2

An Automatic System for Reconstructing High-Quality Seasonal Landsat Time Series

Xiaolin Zhu, Eileen H. Helmer, Jin Chen, and Desheng Liu

CONTENTS

2.1 Introduction

Seasonal time series data from satellites are highly desired by researchers from different fields to study our Earth system. Seasonal time series data contain the temporal aspects of natural phenomena on the land surface, which are extremely helpful for discriminating different land cover types (Zhu and Liu, 2014), monitoring vegetation dynamics (Shen et al., 2011), estimating crop yields (Johnson et al., 2016), assessing environmental threats (Garrity et al., 2013), exploring human-nature interactions (Zhu and Woodcock, 2014a), and revealing ecology-climate feedbacks (Piao et al., 2015).

Since 2008, all Landsat images, archived and newly acquired, have been available at no charge for end users, stimulating the studies of land surface dynamics with seasonal Landsat time series, because these data have a spatial resolution appropriate for heterogeneous land surfaces, such as urban areas (Schneider, 2012; Zhou et al., 2014). However, like all optical satellite images, Landsat images are contaminated by clouds and cloud shadows (Ju and Roy,

2008). In addition, since May 2003, the scan-line corrector (SLC) of the Enhanced Thematic Mapper plus sensor (ETM+) on Landsat 7 has failed permanently. It causes roughly 22% of the pixels to be unscanned in any ETM+ image (referred to as SLC-off images) after 2003 (Chen et al., 2011). These sources of image contamination (i.e., clouds, cloud shadows, and SLC gaps) severely hinder Landsat time series applications and thus must be removed and then filled with data predicted from a different date, from nearby pixels or from both to reconstruct a high quality seasonal time series. Image reconstruction is especially important in cloudy regions such as tropical, subtropical, and high altitude regions.

Generally, reconstructing high quality Landsat time series involves two steps: screening clouds and cloud shadows and interpolating contaminated pixels. For screening clouds and cloud shadows in Landsat images, some promising methods have been developed, including Function of mask (Fmask) (Zhu and Woodcock, 2012), time series filtering method (Goodwin et al., 2013), and multiTemporal mask (Tmask) (Zhu and Woodcock, 2014b). Now, Fmask is used by the United States Geological Survey (USGS) to produce a cloud mask layer for end users. This product can be directly used when reconstructing Landsat time series. For interpolating contaminated pixels in Landsat images, existing techniques can be grouped into three categories: spatial interpolators (Cheng et al., 2014; Meng et al., 2009), temporal interpolators (Melgani, 2006; Zhu et al., 2015), and spatiotemporal interpolators (Chen et al., 2011; Zhu et al., 2012a,b).

Spatial interpolators search for one or several uncontaminated pixels that likely have the same land cover and that have very similar spectral values as the contaminated pixel, referred to as similar pixels. This searching process requires cloud-free ancillary images acquired at different times. Then, the contaminated pixel is replaced by similar pixels. Meng et al. (2009) developed a closest spectral fit (CSF) method to replace spectral values of cloudy pixels by cloud-free pixels using location-based one-to-one correspondence and spectral-based closest fit. Cheng et al. (2014) applied a pixel-offset based Markov random field global function to find the most suitable similar pixels to replace contaminated pixels.

Temporal interpolators estimate values of contaminated pixels through the temporal relationship among images acquired at different time points. In general, there are two ways to model the temporal relationship: as a function of cloud-free observations in the time series sequence (Melgani, 2006) or as a function of observation time (Zhu et al., 2015). To interpolate cloudy pixels in one image of a time series, the contextual multiple linear prediction (CMLP) method developed by Melgani (2006) uses a multiple linear function to model the relationship between the pixel values of the image being reconstructed and other images in the sequence. CMLP can only interpolate cloudy images one at a time in the sequence because it needs to build the temporal model for each cloudy image. Zhu et al. (2015) employed harmonic models to fit the temporal pattern of each pixel in a time series sequence. The temporal model is a function of the satellite image acquisition time, and its coefficients are

estimated from all cloud-free points in the time series sequence. This trained temporal model delineates the seasonality of each pixel. It can be used to interpolate cloudy pixels in all images of the time series simultaneously.

Spatiotemporal interpolators integrate both spatial and temporal information into the prediction of contaminated pixels (Chen et al., 2011; Zhu et al., 2012a,b). Chen et al. (2011) proposed a Neighborhood Similar Pixel Interpolator (NSPI) method to interpolate the gaps caused by the SLC problem. NSPI first searches for similar pixels and then it uses these similar pixels as samples to predict gap pixels in two ways. A spatial prediction is obtained from a weighted average of similar pixel values in the target image, that is, the image to be constructed, to predict the gap pixels. A temporal prediction is obtained by retrieving the temporal change from similar pixel values in both the target image and ancillary images. Then the two predictions are combined based on confidence measures to get the final prediction of gap pixels. The NSPI method was further modified (i.e., MNSPI) for restoring the spectral values of cloudy pixels by considering the difference between narrow wedge-shaped SLC-off gaps and cloud, that is, clouds are randomly shaped clusters and most are much larger than SLC-off gaps (Zhu et al., 2012a). The temporally predicted result is given an increasingly larger weight as the spatial distance of the pixel from the cloud edge increases. The NSPI method was further improved through incorporating Geostatistics, which can help to estimate the uncertainties of interpolation results and reduce the number of empirical parameters (Zhu et al., 2012b).

In general, spatiotemporal interpolators have more advantages compared with either spatial interpolators or temporal interpolators alone. Limitations of spatial interpolators include: (1) accuracy decreases fast with the size of cloud patches; (2) spatial continuity of interpolated images cannot be well reserved in heterogeneous landscapes. Spatiotemporal interpolators use the corresponding pixel information from cloud-free ancillary images, so it is less affected by the size of cloud patches. The weighted combination of two predictions in spatiotemporal interpolators ensures the spatial and radiometric continuity of the reconstructed image in heterogeneous landscapes (Chen et al., 2011). Limitations of temporal interpolators are high uncertainties when applied in very cloudy regions. Temporal interpolators use cloud-free points in the time series to model a temporal pattern for each pixel, then this temporal pattern is used to predict values of cloudy points. However, in areas with persistent clouds, such as tropical regions, there are no adequate cloud-free points in the time series to build a reliable temporal model. Thus, it leads to large errors in the reconstructed seasonal time series. In contrast, spatiotemporal interpolators also use spatial information which is a valuable complement when the temporal information is not reliable.

However, the original NSPI and MNSPI were developed for interpolating gap pixels and cloudy pixels, respectively, in individual images. They are not efficient for reconstructing seasonal Landsat time series when clouds, cloud shadows, and SLC-off gaps exist in most images in a time series. To

this end, this chapter introduces an automatic system for interpolating all types of contaminated pixels in all Landsat images of a time series through integrating NSPI and MNSPI into an iterative process. The input data of this system are a time series of Landsat images, along with associated cloud and cloud shadow masks, and the output is a time series with the same Landsat images but without missing pixels caused by clouds, cloud shadows, and SLC-off gaps. This system will promote the use of Landsat time series to monitor land surface dynamics.

2.2 Methods

Figure 2.1 shows the flowchart of the proposed automatic system for reconstructing Landsat time series. It requires two input data sets. First, multitemporal original Landsat images are collected to composite a time series. These original Landsat images can be either the Level 1 Terrain corrected (L1T) product or land surface reflectance product downloaded from the USGS Earth Explorer system. If a Landsat image is covered by a lot of clouds, cloud shadows, and SLC-off gaps, its useful information is very limited. Therefore, only Landsat images with more than 40% clear pixels of the entire scene are included in the time series. All these selected Landsat images are stacked by their acquisition dates. Considering that original Landsat images have

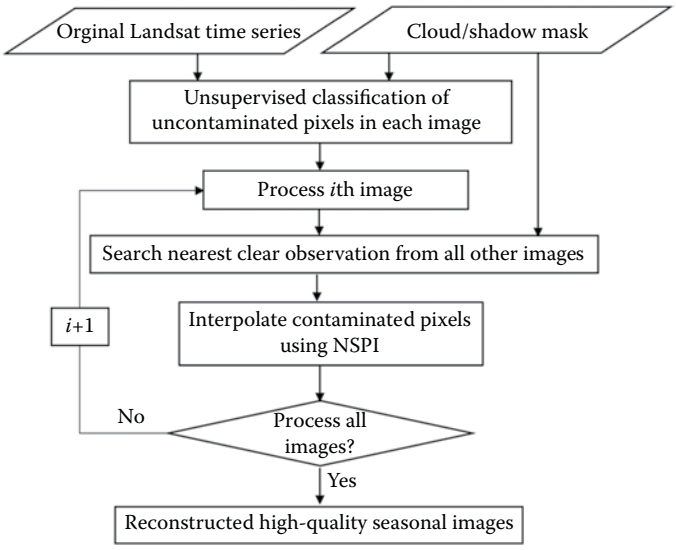

FIGURE 2.1
Flowchart of the automatic system for reconstructing high-quality seasonal Landsat time series.

been geo-rectified and calibrated accurately (Zhu and Woodcock, 2014b), no further preprocessing steps are needed. Second, a mask product is needed to indicate the contaminated pixels in the time series. Cloud masks of Landsat images are available from the USGS Earth Explorer system. Although the cloud mask produced by the Fmask algorithm has acceptable accuracy in many places, the omission of clouds and cloud shadows will lead to errors in the interpolation results. Therefore, two extra steps are used to reduce the omission errors in cloud mask: (1) considering that the omission errors often happen at the edges of cloud patches, a buffer of 1 or 2 pixels can be added to the cloud patches; (2) manual editing is needed if there are still contaminated pixels not marked in the mask. As Figure 2.1 shows, there are 3 main steps for the proposed automatic system. The details will be described below.

2.2.1 Classify Uncontaminated Pixels in Each Image

The NSPI method assumes that pixels with the same class should have high spectral similarity and similar temporal changing pattern (Chen et al., 2011). To speed up the searching of spectrally similar pixels in the original NSPI method, all input images in the time series are classified based on the spectral similarity of pixels. Here, to ensure the automation of the whole system, we use a classic unsupervised classifier, K-means, to classify the uncontaminated pixels in each image. The k-means method uses an iterative procedure. At each iteration, each pixel is assigned to one class based on the closeness with the class means obtained from the last iteration, and then class means are updated using a new class labels of pixels. The iterative process will be ended when the class labels no longer change (Lloyd, 1982). Number of classes, k, is an important parameter in K-means classification. It can be determined empirically based on the study area and the same value is used for all of the images. In most cases, the number of classes ranges from 3 to 10. Specifically, urban areas often have 7–10 land cover types, rural areas have 4–6 land cover types, and areas without human activities have 3–4 land cover types. The final result is a classification map of uncontaminated pixels for each of the original Landsat images.

2.2.2 Select Ancillary Data for Each Contaminated Pixel from the Time Series

In the NSPI method, to interpolate the value of a contaminated pixel, which is named as a target pixel, a clear observation in other images corresponding to this target pixel is needed. This clear observation is named as ancillary data. As a result, to interpolate all contaminated pixels in any one image of the time series, we need to select the best ancillary data from the time series for each contaminated pixel. Assuming that there are K images in the time series, we are interpolating the ith image ($i = 1, \ldots, K$). The searching process of ancillary data for all contaminated pixels in the ith image is as

follow: (1) sort all other *K*-1 images in the time series based on the temporal closeness to the *i*th image; (2) start from the image closest to the *i*th image, clear pixels in this image corresponding to contaminated pixels in the *i*th image are selected as ancillary data of these contaminated pixels; (3) for the remaining contaminated pixels without ancillary data, do the above selecting process in the next closest image; (4) repeat the searching process until all contaminated pixels have ancillary data selected.

2.2.3 Interpolate Contaminated Pixels by NSPI

There are some differences between the NSPI method for SLC-off gap filling (Chen et al., 2011) and MNSPI for cloud removal (Zhu et al., 2012a). To integrate NSPI and MNSPI in one system for interpolating pixels contaminated by both clouds and SLC-off gaps, the original NSPI method needs some modifications. Below will briefly describe the steps of NSPI interpolation, while more details can be find in Chen et al. (2011) and Zhu et al. (2012a).

Let us assume that a contaminated pixel (x, y) in a Landsat image to be interpolated was acquired at t_2, and its corresponding ancillary data were provided by a Landsat image acquired at t_1. The first step is searching for similar pixels based on the ancillary data. The original NSPI and MNSPI use different strategies to search for similar pixels based on the different properties of SLC-off gaps and clouds. Here, a new strategy appropriate for both gaps and clouds is used. As shown in Figure 2.2, a target pixel contaminated by clouds is to be interpolated. First, using the spectral classification map of

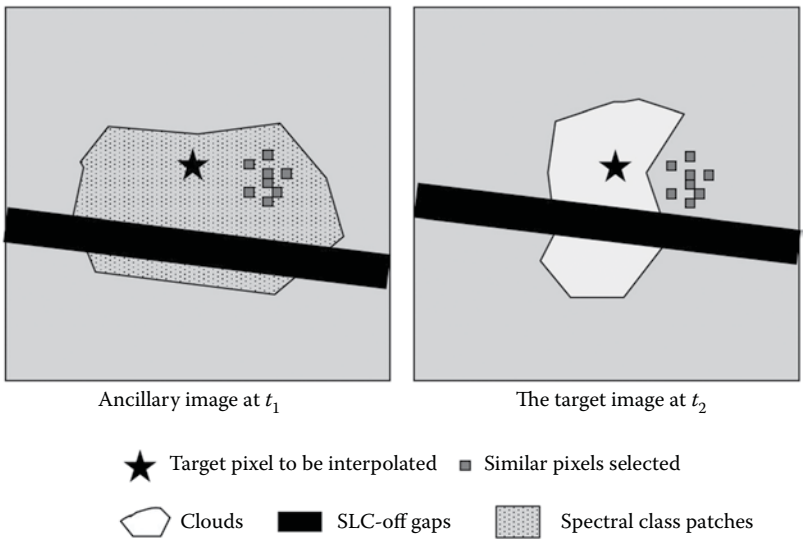

FIGURE 2.2
Schematic diagram of the similar pixel selection.

the ancillary image, pixels that are clear in both the target image and the ancillary image and that are the same class as the target pixel are selected by an adaptive window procedure (Chen et al., 2011) as the candidates for similar pixels. Second, the N (e.g., 20) pixels that are most similar to the target pixel are selected as similar pixels from the candidates. Here, the similarity is defined as the root mean square deviation (*RMSD*) between each candidate pixel and the target pixel as Equation 2.1:

$$RMSD_i = \sqrt{\frac{\sum_{b=1}^{n}(L(x_i, y_i, t_1, b) - L(x, y, t_1, b))^2}{n}}, \tag{2.1}$$

where $L(x_i, y_i, t_1, b)$ is the band b value of ith candidate pixel located in (x_i, y_i) in the ancillary image acquired at t_1, $L(x, y, t_1, b)$ and has the same definition but for a target pixel, and n is the number of spectral bands.

The similar pixels are given different weights when they are used to predict the value of the target pixel. The weight W_j determines the contribution of the jth similar pixel for predicting the value of the target pixel. This is determined by the spatial distance and the spectral similarity between the similar pixel and the 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. Therefore, the weight of the jth similar pixel, W_j, can be calculated by the following equations:

$$D_j = \sqrt{(x_j - x)^2 + (y_j - y)^2}, \tag{2.2}$$

$$CD_j = RMSD_j \times D_j, \tag{2.3}$$

$$W_j = \frac{(1/CD_j)}{\sum_{j=1}^{N}(1/CD_j)}. \tag{2.4}$$

The range of W_j is from 0 to 1, and the sum of all similar pixel weights is 1.

Then, two initial predictions using the spatial information and temporal information can be calculated, respectively. First, since the similar pixels have the same or approximate spectral value as the target pixel when they are observed at the same time, we can use the information of these similar pixels in the image at t_2 to predict the target pixel. Accordingly, the weighted average of all the similar pixels in the target image is used to make the first prediction for the target pixel:

$$L_1(x, y, t_2, b) = \sum_{j=1}^{N} W_j \times L(x_j, y_j, t_2, b). \tag{2.5}$$

Secondly, for the target pixel, the value at t_2 equals the sum of the value at t_1 and the change from t_1 to t_2. Because the value at t_1 can be obtained directly from the ancillary 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 temporal pattern as the target pixel. Accordingly, the weighted average of the change provided by all the similar pixels is used to calculate the value of the target pixel as the second prediction:

$$L_2(x,y,t_2,b) = L(x,y,t_1,b) + \sum_{j=1}^{N} W_j \times (L(x_j,y_j,t_2,b) - L(x_j,y_j,t_1,b)). \quad (2.6)$$

Last, a weighted combination of the two initial predictions is used to compute the final prediction. The weights (T_1 and T_2) are determined by the extent of spatial continuity and the extent of temporal continuity between the ancillary image and the target image estimated from similar pixels. Here, the averaged *RMSD* (R_1) between the similar pixel and the target pixel is used to denote the extent of the spatial continuity:

$$R_1 = \frac{1}{N}\sum_{j=1}^{N}\sqrt{\frac{\left[\sum_{b=1}^{n}(L(x_j,y_j,t_1,b) - L(x,y,t_1,b))^2\right]}{n}}. \quad (2.7)$$

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 temporal continuity between the input image and the target image:

$$R_2 = \frac{1}{N}\sum_{j=1}^{N}\sqrt{\frac{\left[\sum_{b=1}^{n}(L(x_j,y_j,t_1,b) - L(x_j,y_j,t_2,b))^2\right]}{n}}. \quad (2.8)$$

Then, the normalized reciprocal of R_1 and R_2 is used as the weight T_1 and T_2 respectively:

$$T_i = \frac{(1/R_i)}{(1/R_1 + 1/R_2)}, \quad \text{where} \quad i = 1,\ 2. \quad (2.9)$$

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

$$L(x,y,t_2,b) = T_1 \times L_1(x,y,t_2,b) + T_2 \times L_2(x,y,t_2,b) \quad (2.10)$$

The above interpolation process is implemented for all contaminated pixels in each image of the time series until all images are reconstructed.

2.3 Experiments

To evaluate the feasibility and effectiveness of the proposed automatic system, two experiments were carried out: a real cloudy time series reconstruction and a simulated cloudy time series reconstruction.

For the experiment with the real cloudy time series reconstruction, 15 Landsat-7 images covering Mona Island of Puerto Rico (a size of 400×400 pixels) were used. Mona Island lies 66 km west of Puerto Rico. It is a mainly flat limestone plateau surrounded by sea cliffs and covered by tropical dry forests. As is typical in tropical regions, most Landsat images of this island are cloudy. In this experiment, we used all 15 Landsat-7 images acquired from 2008 and 2009 with less than 60% of pixels contaminated by clouds, cloud shadows, and SLC-off gaps. These images were stacked according to their dates (i.e., day of year) to composite a one-year time series (Table 2.1). From Table 2.1, the proportion of contaminated pixels in images varies from 23.67% to 57.29%. Figure 2.3 shows a false-color composite of 3 images with low, medium, and high proportions of contaminated pixels. It is a common issue for existing methods that the accuracy of image reconstruction decreases with

TABLE 2.1

Summary of Landsat-7 Images in Mona Island Site

ID	Year	Day of Year	Contaminated Pixels %
1	2008	11	26.37
2	2008	27	25.96
3	2008	43	25.25
4	2009	61	32.89
5	2009	77	37.62
6	2008	91	25.68
7	2009	109	46.68
8	2008	139	57.29
9	2008	171	48.47
10	2008	203	27.48
11	2009	253	53.11
12	2009	301	24.23
13	2009	317	46.68
14	2009	333	27.63
15	2009	349	23.67

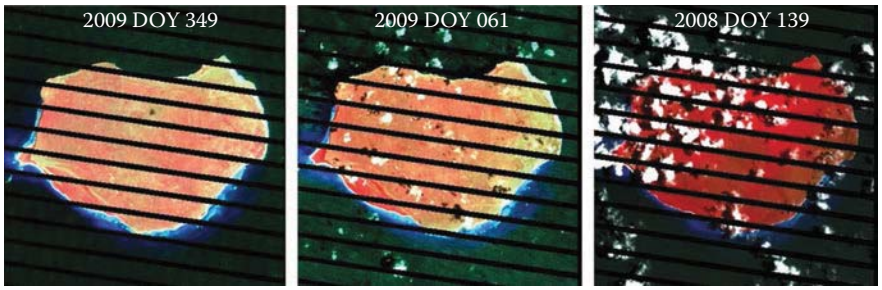

FIGURE 2.3
False-color composite of 3 Landsat-7 images in Mona Island with low, medium, and high proportion of contaminated pixels (from left to right).

the proportion of missing pixels, so the different proportions of contaminated pixels in this site are good for testing whether or not the performance of the proposed method is acceptable for images with large proportions of contaminated pixels.

To more quantitatively assess the accuracy of seasonal image reconstruction, another study site with a lot of cloud-free Landsat images was used to implement a simulation study. Fifteen cloud-free Landsat 7 images were provided by Emelyanova et al. (2013). This site is in southern New South Wales, Australia, and has a heterogeneous landscape. The major land cover types in this area are irrigated rice cropland, dryland agriculture, and woodlands. These images were acquired during 2001 October to 2002 May. Rice croplands are often irrigated in October-November, which leads to large temporal changes in the time series. Both high temporal change and high heterogeneity challenge the reconstruction of contaminated pixels, so this data set is ideal for testing the effectiveness of the proposed method. We cut all images to the image size of Mona Island for this experiment. Cloud masks from the Mona Island site were overlaid onto all 15 cloud-free images in this site to produce pseudo cloudy images. Figure 2.4 shows one image in this site and its corresponding simulated contaminated image using the mask from a Mona Island image shown in Figure 2.3 (the right one).

Then, the proposed method was applied to these pseudo cloudy images. The accuracy was evaluated through the calculation of two statistical indices. The first index is the root mean square error (*RMSE*). This metric is frequently used to assess the differences between values predicted by a model and the values observed or measured. A larger *RMSE* indicates a larger prediction error. The second metric is correlation coefficient (*R*) between the actual values and predicted values of contaminated pixels. *R* is used to show the linear relationship between actual and reconstructed images. An *R* value closer to 1 indicates higher accuracy of reconstructed images.

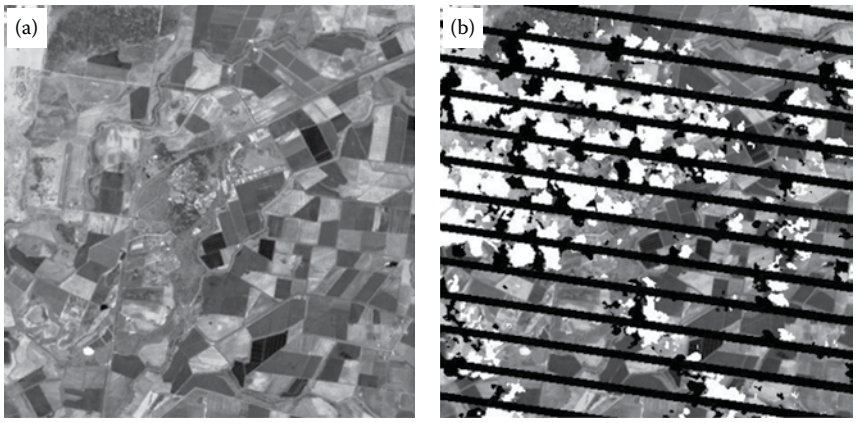

FIGURE 2.4
A Landsat-7 image acquired on February 11, 2002 (a) and its simulated contaminated image using the mask from the Mona Island image of 2008 DOY139 (b).

2.4 Results

2.4.1 Reconstruction of Real Landsat Time Series

All the 15 images in the time series were reconstructed successfully. To save space here, Figure 2.5 only shows the reconstructed results of three images in Figure 2.3, but other images have similar reconstruction results. We can see that regardless of the proportion of contaminated pixels the images have, the reconstructed images appear spatially continuous without displaying the footprints of gaps or cloud/shadow patches. The result suggests that the proposed automatic system can interpolate contaminated pixels accurately to produce high-quality time series.

Although we cannot assess the accuracy quantitatively for Mona, because the true pixel values are unknown, we can analyze the time series based

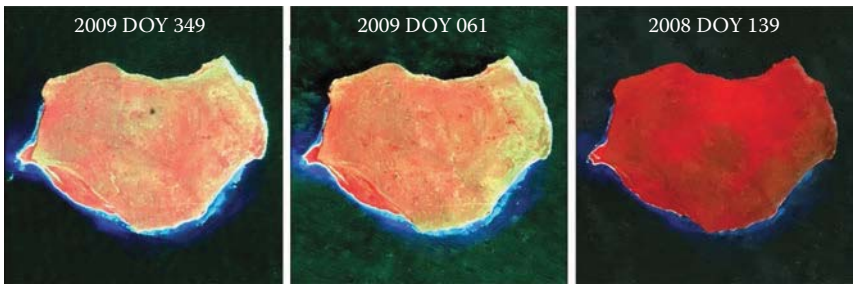

FIGURE 2.5
Reconstruction results of three contaminated images shown in Figure 2.3.

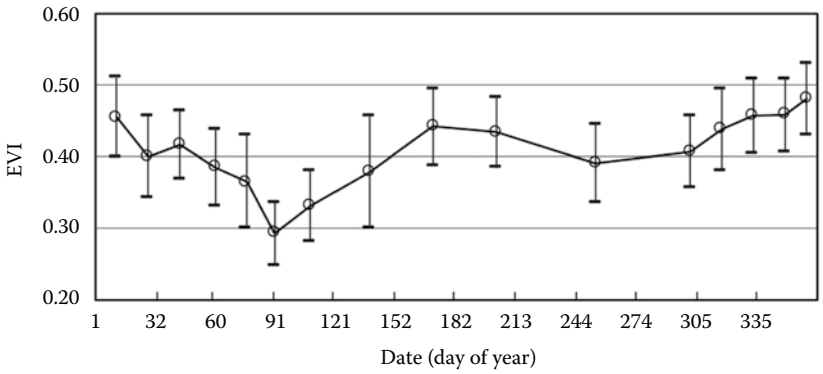

FIGURE 2.6
Enhanced vegetation index (EVI) time series of Mona Island derived from reconstructed 15 Landsat-7 images. Each circle is the mean value of all pixels in this study area and the error bar is one standard deviation.

on the prior knowledge of the vegetation growing cycle to evaluate the usefulness of the reconstructed images. Figure 2.6 shows the average enhanced vegetation index (EVI) time series (Huete et al., 2002) derived from the 15 reconstructed images over Mona Island. The dry season of Mona Island is from January to April. As can be seen from the EVI curve, the EVI values decrease during the dry season and reach the lowest point in the end of March. This temporal pattern is consistent with the local vegetation pattern as the deciduous tree species lose leaves during the dry season. The EVI curve from the reconstructed time series indicates the reliability of the proposed method. It can help us to monitor vegetation seasonality at high frequency.

2.4.2 Reconstruction of Simulated Landsat Time Series

Quantitative accuracy assessment of reconstruction results of all 15 simulated contaminated images are reported in Table 2.2. The proposed method was used to reconstruct all 6 bands with 30-m resolution. To save space and because results from the two shortwave bands are similar to those for the first four bands, only results from the first 4 bands were listed in Table 2.2. *R* values of all images are higher than 0.8, and most RMSE values are less than 0.02, suggesting that the proposed automatic system interpolated all contaminated pixels very accurately. In addition, most existing methods do not usually yield acceptable results when there are large proportions of contaminated pixels. Figure 2.7 shows the scatter plot between the accuracy (both RMSE and R) of near infrared (NIR) band and the proportion of contaminated pixels. No clear relationship exists in the scatter plot, suggesting that the proposed method yields a robust result, regardless the proportion of pixels contaminated in an image.

TABLE 2.2

RMSE and R Values of Reconstructed 15 Simulated Landsat-7 Cloudy Images

				RMSE				R			
ID	**Year**	**DOY[a]**	**CP %[b]**	**Band1**	**Band2**	**Band3**	**Band4**	**Band1**	**Band2**	**Band3**	**Band4**
1	2001	17	26.37	0.0075	0.0092	0.0141	0.0300	0.90	0.90	0.91	0.96
2	2001	33	25.96	0.0080	0.0104	0.0152	0.0304	0.88	0.88	0.90	0.95
3	2001	40	25.25	0.0081	0.0105	0.0165	0.0359	0.84	0.84	0.86	0.90
4	2001	56	32.89	0.0073	0.0096	0.0145	0.0230	0.91	0.92	0.93	0.94
5	2001	65	37.62	0.0098	0.0146	0.0231	0.0319	0.89	0.90	0.91	0.91
6	2002	97	25.68	0.0081	0.0117	0.0183	0.0273	0.96	0.96	0.97	0.93
7	2002	104	46.68	0.0094	0.0142	0.0219	0.0285	0.94	0.94	0.95	0.92
8	2002	136	57.29	0.0098	0.0131	0.0193	0.0281	0.91	0.91	0.92	0.91
9	2002	145	48.47	0.0092	0.0118	0.0182	0.0296	0.89	0.89	0.90	0.91
10	2002	168	27.48	0.0082	0.0109	0.0148	0.0250	0.90	0.88	0.92	0.95
11	2002	184	53.11	0.0074	0.0089	0.0130	0.0255	0.91	0.90	0.93	0.94
12	2002	193	24.23	0.0082	0.0108	0.0162	0.0269	0.85	0.87	0.89	0.92
13	2002	200	46.68	0.0088	0.0112	0.0158	0.0262	0.84	0.85	0.88	0.91
14	2002	209	27.63	0.0075	0.0086	0.0118	0.0164	0.91	0.93	0.94	0.97
15	2002	216	23.67	0.0065	0.0078	0.0114	0.0177	0.91	0.94	0.95	0.96

[a] DOY: days starting from October 1, 2001.
[b] CP%: proportion of simulated contaminated pixels.

Figure 2.8 shows the NIR band results of two simulated contaminated images. One image has a low proportion of contaminated pixels (upper row in Figure 2.8), while another one has a high proportion of contaminated pixels (lower row in Figure 2.8). By visual comparison of both of the reconstructed images shown in Figure 2.8, the reconstructed images are very close to the actual images. We can see that all the spatial details were retrieved well,

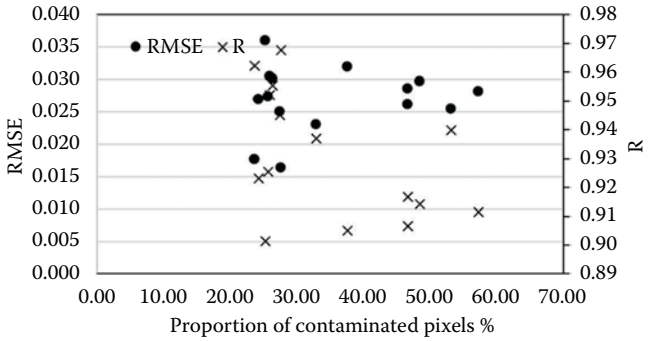

FIGURE 2.7
Scatter plot between proportion of contaminated pixels in each image and the accuracy of reconstructed images: root mean square error (RMSE) and correlation coefficient R.

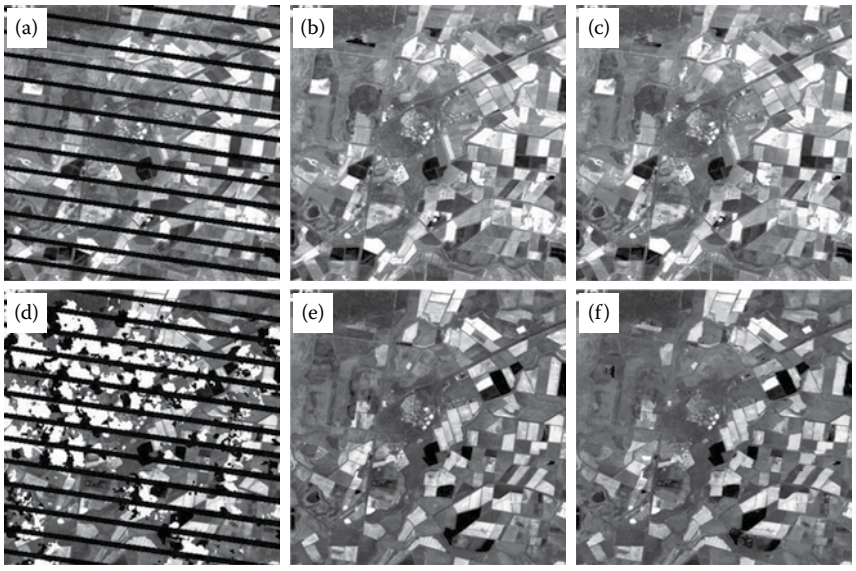

FIGURE 2.8
Images in upper row are NIR band of simulated contaminated Landsat-7 image of DOY*136 (a), the true image (b), and the reconstructed image (c). Images in lower row are NIR band of simulated contaminated Landsat-7 image of DOY*216 (d), the true image (e), and the reconstructed image (f).

indicating that the contaminated pixels interpolated by our method have high accuracy, even when the landscape is very heterogeneous.

Figure 2.9 shows the average NDVI time series derived from the 15 cloud-free Landsat-7 images compared with the time series as reconstructed with the simulated contaminated images. It is clear that the two NDVI temporal curves are very consistent with each other. Both curves describe well the temporal dynamics associated with crop phenology over a single growing season. During the irrigation period, the NDVI values decrease. Then, NDVI values increase during the crop growing period (Emelyanova et al., 2013). This temporal pattern from reconstructed images shows that the reconstructed time series can be used to monitor a fast-changing agricultural landscape.

2.5 Conclusion and Discussions

Landsat time series data have been widely used to study seasonal land surface dynamics at regional and global scales, but the SLC failure in Landsat-7 and cloud contamination reduce the chances of obtaining a

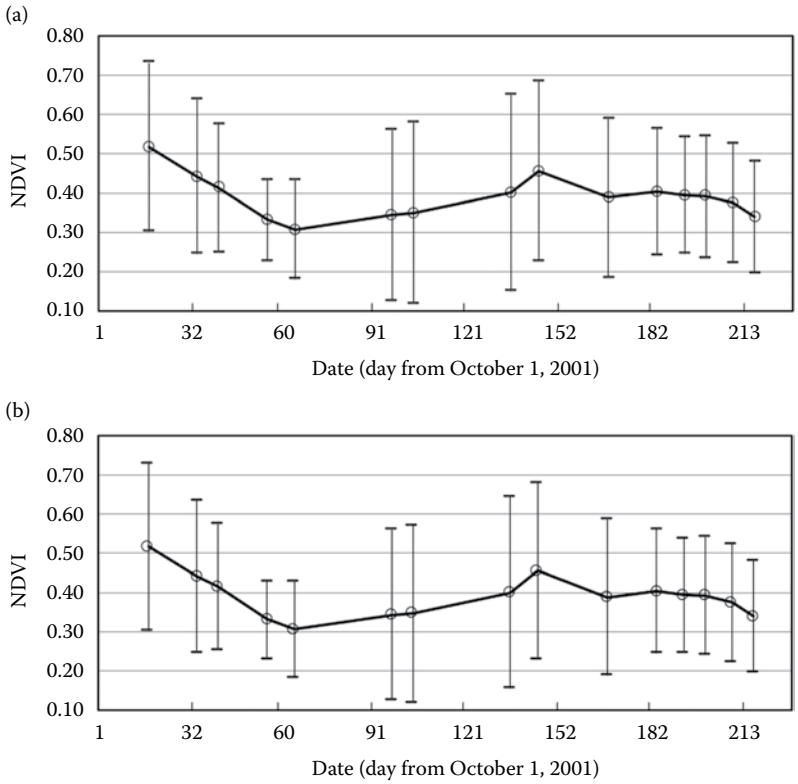

FIGURE 2.9
NDVI time series derived from original cloud-free Landsat-7 images (a) and from reconstructed results of simulated contaminated images (b). Each circle is the mean value of all pixels and the error bar is one standard deviation.

high-quality time series. Therefore, the necessary and feasible way to address this problem is to interpolate contaminated pixels to reconstruct the time series. This chapter introduces an automatic system to reconstruct all contaminated images simultaneously in a time series, which is based on the NSPI method and an iterative process. The real and simulation experiments show that the proposed method can reconstruct dense Landsat time series successfully. The vegetation indices derived from the reconstructed time series can be used to monitor the vegetation seasonality, including forest phenology and crop growing stages. The robustness of the proposed method for interpolating images with large proportions of missing pixels is very important because it is difficult to acquire good quality images in cloudy regions (Ju and Roy, 2008). Considering that the Landsat series have collected data over 40 years, the method introduced in this chapter will promote the use of seasonal Landsat time series in various disciplines, including ecology, hydrology, earth science, environmental science, agriculture, and

even sociology. Studying seasonality can advance our understanding of interactions between different Earth systems and improve our ability to predict future scenarios.

In general, the proposed method has the following strengths: (1) it makes full use of temporal and spatial information contained in all images in the time series and produces the contamination-free images in all seasons simultaneously; (2) it integrates cutting edge techniques for SCL-off gap filling and cloud removal into one system so that various types of missing pixels can be processed at the same time; (3) it is highly automated with minimal predefined parameters and human-computer interactions, so it can be user friendly for people from fields other than remote sensing. The proposed method has been successfully applied to study vegetation phenology in Puerto Rico and the US Virgin Islands (Gwenzi et al., 2017). The open-source program of the proposed method can be downloaded from the developer's website: xiaolinzhu.weebly.com.

The proposed method may face some challenges in some specific situations. First, in regions with extremely persistent clouds, it would be very difficult to reconstruct Landsat time series with acceptable accuracy because clouds totally cover many Landsat images. A possible way to solve this problem is to incorporate data from other sensors with short revisit cycles, such as MODerate resolution Imaging Spectroradiometer (MODIS) images. Currently, some spatiotemporal data fusion methods have been developed to fuse MODIS and Landsat images to increase the frequency of Landsat observations (Gao et al., 2006; Zhu et al., 2016, 2010). These techniques could be integrated into our automatic system in the future. Second, the proposed method contains a similar pixel searching process for each contaminated pixel, so it may need a lot of computing time when processing massive images over large area. Therefore, we recommend using high performance computers or parallel computing to increase the computing speed when it is used to reconstruct Landsat time series over large area and long period.

Acknowledgments

This study was supported by the Research Grants Council of Hong Kong (project no. 25222717), the National Natural Science Foundation of China (project no. 41701378), the USDA Forest Service International Institute of Tropical Forestry (Cooperative Agreement 13-CA-11120101-029), the Southern Research Station Forest Inventory and Analysis Program, the National Science Foundation of the USA (1010314), and it was conducted in cooperation with the University of Puerto Rico and the Rocky Mountain Research Station.

References

Chen, J., Zhu, X., Vogelmann, J.E., Gao, F., Jin, S., 2011. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* 115, 1053–1064.

Cheng, Q., Shen, H., Zhang, L., Yuan, Q., Zeng, C., 2014. Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model. *ISPRS J. Photogramm. Remote Sens.* 92, 54–68. doi:10.1016/j.isprsjprs.2014.02.015

Emelyanova, I. V., McVicar, T.R., Van Niel, T.G., Li, L.T., van Dijk, A.I.J.M., 2013. Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sens. Environ.* 133, 193–209. doi:10.1016/j.rse.2013.02.007

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 Trans. Geosci. Remote Sens.* 44, 2207–2218.

Garrity, S.R., Allen, C.D., Brumby, S.P., Gangodagamage, C., McDowell, N.G., Cai, D.M., 2013. Quantifying tree mortality in a mixed species woodland using multitemporal high spatial resolution satellite imagery. *Remote Sens. Environ.* 129, 54–65. doi:10.1016/j.rse.2012.10.029

Goodwin, N.R., Collett, L.J., Denham, R.J., Flood, N., Tindall, D., 2013. Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series. *Remote Sens. Environ.* 134, 50–65. doi:10.1016/j.rse.2013.02.019

Gwenzi, D., Helmer, E.H., Zhu, X., Lefsky, M.A., Marcano-Vega, H., 2017. Predictions of tropical forest biomass and biomass growth based on stand height or canopy area are improved by Landsat-scale phenology across Puerto Rico and the US Virgin Islands. *Remote Sens.* 9(2), 1–18. doi:10.3390/rs9020123

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi:10.1016/S0034-4257(02)00096-2

Johnson, M.D., Hsieh, W.W., Cannon, A.J., Davidson, A., Bédard, F., 2016. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. *Agric. For. Meteorol.* 218–219, 74–84. doi:10.1016/j.agrformet.2015.11.003

Ju, J., Roy, D.P., 2008. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sens. Environ.* 112, 1196–1211. doi:10.1016/j.rse.2007.08.011

Lloyd, S.P., 1982. Least Squares Quantization in PCM. *IEEE Trans. Inf. Theory* 28, 129–137. doi:10.1109/TIT.1982.1056489

Melgani, F., 2006. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Trans. Geosci. Remote Sens.* 44, 442–455. doi:10.1109/TGRS.2005.861929

Meng, Q., Borders, B.E., Cieszewski, C.J., Madden, M., 2009. Closest Spectral Fit for Removing Clouds and Cloud Shadows. *Photogramm. Eng. Remote Sens.* 75, 7505–7505. doi:10.14358/PERS.75.5.569

Piao, S., Yin, G., Tan, J., Cheng, L., Huang, M., Li, Y., Liu, R. et al., 2015. Detection and attribution of vegetation greening trend in China over the last 30 years. *Glob. Chang. Biol.* 21, 1601–1609. doi:10.1111/gcb.12795

Schneider, A., 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. *Remote Sens. Environ.* 124, 689–704. doi:10.1016/j.rse.2012.06.006

Shen, M., Tang, Y., Chen, J., Zhu, X., Zheng, Y., 2011. Influences of temperature and precipitation before the growing season on spring phenology in grasslands of the central and eastern Qinghai-Tibetan Plateau. *Agric. For. Meteorol.* 151, 1711–1722.

Zhou, W., Qian, Y., Li, X., Li, W., Han, L., 2014. Relationships between land cover and the surface urban heat island: Seasonal variability and effects of spatial and thematic resolution of land cover data on predicting land surface temperatures. *Landsc. Ecol.* 29, 153–167. doi:10.1007/s10980-013-9950-5

Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* 114, 2610–2623. doi:10.1016/j.rse.2010.05.032

Zhu, X., Gao, F., Liu, D., Chen, J., 2012a. A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images. *IEEE Geosci. Remote Sens. Lett.* 9, 521–525. doi:10.1109/LGRS.2011.2173290

Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sens. Environ.* 172, 165–177. doi:10.1016/j.rse.2015.11.016

Zhu, X., Liu, D., 2014. Accurate mapping of forest types using dense seasonal landsat time-series. *ISPRS J. Photogramm. Remote Sens.* 96, 1–11.

Zhu, X., Liu, D., Chen, J., 2012b. A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote Sens. Environ.* 124, 49–60.

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* 118, 83–94. doi:10.1016/j.rse.2011.10.028

Zhu, Z., Woodcock, C.E., 2014a. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* 144, 152–171. doi:10.1016/j.rse.2014.01.011

Zhu, Z., Woodcock, C.E., 2014b. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sens. Environ.* 152, 217–234. doi:10.1016/j.rse.2014.06.012

Zhu, Z., Woodcock, C.E., Holden, C., Yang, Z., 2015. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sens. Environ.* 162, 67–83. doi:10.1016/j.rse.2015.02.009