammetric Engineering and Remote Sensing*, *70*, 1109–1111.

Maxwell, S. K., Schmidt, G. L., & Storey, J. C. (2007). A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images. *International Journal of Remote Sensing*, *28*, 5339–5356.

Melgani, F. (2006). Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 442–455.

Papritz, A., & Stein, A. (1999). Spatial prediction by linear kriging. In A. Stein, F. D. van der Meer, & B. G. H. Gort (Eds.), *Spatial statistics for remote sensing* (pp. 83–113). Dordrecht: Kluwer Academic.

Pringle, M. J., Schmidt, M., & Muir, J. S. (2009). Geostatistical interpolation of SLC-off Landsat ETM plus images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 654–664.

Riano, D., Chuvieco, E., Salas, J., & Aguado, I. (2003). Assessment of different topographic corrections in Landsat-TM data for mapping vegetation types. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 1056–1061.

Roy, D. P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., et al. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment*, *112*, 3112–3130.

Tarabalka, Y., Benediktsson, J. A., & Chanussot, J. (2009). Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 2973–2987.

Teillet, P. M., Barker, J. l., Markham, B. L., Irish, R. R., Fedosejevs, G., & Storey, J. C. (2001). Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets. *Remote Sensing of Environment*, *78*, 39–54.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, *46*, 234–240.

USGS (2004). Phase 2 gap-fill algorithm: SLC-off gap-filled products gap-fill algorithm methodology. Available online at landsat.usgs.gov/documents/L7SLCGapFilledMethod.pdf (accessed on November 2, 2011).

USGS (2012). Landsat 5 suspension of operations extended. Available online at http://www.usgs.gov/newsroom/article.asp?ID=3109 (accessed on March 25, 2012).

Webster, R., & Oliver, M. A. (2001). *Geostatistics for environmental scientists.* Chichester: John Wiley & Sons.

Williams, D. L., Goward, S., & Arvidson, T. (2006). Landsat: Yesterday, today, and tomorrow. *Photogrammetric Engineering and Remote Sensing*, *72*, 1171–1178.

Zhang, C., Li, W., & Travis, D. (2007). Gaps-fill of SLC-off Landsat ETM plus satellite image using a geostatistical approach. *International Journal of Remote Sensing*, *28*, 5103–5122.

Zhu, X., Chen, J., Gao, F., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, *114*, 2610–2613.

Zhu, X., Gao, F., Liu, D., & Chen, J. (2012). A modified neighborhood similar pixel interpolator approach for removing thick clouds in Landsat images. *IEEE Geoscience and Remote Sensing Letters*, *9*(3), 521–525.