



A simple and effective method for filling gaps in Landsat ETM+ SLC-off images

Jin Chen ^a, Xiaolin Zhu ^{a,b}, James E. Vogelmann ^{c,*}, Feng Gao ^{d,e}, Suming Jin ^f

^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 43210-1361, USA

^c USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA

^d Biospheric Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

^e Earth Resources Technology, Inc., Laurel, MD 20707, USA

^f ASRC Research and Technology Solutions (ARTS), Contractor to U.S. Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD 57198, USA

ARTICLE INFO

Article history:

Received 30 August 2010

Received in revised form 6 December 2010

Accepted 11 December 2010

Available online 15 January 2011

Keywords:

Landsat ETM+

SLC-off

Gap filling

ABSTRACT

The scan-line corrector (SLC) of the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor failed in 2003, resulting in about 22% of the pixels per scene not being scanned. The SLC failure has seriously limited the scientific applications of ETM+ data. While there have been a number of methods developed to fill in the data gaps, each method has shortcomings, especially for heterogeneous landscapes. Based on the assumption that the same-class neighboring pixels around the un-scanned pixels have similar spectral characteristics, and that these neighboring and un-scanned pixels exhibit similar patterns of spectral differences between dates, we developed a simple and effective method to interpolate the values of the pixels within the gaps. We refer to this method as the Neighborhood Similar Pixel Interpolator (NSPI). Simulated and actual SLC-off ETM+ images were used to assess the performance of the NSPI. Results indicate that NSPI can restore the value of un-scanned pixels very accurately, and that it works especially well in heterogeneous regions. In addition, it can work well even if there is a relatively long time interval or significant spectral changes between the input and target image. The filled images appear reasonably spatially continuous without obvious striping patterns. Supervised classification using the maximum likelihood algorithm was done on both gap-filled simulated SLC-off data and the original “gap free” data set, and it was found that classification results, including accuracies, were very comparable. This indicates that gap-filled products generated by NSPI will have relevance to the user community for various land cover applications. In addition, the simple principle and high computational efficiency of NSPI will enable processing large volumes of SLC-off ETM+ data.

Published by Elsevier Inc.

1. Introduction

The Landsat series of satellites provides an unparalleled data source for land surface mapping and monitoring (Byrne et al., 1980; Cohen & Goward, 2004; Hansen et al., 2008; Healey et al., 2005; Masek et al., 2008; Vogelmann et al., 2001). The Landsat sensors include the Landsat 5 Thematic Mapper (TM), the Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and the Landsat 1–5 Multispectral Scanners (MSS). The high value of the data from Landsat can be attributed in part to long-term repeat coverage (1972–present) and relatively high spatial resolution (30 m for the TM and the ETM+, and 80 m for the MSS sensors). Both Landsat 5 and Landsat 7 are still functioning, although both have substantially exceeded their planned design lives.

On May 31, 2003, the scan-line corrector (SLC) for the ETM+ sensor on board Landsat 7 failed permanently. The SLC compensates for the forward motion of the satellite, and without an operating SLC, images have wedge-shaped gaps that range from a single pixel in

width near the image-nadir, to about 12 pixels towards the edges of the scene. Missing pixels comprise about 22% pixels of these images (Arvidson et al., 2006; Ju & Roy, 2008). The deteriorated image quality resulting from SLC failure has become a major obstacle for Landsat ETM+ data applications. Accordingly, images acquired before the SLC failure are designated SLC-on images, while those acquired after the SLC failure are designated SLC-off images.

Soon after SLC failure, a joint United States Geological Survey/National Aeronautics and Space Administration (USGS/NASA) Landsat team explored a number of different options for filling in the data gaps in the SLC-off images. One method developed was a local linear histogram-matching method using one or more SLC-off or SLC-on images (USGS, 2004). This method applies a local linear histogram matching in a moving window of each missing pixel to derive the re-scaling function. This re-scaling function is then used to convert the radiometric values of one input scene into equivalent radiometric values of the scene being gap-filled, and the transformed data are then used to fill the gaps of that scene. This method is very simple and easy to implement, and can resolve many of the missing-data problems if the input scenes are of high quality (e.g., negligible cloud and snow cover) and represent comparable seasonal conditions (USGS, 2004).

* Corresponding author. Tel.: +1 605 594 6062.

E-mail address: vogel@usgs.gov (J.E. Vogelmann).

As an alternative to this method, Roy et al. (2008) proposed to use the information observed by MODIS to estimate reflectance of the un-scanned pixels. Maxwell et al. (2007) developed another approach, whereby multi-scale segmentation was used to fill gaps in the Landsat 7 ETM+ SLC-off images. This approach was applied to three different areas, demonstrating that the filled products were useful for a wide variety of applications, such as general land cover mapping and visual assessment (Bédard et al., 2008). Geostatistics based methods have also been employed (Zhang, et al., 2007; Pringle, et al., 2009), in which kriging or co-kriging techniques have been used to fill the data gaps. The case studies showed that these geostatistical methods can also be very effective for interpolating the missing pixels in the SLC-off imagery.

While the above methods can restore the un-scanned gaps in ETM+ SLC-off imagery, sometimes with very good results, it should be noted that the above methods also suffer from a number of limitations that have precluded their widespread use, especially for quantitative application. For instance, the local linear histogram-matching method can yield satisfactory results in homogenous regions such as forests, but it tends to have difficulty with heterogeneous landscapes where the size of surface objects are smaller than the local moving window size (USGS, 2004). In general, using information from non-Landsat sensors is constrained by spectral compatibility and spatial resolution issues. Few instruments with high spatial resolution are spectrally similar to ETM+; one sensor that has comparable spectral bands is MODIS, but this sensor has much coarser spatial resolution (Roy et al., 2008). The multi-scale segmentation approach has a disadvantage in having lower reflectance prediction accuracy at the pixel level, especially for narrow or small objects, such as roads and streams (Maxwell et al., 2007). The geostatistical interpolation methods also have two major drawbacks. First, these methods do not predict the reflectance well at the pixel-level, and thus are not optimal for small and discrete objects. Secondly, these geostatistical approaches are very computationally intensive, which limit their implementation for mass production (Zhang, et al., 2007; Pringle, et al., 2009).

Based on the shortcomings in the above methods, the aim of this study is to demonstrate the application of a simple and effective method to fill the gaps in SLC-off ETM+ imagery. This new method, which we will refer to hereafter as the Neighborhood Similar Pixel Interpolator (NSPI), has the potential to interpolate the value of pixels located in the gaps accurately, especially improving results in heterogeneous landscape areas. In this paper, we will first describe this approach and the algorithm, and later on we will demonstrate its use and performance on simulated and actual SLC-off images.

2. Algorithm development

It is reasonable to assume that neighboring pixels in close proximity to SLC-off gaps share similar spectral characteristics and temporal patterns of changes with the missing pixels located within the gaps, if they belong to the same land cover type. Thus it is logical to make use of the information of the same-class neighboring pixels to restore spectral reflectance of missing pixels. Here, for convenience in this paper, the SLC-off image that will be filled is defined as the target image, while the other images that are selected to fill the gaps in the target image are referred as the input images. There are two data sources that can be used to fill the gaps in target image: (1) an appropriate TM image or SLC-on ETM+ image, and (2) SLC-off images acquired at different dates, whereby the scanned parts of these images partly overlap with the gaps in the target image. The steps for gap filling implementation for these two data sources will be introduced respectively below. Since the local atmospheric conditions are usually relatively homogenous, pixels within a given neighborhood will normally be under similar atmospheric effects. The NSPI can be applied either (1) to top-of-atmosphere radiance or DN value (if the

radiometric calibration formula is the same between input and target images), or (2) to top-of-atmosphere reflectance or reflectance products after the atmospheric correction.

2.1. Using a single TM or SLC-on ETM+ image

A TM image acquired reasonably close to the date of the target image can be used as the input image. Target images that are the most similar to the input images in seasonality and acquired under comparable sun illumination conditions are the best scenes to use. Similarly, an SLC-on ETM+ image can also be used, assuming that there has not been a significant amount of land use and land cover change between data acquisition times. Before implementing the filling process, the input image must be geometrically rectified to match the target image. Fig. 1 presents a flowchart of the gap filling method using a single input image. All steps will be discussed in detail below.

2.1.1. Selection of neighboring similar pixels

Based on the assumption that the same land cover class pixels in close proximity to the gaps have similar spectral characteristics and temporal patterns of change with the target missing pixel, it is necessary to search for similar pixels near the gaps. We assume that no major land cover changes occurred during the period between acquisitions of the input image and target image. Assuming that the time interval between input and target scene acquisitions is short, we think that this assumption will generally be valid. Accordingly, we can select the similar pixels from the input image and assume that these pixels are also spectrally similar with the target missing pixel at the target image. Here, an adaptive moving window searching procedure is employed. As shown in Fig. 2, all common pixels that are located in the moving window but outside the gaps with valid values both in target and input images are selected. When using TM or SLC-on ETM+ images as input images, the gaps only exist in the target image within the window (black part in Fig. 2). Similar pixels are then selected from these common pixels according to spectral similarity. Here the spectral similarity is defined as root mean square deviation (RMSD) between each common pixel and the target pixel as Eq. (1). The target pixel is a pixel located in the gaps of the target image without a valid value.

$$RMSD_i = \sqrt{\frac{\sum_{b=1}^n (L(x_i, y_i, t_1, b) - L(x, y, t_1, b))^2}{n}} \quad (1)$$

$L(x_i, y_i, t_1, b)$ is the value of i th common pixel located in (x_i, y_i) in band b for the input image acquired at t_1 , $L(x, y, t_1, b)$ is with same definition but for a target pixel, and n is the number of spectral bands. A large RMSD denotes a large spectral difference. Then, a threshold is used to identify similar pixels that have an RMSD values lower than the threshold. The threshold can be determined by the standard deviation of a population of pixels from the input image and the estimated number of land cover classes (Gao et al., 2006). If RMSD of the i th common pixel satisfies Eq. (2), the i th common pixel is selected as a similar pixel:

$$RMSD_i \leq \left[\sum_{b=1}^n \sigma(b) \times 2 / m \right] / n \quad (2)$$

where $\sigma(b)$ is the standard deviation of the whole input image for band b , and m is the number of classes. The estimated number of classes (m) needs to be predefined. This value is an empirical threshold and varies with the complexity of the landscape. It can be estimated by visual interpretation of the input images, or using a prior land cover map. In this experiment we used the value of five for m . Use of a larger number of classes represents a stricter condition for

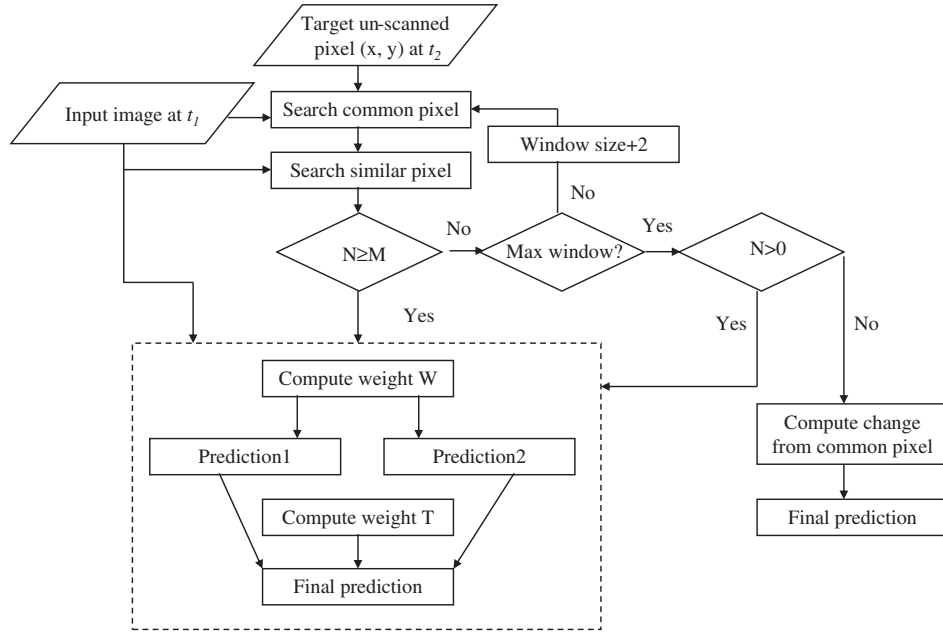


Fig. 1. The flowchart of the NSPI using a single input image.

selection of similar pixels from the input images. Considering that artifacts such as random noise, and atmospheric haze and small clouds might cause some uncertainty for prediction if the number of selected similar pixels is too small, the sample size of selected similar pixels must meet a minimum requirement. Here, we defined the minimum number of similar pixels as M . This value also needs to be predefined. For the test area, we empirically determined that a value of 20 was appropriate (discussed in further detail below). The initial moving window size (IWS) can then be calculated from the minimum sample size M from the following equation:

$$IWS = \left[\left(\sqrt{M} + 1 \right) / 2 \right]^* 2 + 1 \quad (3)$$

The center of the moving window is the target pixel located at (x, y) , and $[*]$ is the integer part of $(\sqrt{M} + 1) / 2$. Beginning with this initial

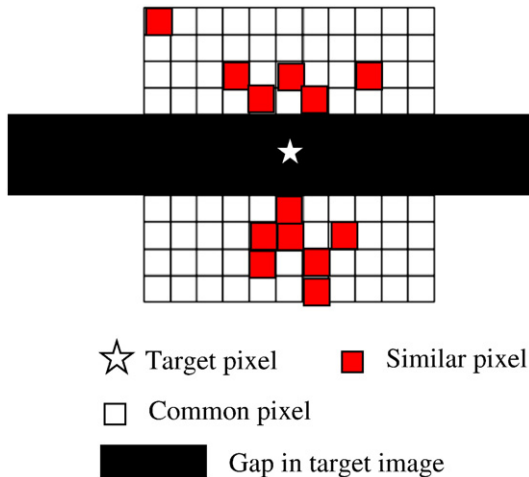


Fig. 2. Schematic diagram of the similar pixels selection when using a single input image.

window, the number of selected similar pixels (N) is counted. If the number of similar pixels cannot meet the required sample size of selected similar pixels M , the window is expanded by two pixels (i.e., 5×5 to 7×7 ...) to search similar pixels in a larger area. The searching procedure iterates many times until the minimum number of similar pixels M is met. Because large window sizes will take much computing time, we set a maximum window size with the size of 17×17 . If the minimum number of similar pixels (M) is not met before window size reaches the maximum window size, all selected similar pixels are used, regardless of the total number of similar pixels. For the special situation when there is no similar pixel selected when the maximum window size has been reached, then all of the common pixels will be used to predict the value of target pixel.

2.1.2. Calculation of the weights for similar pixels

The information of all the similar pixels can be used to predict the value of the target pixel. However, the contribution of the similar pixels might vary because some similar pixels are likely to be more spectrally comparable to the target pixel than others. The weight W_j determines the contribution of j th similar pixel for predicting the value of the target pixel. This is determined by the location of the similar pixel and the spectral similarity between the similar pixel and target pixel (Gao et al., 2006). Higher spectral similarity and smaller distance of a similar pixel to the target pixel will increase the weight of that given pixel. In Eq. (4), the geographic distance D_j between the j th similar pixel (x_j, y_j) and the target pixel (x, y) is calculated:

$$D_j = \sqrt{(x_j - x)^2 + (y_j - y)^2} \quad (4)$$

The spectral similarity is determined by the RMSD between each similar pixel and the target pixel as in Eq. (2) above. Combining spectral similarity and geographic distance, a synthetic index CD can be computed as:

$$CD_j = RMSD_j \times D_j \quad (5)$$

As described above, a similar pixel with a larger CD value contributes less to the computed value for the target pixel, so we used the normalized reciprocal of CD as the weight W_j :

$$W_j = (1/CD_j) / \sum_{j=1}^N (1/CD_j) \quad (6)$$

The range of W_j is from 0 to 1, and the sum of all similar pixel weights is 1. For the special situation when p similar pixels hold the exact same spectral characteristics as the target pixel (i.e. $RMSD=0$), we set W for these special similar pixels as $1/p$, that is, all the information is given by these special similar pixels.

2.1.3. Calculation of the target pixel value

There are two methods to predict the radiometric value of the target pixel using the information provided by similar pixels. First, since the similar pixels have the same or approximate spectral value with the target pixel when they are observed at the same time, we can use the information of these similar pixels in the target image to predict the target pixel. In addition, the reliability of information provided by each similar pixel might be different. The larger weight W the similar pixel holds, the more reliable it is. Accordingly, the weighted average of all the similar pixels in the target image is used to make the first prediction for the target pixel:

$$L_1(x, y, t_2, b) = \sum_{j=1}^N W_j \times L(x_j, y_j, t_2, b) \quad (7)$$

Secondly, for the target pixel, the radiance value at t_2 equals the sum of radiance value at t_1 and the radiance change from t_1 to t_2 . For the reflectance product, this change can be caused by land surface variations, such as changes in phenology of vegetation and soil moisture. For the radiance (or DN) or top-of-atmosphere (TOA) reflectance products, this change also includes the atmospheric status variation besides the land surface state changes. Because the value at t_1 can be obtained directly from the input image, we only need to estimate the change of the target pixel from t_1 to t_2 . It is reasonable to assume that the change of similar pixels can represent the change of the target pixel, because the similar pixels have the same land cover type as the target pixel and are close to the target pixel in geographic space. Accordingly, the weighted average of the change provided by all of the similar pixels is used to calculate the value of the target pixel as the second prediction:

$$L_2(x, y, t_2, b) = L(x, y, t_1, b) + \sum_{j=1}^N W_j \times (L(x_j, y_j, t_2, b) - L(x_j, y_j, t_1, b)) \quad (8)$$

It should be noted that the spectral properties of Landsats TM and ETM+ are not identical (Teillet et al., 2001). Using a TM image as the input image might confer a small amount of error in the prediction. To alleviate the error, the $L(x_i, y_i, t_1, b)$ of the TM images can be normalized to the value of the ETM images before implementing the filling process.

The accuracy of prediction of Eqs. (7) and (8) might be different for different areas. For homogenous areas, the first prediction $L_1(x, y, t_2, b)$ might be more reliable than the second prediction $L_2(x, y, t_2, b)$ because the target pixel and all the similar pixels probably belong to the same object and should hold the approximate radiometric value. For heterogeneous areas, the second prediction might be more trustworthy because the target pixel and all the similar pixels probably belong to different objects. Thus, the radiometric value provided by the same location pixel in the input image will likely be more similar to the target pixel after accounting for the changes between the acquired dates of the target image and the input image. In addition, the first prediction comes from the target image and thus

is more radiometrically consistent. The second prediction takes into account the same location from input image and thus is more spatially consistent. Accordingly, a more accurate prediction can be obtained by a weighted combination of these two predicted results with consideration given to both landscape homogeneity and the extent of change. Here, we use the averaged $RMSD$ (R_1) between the similar pixel and the target pixel to denote the extent of the landscape homogeneity:

$$R_1 = \frac{1}{N} \sum_{j=1}^N \sqrt{\left[\sum_{b=1}^n (L(x_j, y_j, t_1, b) - L(x, y, t_1, b))^2 \right] / n} \quad (9)$$

In the same way, the averaged $RMSD$ (R_2) of similar pixels between observations at t_1 and t_2 is used to denote the extent of change between the input image and the target image:

$$R_2 = \frac{1}{N} \sum_{j=1}^N \sqrt{\left[\sum_{b=1}^n (L(x_j, y_j, t_1, b) - L(x_j, y_j, t_2, b))^2 \right] / n} \quad (10)$$

Then, according to above discussion, we use the normalized reciprocal of R_1 and R_2 as the weight T_1 and T_2 respectively:

$$T_i = (1/R_i) / (1/R_1 + 1/R_2), \text{ where } i = 1, 2 \quad (11)$$

The final predicted value of the target pixel located in the gap is calculated as:

$$L(x, y, t_2, b) = T_1 \times L_1(x, y, t_2, b) + T_2 \times L_2(x, y, t_2, b) \quad (12)$$

For the special situation where there is no similar pixel selected, the local linear histogram matching approach (USGS, 2004) is applied to predict the value of the target pixel. A gain and bias is calculated using the mean and standard deviation of the common pixels:

$$gain = \frac{\sigma_P}{\sigma_F} \quad (13)$$

$$bias = \mu_P - \mu_F \times gain$$

where μ_P and μ_F are the mean value of common pixels in the target and input image respectively, σ_P and σ_F are the standard deviations of common pixels in the target and input image respectively. Then, the value of the target pixel can be computed as:

$$L(x, y, t_2, b) = gain \times L(x, y, t_1, b) + bias \quad (14)$$

2.1.4. Data quality flag generation

A data quality flag layer as a supplement to the final filled products is helpful for users to evaluate data reliability. Accordingly, we generated such a layer, in which integer 0 to 3 indicates which gap-filling approach is used for each pixel (Table 1).

2.2. Using multiple SLC-off ETM+ images

Although Landsat 5 continues to amaze the scientific community with its longevity, instrument failure is inevitable in the near future (Bédard et al., 2008). A single Landsat 5 TM image as the input image

Table 1
Data quality flag for using a single input image.

Value	Meaning
0	This pixel is not in gaps
1	Predicted by similar pixels more than M
2	Predicted by similar pixels less than M
3	Predicted by local linear histogram method (no similar pixel were selected)

to fill SLC-off gaps may not be a long-term viable option. It is possible to fill gaps using multiple SLC-off ETM+ images because the gaps in these images do not overlap completely among scenes. The criteria for selecting the input images are (1) that their acquisition dates are close to that of the target image and (2) that they are mostly free of cloud contamination or snow cover. The number of necessary input images can be determined by the degree of gap overlapping. If gaps in input and target images do not overlap, then fewer input images are needed.

When multiple SLC-off ETM+ images are used as input images, the process of gap filling can be performed as shown in Fig. 3. First, all input images are sorted by acquisition date. The input image acquired closest in time to the target image is given first priority, and is the first one used in the process. Images acquired further from the target image acquisition date are used correspondingly later in the process. Next, as Fig. 3 shows, if the corresponding pixel of the target pixel in the first priority input image is scanned, the first input image is used to predict the value of the target pixel according to the method described in Section 2.1. Otherwise, the second priority input image is used. This process is repeated until all un-scanned target pixels in the target image are filled.

Since the input and target images both have un-scanned gaps, the maximum window size needs to be modified when using multiple SLC-off images as input images. As Fig. 4 shows, the number of common pixels within the window is less than that of using an input image without gaps. Accordingly, the threshold of the maximum window size is increased from 17 to 31. After finishing the gap-filling according to the procedure shown in Fig. 3, the data quality flag layer is also produced to identify which input image is used and which gap-filling approach is used for each target pixel (Table 2).

3. Data and experiment design

3.1. Study area and data

The study region is located in eastern Maryland around 39.10°N and 76.14°W, and is covered by World Reference System 2 Path 15 and Row 33. Within this Path/Row, we selected an intensive study area, which covers 15 km × 15 km (500 × 500 Landsat pixels). The major land cover types in this area include forest, water and arable land. The arable land in these scenes includes both “green” and “non-green” crops (the cover of the latter consists largely of bare ground, and is depicted as grey in the images in Fig. 5). Water covers almost one third of this study area and is relatively spectrally homogenous, while the other areas are (i.e., forest and arable land) are spectrally heterogeneous.

Five Landsat 5 TM images and three Landsat 7 ETM+ SLC-off images (Level 1 product) covering the study area were selected to

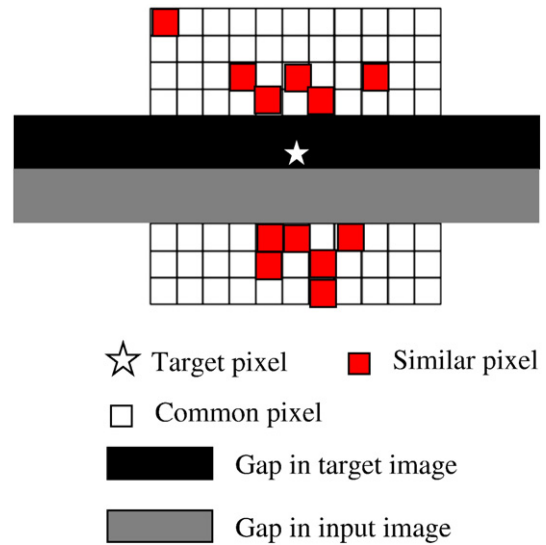


Fig. 4. Schematic diagram of the similar pixels selection when using multiple input images.

validate this gap-filling method. All images were clear of cloud contamination and snow cover. All of these images were calibrated to TOA reflectance with a range from 0 to 1.0. For this portion of the study, we only filled the gaps of the green, red, and near infrared (NIR) bands (bands 2, 3 and 4) to evaluate the performance of NSPI. For a later portion of the study, whereby we evaluated classification results of a gap-filled product, we used all six reflective 30 m Landsat bands (i.e., bands 1 through 5 and 7).

3.2. Experiment design

3.2.1. Experiments by using simulated SLC-off images

The NSPI was tested with simulated SLC-off images, which helps to understand its accuracy quantitatively. Fig. 5 shows four Landsat 5 TM images used to test the proposed method for a single TM image as the input image case (displayed as 4, 3 2 false color composites). They were acquired on May 25, 2008, June 10, 2008, February 8, 2010 and April 29, 2010. From Fig. 5, it can be seen that two TM images acquired in 2008 (a and b) are very similar to each other, while the two TM images acquired in 2010 (d and e) are less similar to each other, mainly due to the difference between winter (d) and spring (e) vegetation condition.

SLC-off images (Fig. 5c and f) were simulated for the images acquired on June 10, 2008 and April 29, 2010. For these images, simulated gaps were generated using an actual SLC-off image,

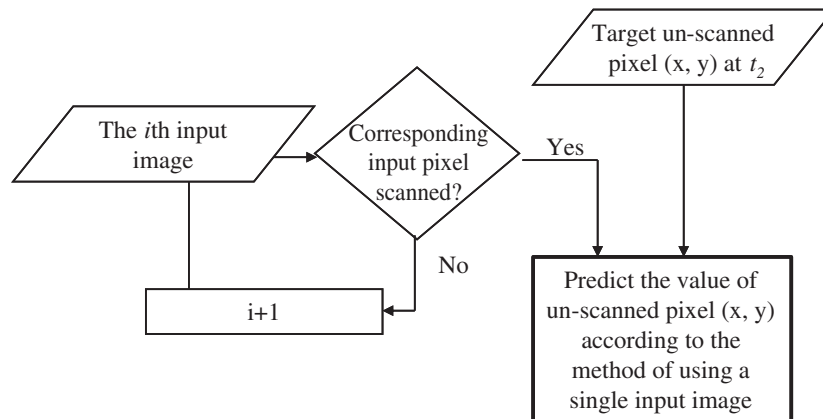


Fig. 3. The flowchart of the NSPI using multiple input images.

Table 2
Data quality flag for using multiple input images.

Value	Meaning
0	This pixel is without gaps
11	Predicted by similar pixels more than M using 1st input image
12	Predicted by similar pixels less than M using 1st input image
13	Predicted by common pixels using 1st input image
...	...
m1	Predicted by similar pixels more than M using m th input image
m2	Predicted by similar pixels less than M using m th input image
m3	Predicted by common pixels using m th input image

whereby zero values of the gap pixels from that image replaced the values for those same locations on the test images. There were a total of 250,000 un-scanned pixels per simulated SLC-off image. After generating the simulated SLC-off images, the TM images from May 25, 2008 (Fig. 5a) and February 8, 2010 (Fig. 5d) were used as input images to fill the June 10, 2008 and April 29, 2010 images, respectively. Gap filling was done both using the NSPI, and the local linear histogram method described by the USGS. We conducted pixel-based comparisons between NSPI and the local linear histogram method. However, due in part to the nature of the different interpolation approaches, we felt that conducting similar pixel-based comparisons between NSPI and the multi-scale segmentation approach (Maxwell et al., 2007) or the geostatistics based methods (Zhang, et al., 2007; Pringle et al., 2009) was statistically problematic and beyond the scope of the study. Because images a and c were acquired closer in time than that of the other two (d and f), we were able to assess the effect of spectral similarity between input and target image on the filled results. Finally the actual TM images acquired on June 10, 2008 and April 29, 2010 were used to validate performance of the NSPI and the local linear histogram method by comparing the filled images with the actual TM images.

For the case of using multiple SLC-off images as input images, two new SLC-off images were also simulated, in which a simulated SLC-off

TM data set was generated for the February 8, 2010 image by moving the gaps in Fig. 5f downwards 6 pixels. This data set was used to simulate the first priority SLC-off input image (Fig. 6a). Similarly, the second priority simulated SLC-off data set was generated using the January 23, 2010 data set (Fig. 6b) by moving the gaps in Fig. 5f upwards 6 pixels. The parts outside of gaps of these two simulated SLC-off images can cover all the gaps of Fig. 5f. Accordingly, Fig. 6a and b were used as input images to fill the gaps of Fig. 5f.

3.2.2. Experiments by using actual SLC-off images

The NSPI was also applied to actual SLC-off ETM+ images to evaluate its performance. The SLC-off ETM+ image acquired on September 22, 2008 (Fig. 7c) was selected as the target image. For the case of using a single input image to fill the gaps, the TM image from June 10, 2008 (Fig. 5b) was used. For the case of using multiple input images, two SLC-off ETM+ images acquired on February 11, 2008 and June 1, 2008 (Fig. 7a and b) were used to fill the gaps in the target image.

3.2.3. Classification experiments using gap-filled data from an entire scene

The NSPI was run on an entire scene (path 15 Row 33), and using all six 30-m reflective TM bands. For this portion of the study, we generated a simulated SLC-off data set using the April 29, 2010 scene (target scene), and using a May 7, 2007 scene as the input data. After this, we generated a land cover classification for the simulated SLC-off gap-filled data, and also for the actual (original non-gap) data. The comparison is described in further detail below.

3.3. Validation

3.3.1. Assessments of image characteristics

Both qualitative and quantitative evaluations were performed. For the qualitative assessment, the filled results were visually checked to determine whether they were spatially continuous, and whether

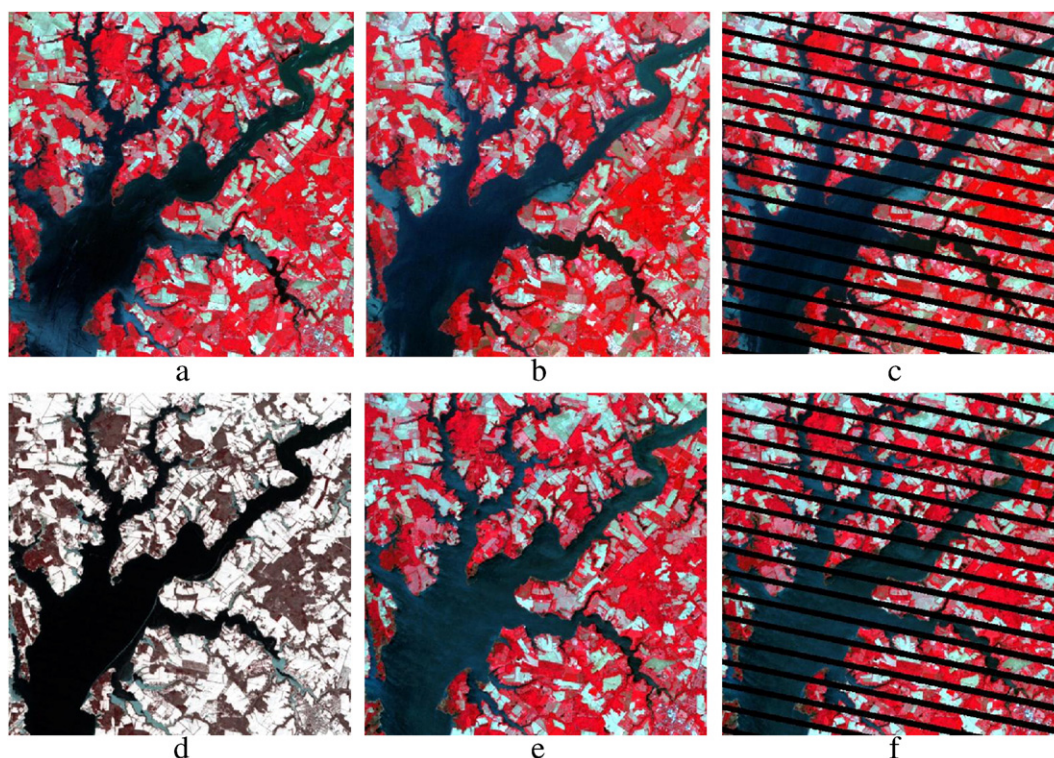


Fig. 5. NIR–red–green composites of Landsat TM images for the simulation test for using a single input image. (a) Acquired on May 25, 2008; (b) acquired June 10, 2008; (d) acquired February 8, 2010; (e) acquired April 29, 2010; (c) and (f) SLC-off image were simulated based on (b) and (e) respectively.

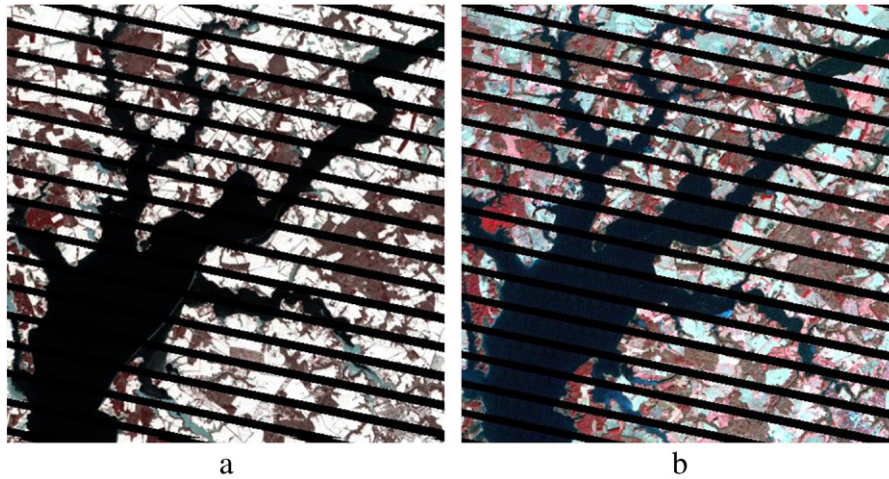


Fig. 6. Two simulated SLC-off images to test using multiple input images. (a) Simulated based on TM image from February 8, 2010 and (b) based on TM image from January 23, 2010.

there were stripes or other anomalies in the filled images. For the quantitative assessment, the filled images were compared with the actual image data through the calculation of two statistical indices. The first index used was the root mean square error (*RMSE*). This metric is frequently used to assess the differences between values predicted by a model and the values actually observed or measured. It is defined as Eq. (15).

$$RMSE = \sqrt{\frac{\sum_{f=1}^N [L(x_f, y_f, t_2, b) - L_0(x_f, y_f, t_2, b)]^2}{N}} \quad (15)$$

where N is the total number of un-scanned pixels, $L(x_f, y_f, t_2, b)$ and $L_0(x_f, y_f, t_2, b)$ are predicted and actual value of the f th un-scanned pixel respectively. A larger *RMSE* indicates a larger prediction error.

The second metric used was average difference (*AD*). This is used to evaluate either underestimation or overestimation of prediction, and is defined as Eq. (16). A positive value of *AD* indicates an over-prediction while a negative value represents an under-prediction.

$$AD = \frac{\sum_{f=1}^N [L(x_f, y_f, t_2, b) - L_0(x_f, y_f, t_2, b)]}{N} \quad (16)$$

3.3.2. Assessments of gap-filled imagery for applications

In order to determine whether or not NSPI can yield results relevant to the applications community, we generated a land cover product using the gap-filled data generated from the simulated SLC-

off data set, and compared this land cover product with an analogous one generated from the original image data source. For this test, we used the entire eastern Maryland scene, which is dominated by forest, urban areas, water, and arable lands. For both simulated gap-filled and original imagery, we converted data to TOA reflectance using coefficients from Chander et al. (2009). Land cover data were generated using the maximum likelihood supervised classifier. Training statistics were derived using 49 polygons; high resolution imagery from Google Earth was used to help define and categorize the polygons. Polygons were distributed throughout the entire image, and were selected to be representative of the different land cover types occurring within the region. Polygons were generally at least several hundred pixels in size, except for those polygons used to characterize urban areas, which were generally smaller. The classification scheme was relatively simple, and was comprised of the following land cover categories: water, urban, agricultural fields/grass, forest, and wetlands. The training polygons were from areas not covered by the simulated SLC-off gaps. The same polygons and accompanying statistics were used to classify both the gap-filled and the original Landsat data sets.

To help assess the accuracy of the classification results, 300 points were randomly selected from throughout the Landsat scene. Using Google Earth, a land cover class determination was made for each point. The points used did not overlap with training polygon areas. Points that were difficult to characterize (e.g., located on interfaces between two or more land cover categories, or land cover type was not easily determinable) were excluded from the analysis. Thus, we only used those points where we had high confidence that our call was correct. These reference data points were then compared with the

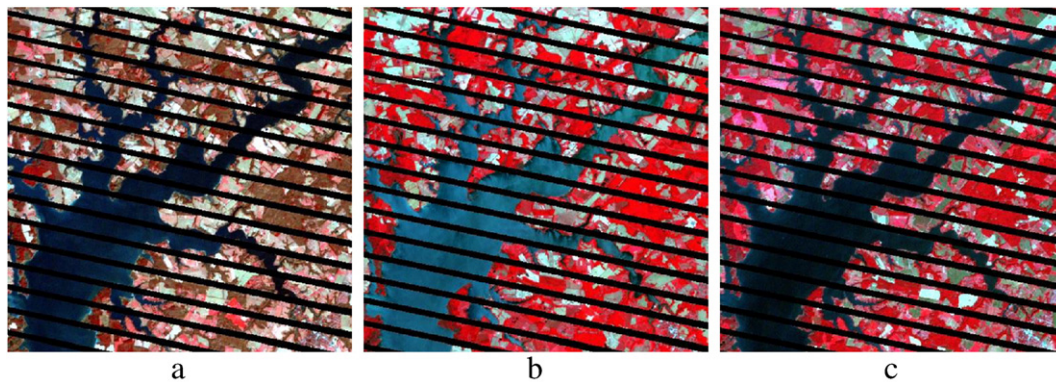


Fig. 7. Actual SLC-off ETM+ images for test. (a) Acquired February 11, 2008; (b) acquired June 8, 2008; (c) acquired September 22, 2008.

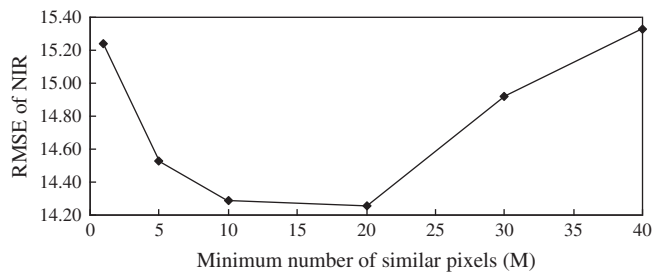


Fig. 8. The RMSE of NIR band of the fill results when using different values for parameter M , in the case of using Fig. 5d to fill gaps in Fig. 5f.

classification results to assess comparative accuracies between actual and gap-filled data sets. Of the 300 random points selected, 243 points were retained for accuracy assessment. Of these, 47 points were located within the gap fill areas. To augment the number of points associated with these fill areas, an additional 159 points were randomly selected from within the gap areas, bringing the total number of accuracy assessment points associated with the gap fill areas to 206. It should be noted that the goal of this portion of the investigation was not necessarily to generate the best classification possible (which can generally be done using more sophisticated approaches employing multiple dates of imagery, ancillary sources of data, and different algorithms). Rather, the goal was to generate comparative results between the actual and the NSPI data sets, thereby helping determine the change in usability of the NSPI data sets versus “actual” data.

4. Results

4.1. Optimization of parameter M

Using NSPI, the only parameter that needs to be set is M , which is defined as the minimum number of similar pixels. In an experiment, selection of M was optimized through comparative experimental

results in which this parameter was varied. Fig. 8 shows the NIR band RMSE of filled results of Fig. 5f by single input image using different M values. It is apparent that too small or too large a sample size decreases the accuracy of the filled result. The highest accuracy was obtained when the M was set as 20. In addition, filling accuracy reflected by RMSE is reasonably insensitive to the minimum sample size M when M value changes from 5 to 25. All the other bands and other filled results showed similar patterns, with optimized M ranging from 10 to 30. Therefore, considering that larger M not only decreases the computing efficiency but also improves filling accuracy insignificantly, we recommend 20 as an appropriate value of M in the tradeoff between computing efficiency and filling accuracy. From the flag layer of all the filled results, more than 98% un-scanned pixels in our intensive study area could be predicted by using similar pixels more than 20, indicating that 20 is an appropriate parameter using this method. There were only 180 pixels (0.07% of the intensive study area) for which there were no similar pixels. Thus, under normal situations, most target pixels are expected to have at least some similar pixels.

4.2. Filled results of simulated SLC-off images

Fig. 9a and b show the filled images of Fig. 5c using Fig. 5a as the input image by using the local histogram matching method and NSPI respectively. It is apparent that both filled image data sets are very similar to the actual image (Fig. 9c). Meanwhile, Fig. 9d and e are the filled images of Fig. 5f using Fig. 5d as the input image by the histogram matching method and NSPI respectively. In this case, the image filled by the new method appears much closer to the actual image Fig. 9f than the one filled by using the histogram matching method. There are visible stripes left in Fig. 9d, mainly located on the land surfaces rather than the water surface, indicating larger errors of the prediction by the histogram matching method in heterogeneous areas. Scatter plots in Fig. 10 show the relationships between the predicted values by these two methods and the actual values for NIR

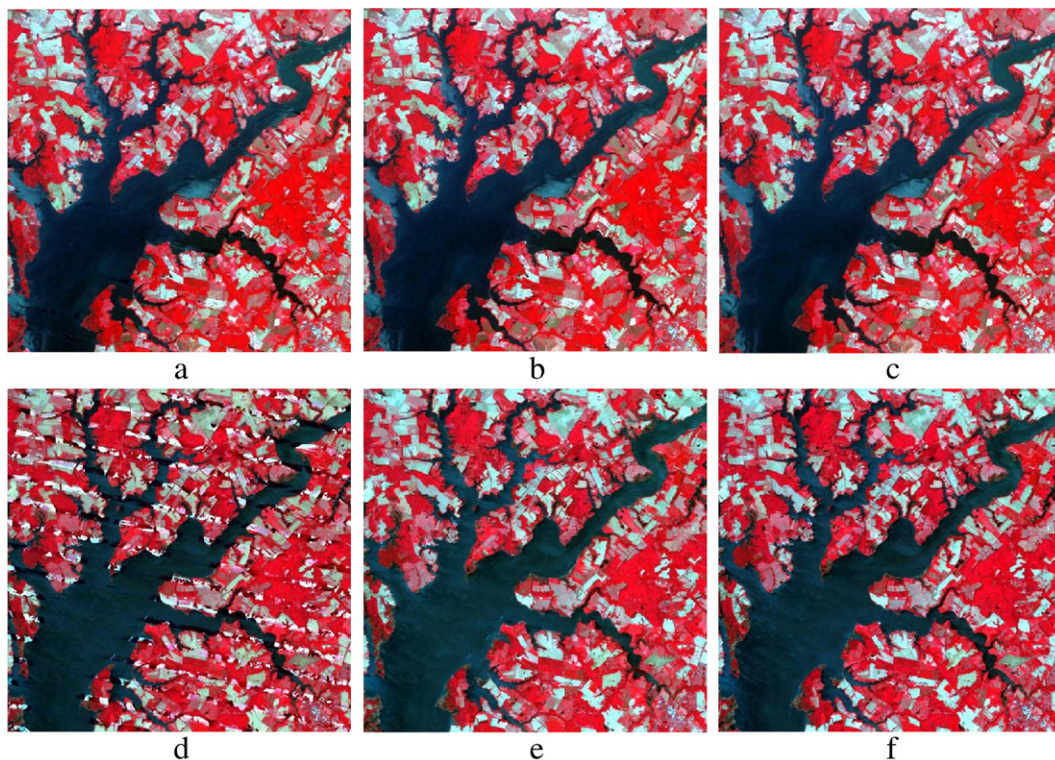


Fig. 9. Results of gap-fill for the simulation test of using a single input image. Panels (a) and (b) are the filled images of Fig. 5c by local linear histogram matching and NSPI respectively, and (c) is the actual image; panels (d) and (e) are the filled images of Fig. 5f by local linear histogram matching and NSPI respectively, and (f) is the actual image.

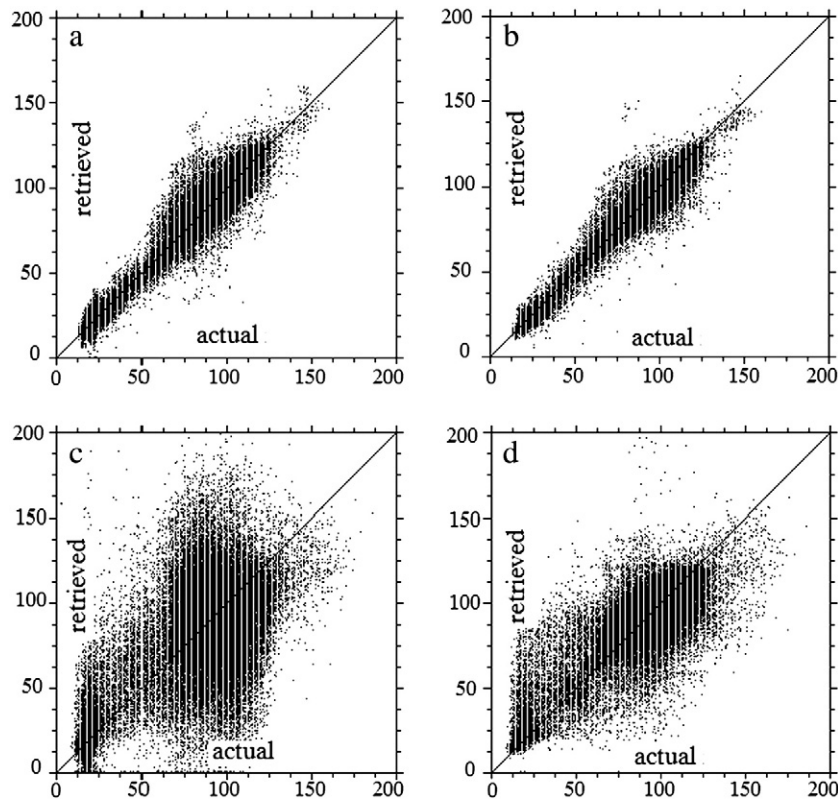


Fig. 10. Scatter plots of the real and the predicted (retrieved) value of all un-scanned pixels. Panels (a)–(d) are the scatter plots of Fig. 9a, b, d, e respectively. Axes denote top-of-atmosphere reflectance.

band. The data points in the scatter plot of the new method fall closer to the 1:1 line than those of the histogram equalization method, especially for the filled result of Fig. 5f, indicating that NSPI can more accurately predict the value of un-scanned pixels on the pixel level even when the acquisition date of input image is relatively far away from that of the target image.

In order to quantitatively compare the prediction of un-scanned pixels for the above two images, *RMSE* and *AD* of all the predictions are compared (Table 3). Generally, NSPI has higher prediction accuracy compared with results from the local linear histogram matching method. This is especially apparent when the time interval between acquisition dates of input and target images is larger. For example, compared to the histogram matching method, NSPI slightly reduces *RMSE* (from 0.0195 to 0.0153) for the NIR band when using the temporally closer input image. The *RMSE* is greatly reduced (from 0.0960 to 0.0398) using NSPI when the more temporally distant input image is used.

The same experiment was performed in Fig. 5f using multiple input images (Fig. 6a and b), and results are displayed in Fig. 11. It is evident that the restored image by NSPI seems much closer to the actual image (Fig. 11c) than that by the histogram matching method.

Table 3

The accuracy of filled results of Fig. 5c and f using a single input image.

Band	Method	Filled result of Fig. 5c		Filled result of Fig. 5f	
		<i>RMSE</i>	<i>AD</i>	<i>RMSE</i>	<i>AD</i>
Green	Histogram matching	0.0070	0.0000	0.3696	−0.1600
	NSPI	0.0052	0.0000	0.0121	−0.0007
Red	Histogram matching	0.0107	0.0000	0.4006	−0.1817
	NSPI	0.0079	0.0000	0.0173	−0.0008
NIR	Histogram matching	0.0195	−0.0002	0.0960	−0.0036
	NSPI	0.0153	0.0001	0.0398	0.0005

From the scatter plots for NIR band (Fig. 12), it is apparent that the histogram matching method estimated the values of the un-scanned pixels with larger errors. This is further supported by the quantitative accuracy assessment shown in Table 4. The *RMSE* values for green, red and NIR bands were notably lower using our method.

4.3. Filled results of actual SLC-off images

Fig. 13a and b shows the filled SLC-off ETM+ image from September 22, 2008 using the local linear histogram matching and NSPI respectively, both using the June 10, 2008 TM image as a single input. Both filled images appear spatially continuous without stripes, and both filled images appear similar to each other. This suggests both methods have comparable abilities for gap filling if the states of land surface of the input image are comparable to the target image.

The filled images by the histogram matching and NSPI using multiple input images are shown in Fig. 13c and d respectively. The quality of the filled image by NSPI appears to be better than that of the histogram matching in this case. The reason for the lower quality for the histogram matching result is likely because one of the input images (Fig. 6a) was acquired a relatively long time from the target image.

4.4. Land cover classification accuracy assessment

Classification of the entire simulated SLC-off gap filled data set using NSPI yielded results very similar to the classification generated for the actual data (non-gapped) data set. Aerial estimates for each land cover class were comparable between the gap-filled and reference data sets (Table 5). For example, deciduous forest covered 9818.3 km² in the gap-fill classification (31.9% of the scene), as compared with 9882.5 for the reference classification (32.1% of the scene). Urban areas, which are often difficult to classify in imagery due to the high degree of spectral heterogeneity that typifies this class,

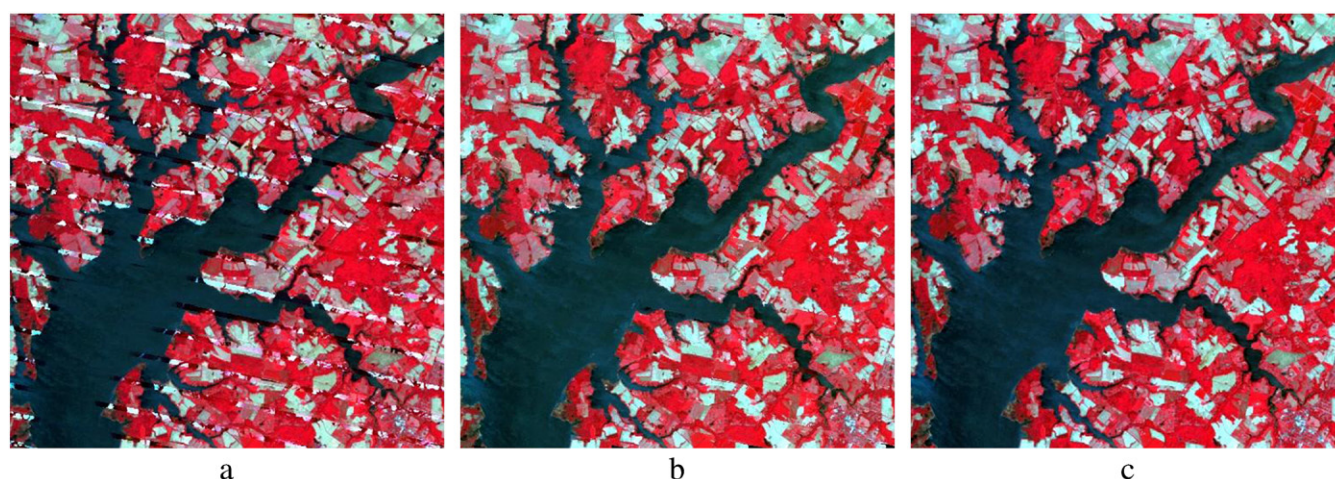


Fig. 11. Results of gap-fill for the simulation test of using multiple input images. Panels (a) and (b) are the filled images of Fig. 5f by local linear histogram matching and NSPI respectively, and (c) is the actual image.

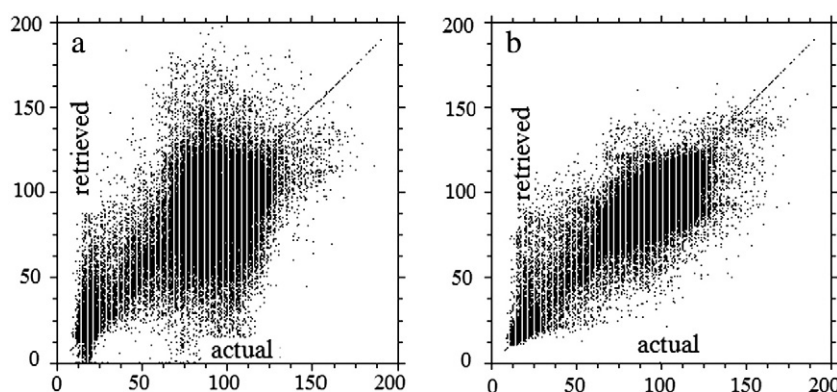


Fig. 12. Scatter plots of the real and the predicted value of all un-scanned pixels. Panels (a)–(b) are the scatter plots of Fig. 11a and b respectively. Axes denote top-of-atmosphere reflectance.

covered 4149.4 km² in the gap-fill classification (13.5% of the scene), as compared with 4236.2 km² for the reference classification (13.8% of the scene). Comparison between the gap-filled classification results versus the classification results from the actual data using standard accuracy assessments (Congalton, 1991) indicated that the two classification products are very comparable (Table 6). Overall accuracy values of 90.5% and 91.8% for the gap-filled and actual data classifications, respectively. For just those pixels within the gap areas, accuracy was 90.8% from the NSPI-generated land cover classification as compared to 92.7% from the classification generated using the actual data (Table 7). This supports the assertion that gap-filled results derived from NSPI can be used in support of normal applications such as land cover mapping.

Table 4
The accuracy of filled results of Fig. 5f using multiple input images.

Band	Gap-fill method	Filled result of Fig. 5f	
		RMSE	AD
Green	Histogram matching	0.3128	−0.1152
	NSPI	0.0107	−0.0004
Red	Histogram matching	0.3338	−0.1272
	NSPI	0.0166	−0.0005
NIR	Histogram matching	0.0771	−0.0014
	NSPI	0.0357	−0.0002

5. Conclusion and discussion

Despite the SLC failure, the quality of the radiometry and geometry of the Landsat 7 ETM+ data is still excellent for many applications (Roy et al., 2010; Chander et al., 2010). Therefore, it is very appropriate and necessary to develop techniques to fill in the un-scanned gaps in the SLC-off imagery. Unfortunately, existing gap filling methods have limitations. This paper proposes a new and effective method to fill the un-scanned gaps in Landsat 7 SLC-off ETM+ images. The NSPI can restore the value of un-scanned pixels very accurately, especially for heterogeneous landscapes and when there is a longer time interval between the input image and target image.

The major improvement of NSPI is that it makes better use of appropriate and relevant information of the scanned pixels located in close proximity to the gaps. First, according to Eq. (8), the similar pixels in the neighborhood of the gap are used to provide the difference information between the input image and the target image, allowing radiometric differences between two images to be taken into account for each land cover type. This is in part why NSPI is more effective when the time interval between the input and target images is longer. Second, NSPI selects a similar number of pixels over a certain sample size (M), which helps create greater statistical reliability, especially in heterogeneous areas. Third, a local rule based on spectral similarity is employed to search similar pixels, which helps to ensure that the scanned pixels belonging to the same land cover features from both input and target are being used. In comparison, the multi-scale segment method (Maxwell et al., 2007) uses only one scanned

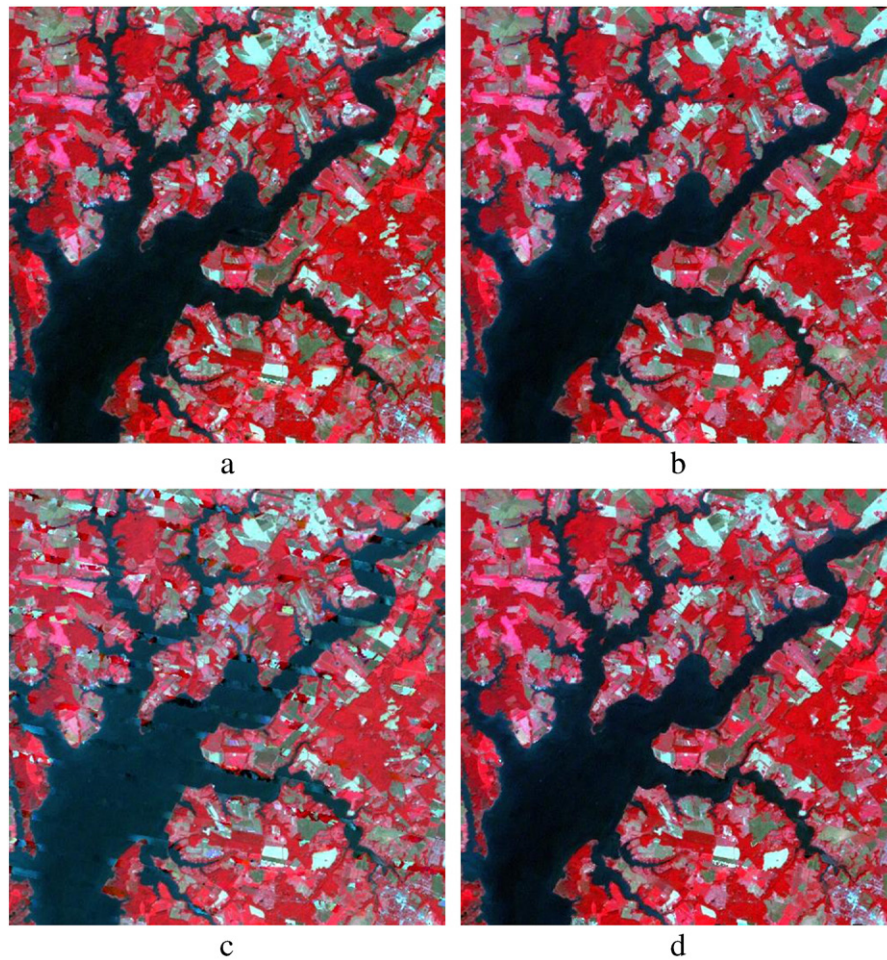


Fig. 13. Results of gap-fill for the actual SLC-off image. Panels (a) and (b) are the filled images of Fig. 7c using Fig. 5b by local linear histogram matching and NSPI, respectively; Panels (c) and (d) are the filled images of Fig. 7c using Fig. 7a and b by local linear histogram matching and NSPI, respectively.

pixel located outside the gaps in the target image. The pixel selected to replace the target pixel is the one that is the closest spatially to the un-scanned target pixel. Use of only one pixel does not statistically ensure the replacement with an appropriate pixel value. Similar to NSPI, the local linear histogram matching method selects the common pixels within a neighborhood, but it does not take spectral similarity of these common pixels into account. In addition, the histogram matching method just utilizes these common pixels as training data to build a re-scaling model, and then this model is employed to restore the value of the target pixels by only using the value in the input image as model input. If the values of common pixels are within a small range (as often happens), a linear regression model is difficult to build with high statistical significance. Accordingly the uncertainties included in the regression model can be correspondingly large.

Table 5

Aerial coverage of five land cover classes derived from classification of NSPI gap-fill data versus reference data for the entire Landsat scene. First value represents number of square km; the second represents the percentage of total area.

Land cover class	NSPI gap-fill classification	Reference classification
Water	4987.4 (16.2%)	4986.4 (16.2%)
Agricultural fields/grass	11353.6 (36.9%)	11200.0 (36.4%)
Forest	9818.3 (31.9%)	9882.5 (32.1%)
Urban	4149.4 (13.5%)	4236.2 (13.8%)
Wetlands (salt marsh)	459.8 (1.5%)	463.4 (1.5%)
Total	30768.5 (100%)	30768.5 (100%)

Another strength of NSPI is that it can help ensure the spatial continuity of the filled results. For the similar pixels located near the gaps, their spectral values are likely to be very close to the target un-scanned pixel. Therefore, the value of the target pixel is also estimated by the weighted sum of similar pixels in the target image according to Eq. (7). Then, this prediction is combined with the prediction by Eq. (8), which can help create consistency between the predicted un-scanned pixels and the scanned pixels outside the gaps.

One additional strength of NSPI is that its principle is conceptually simple. Although NSPI needs more computing time than a number of other approaches, it is not particularly computationally intensive. While initial versions of NSPI required too much computer memory to run on an entire Landsat scene, this was solved by dividing the image into 1000×1000 pixel blocks, and then mosaicking the blocks together. However, to ensure that optimal pixel values are used for gap filling, the computation window size needs to be gradually increased to locate the “best” neighborhood pixels, and there are multiple calculations that need to be done to fill each target pixel. This can require more computing time than other methods. Nonetheless, NSPI can be applied to process large amounts SLC-off ETM+ images in production sequence, if necessary.

We also recognize that there are also several potential limitations regarding this NSPI. The approach does require the availability of one or more reasonably clear ancillary TM or ETM+ image(s). Frequent cloud cover will likely hinder the use of NSPI, and thus we suspect problems in using it in some areas, such as in humid tropical forest ecosystems. In addition, we are unclear how much changing land cover type between input and target acquisition dates will impact final results.

Table 6
Standard accuracy assessment of land cover classifications from (1) NSPI gap filled data using simulated SLC-off data (first number) and (2) actual imagery (number in bold). Number of assessment points was 243. Overall classification accuracies were 90.5% and 91.8% for gap-filled versus reference data sets, respectively.

Class	Reference totals		Classified totals		Number correct		Producer's accuracy		User's accuracy	
Water	52	52	50	50	49	49	94.2%	94.2%	98.0%	98.0%
Agriculture and Grass	62	62	76	69	59	58	95.2%	93.6%	77.6%	84.1%
Forest	82	82	78	84	76	80	92.7%	97.6%	97.4%	95.2%
Urban	44	44	36	37	34	35	77.3%	79.6%	94.4%	94.6%
Wetlands	3	3	3	3	2	1	66.7%	33.3%	66.7%	33.2%

Table 7
Standard accuracy assessment of land cover classifications from just the gap areas. The first number relates to accuracies of land cover data generated from NSPI gap filled simulated SLC-off data, whereas the second (in bold) refers to accuracies developed using the actual imagery. Number of assessment points was 206. Overall classification accuracies within the gap areas were 90.8% and 92.7% for gap-filled versus reference data sets, respectively.

Class	Reference totals		Classified totals		Number correct		Producer's accuracy		User's accuracy	
Water	47	47	45	43	45	43	95.7%	91.5%	100.0%	100.0%
Agriculture and grass	67	67	85	75	67	66	100.0%	98.5%	78.8%	88.0%
Forest	74	74	61	70	61	68	82.4%	91.9%	100.0%	97.1%
Urban	15	15	12	14	12	12	73.3%	80.0%	91.7%	85.7%
Wetlands	3	3	3	4	3	2	100.0%	66.7%	100.0%	50.0%

Another issue worth mentioning is that the NSPI approach uses spectral similar pixels for the prediction and determines weights based on spectral similarity and spatial distance. It is a type of deterministic linear interpolation approach that cannot produce statistical uncertainty for each prediction. Conversely, a geostatistical interpolation approach based on spatial autocorrelation can produce variance for each prediction and this may be appropriate for the gap-filling process. We will explore the geostatistical approach to determine if it can be incorporated in the NSPI processing in future.

At present, selection of the best input image is done on a case by case basis, but should NSPI become more widely used, we believe that there will be number of ways in which the image selection process could be automated. For the test scenes used in this investigation, we did not see any obvious flaws in the scenes that were gap-filled scenes using the NSPI. However, we recognize that more tests will be necessary to assess the robustness of NSPI for gap filling across a wide array of land cover types and conditions.

Acknowledgements

This study was supported by 863 project (2009AA12004) from Ministry of Science and Technology of China. We would also like to acknowledge the support from the USGS Landsat Science Team. We thank Tom Maersperger, Kurtis Nelson and an anonymous reviewer 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.
- 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.
- 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.
- Chander, G., Xiong, X., Choi, T., & Angal, A. (2010). Monitoring on-orbit calibration stability of the Terra MODIS and Landsat 7 ETM+ sensors using pseudo-invariant test sites. *Remote Sensing of Environment*, 114, 925–939.
- Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 sensors. *Remote Sensing of Environment*, 113, 893–903.
- 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.
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 2207–2218.
- Hansen, M. C., Roy, D. P., Lindquist, E., Aducci, B., Justice, C. O., & Altstatt, A. (2008). A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sensing of Environment*, 112, 2495–2513.
- Healey, S. P., Cohen, W. B., Yang, Z. Q., & Kravinka, O. N. (2005). Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment*, 97, 301–310.
- 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.
- 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.
- 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.
- 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.
- Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalsky, V., Hansen, M., et al. (2010). Web-enabled Landsat data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114, 35–49.
- 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.
- Teillet, P. M., Barker, J. I., 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.
- USGS (2004). Phase 2 gap-fill algorithm: SLC-off gap-filled products gap-fill algorithm methodology. landsat.usgs.gov/documents/L7SLCGapFilledMethod.pdf (Available online at (accessed 28 November 2010)).
- Vogelmann, J. E., Howard, S. M., Yang, L. M., Larson, C. R., Wylie, B. K., & Van Driel, N. (2001). Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and Ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, 67, 650–662.
- 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.