We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Making Use of the Landsat 7 SLC-off ETM+ Image Through Different Recovering Approaches

Feng Chen, Xiaofeng Zhao and Hong Ye

1. Introduction

Landsat program represents the world’s longest continuously acquired collection of space-based moderate-resolution land remote sensing data. So far, the sensors include the Landsat 1-5 Multispectral Scanners (MSS), the Landsat 5 Thematic Mapper (TM) and the Landsat 7 Enhanced Thematic Mapper Plus (ETM+). The Landsat imagery with nearly four decades (from 1972 to present, see Fig.1) provides a unique resource for researchers and common users who work in agriculture, geology, forestry, regional planning, education, mapping, and global change research. Currently, both Landsat 5 and Landsat 7 are still being operated in orbit for data collection, although they have exceeded the expected service time and always readily experienced malfunction.

The scan-line corrector (SLC) for the ETM+ sensor, on board the Landsat 7 satellite, failed permanently on May 31, 2003. Normally, the SLC compensates for the forward motion of the satellite. Under the abnormal situation, without an operating SLC, images have wedge-shaped gaps that range from a single pixel in width near the nadir point, to about 12 pixels towards the edges of the scene. The SLC anomaly results in about 22% pixels in these images are un-scanned. The consequence of the SLC failure (called SLC-off problem) hampers the use of the Landsat 7 ETM+ data. At the same time, Landsat 5 has suffered problems with its solar array drive which have affected data availability. It suggests that Landsat 5 would be at the end of its operational life and not be reliable on as a source of future imagery (Pringle et al., 2009). Additionally, TM sensor onboard the Landsat 5 has been suspended from 2011 November resulted from the failure of its electronic component vital to the data transmission, according to the news provided by the U.S. Geological Survey (USGS). The Landsat Data Continuity Mission (LDCM), to be called Landsat 8, which is regarded as a successor of the preliminary series (e.g. the Landsat 1-4, Landsat 5, and
Landsat 7), will be launched in December 2012 or January 2013 according to the passed schedule. Under the condition that no replacement of Landsat or Landsat-like sensor from the LDCM has been launched presently, along with the sensors failures mentioned above, there would be a period of data gap, probably ranging from 2003 to 2013. Fortunately, SLC-off has not affected the radiometric and geometric quality of the sensor, with approximately 80% of the pixels in each image being scanned (The USGS & National Aeronautics Space Administration [NASA], 2003). Although the SLC anomalies produce obviously negative impacts on the post-event Landsat 7 ETM+ data usability, for some applications the SLC-off ETM+ retains significant and important utility for scientific applications and some users still prefer these data over more costly alternatives (USGS & NASA, 2003). In order to resolve the SLC-off problem and enable the continuous usage of ETM+, we must find suitable way to estimate the values at un-scanned pixels.

Soon after SLC-off occurred, a report was compiled by the USGS (USGS & NASA, 2003). It suggested that, at the un-canned location, the reflectance of a particular ETM+ band could be estimated by compositing the target image, the SLC-off one need to be recovered, with the corresponding band of laterally overlapping ETM+ images, SLC-off or SLC-on ETM+ images (called fill image) of the same area from other dates (USGS & NASA, 2003). The means and the variances of the bands in the additional images can be adjusted, either globally or locally, to the means and variances of the bands in a target image (USGS & NASA, 2003). Based on this assumption, an expert group organized by the USGS proposed several recovering methodologies to tackle this issue, including global linear histogram match (GLHM), localized linear histogram match (LLHM), and adaptive window linear histogram match (AWLHM) (Scaramuzza et al., 2003; USGS & NASA, 2004). These methods are simple and easily to be put into application. Therefore, the un-scanned pixels in the SLC-off ETM+ images are able to be recovered with acceptable accuracy, given that the input scenes (the target and fill ones) are provided with high quality and represent comparable seasonal conditions (USGS & NASA, 2004).

However, in practice, there are shortcomings associated with these original methods, of which, the most obvious one is the restrict requirement on selecting the fill image (F. Chen et al., 2010). Thus, as an alternative to those methods, it is possible to use the information from a sensor other than ETM+, which is observed at the close time as a SLC-off ETM+ image, to estimate the un-scanned pixels. This assumption has been verified by several previous
researches. For example, Reza and Ali (2008) used IRS products to recover the SLC-off ETM+ image, while Boloorani et al (2008a) and F. Chen et al (2010) demonstrated the possibility of estimating un-scanned pixels in SLC-off image by using EO-1/ALI and CBERS data as auxiliary information, respectively. Furthermore, Roy et al (2008) used the information observed by MODIS to calculate reflectance of the missing pixels, although the spatial resolution of MODIS sensor is much coarser than that of Landsat 7 ETM+.

Recovering procedures based on other concepts or methodologies were also discussed separately. Multi-scale segmentation approach was developed by Maxwell et al (2007) to fill gaps in the Landsat 7 SLC-off ETM+ images. By means of the geostatistical techniques, in which kriging or co-kriging was used, not only the value of the un-scanned pixels can be estimated, but also the uncertainty of the prediction can be quantified (Zhang et al., 2007; Pringle et al., 2009). Boloorani et al (2008b) developed a methodology called ‘projection transformation’. A simple but effective method was proposed by J. Chen et al (2011) to fill the gaps, in which information about the neighbourhood similar pixels were incorporated for estimating the target un-scanned pixel. The recovered accuracy was improved consequently, especially in heterogeneous landscape areas (J. Chen et al., 2011).

As mentioned above, researches have been conducted widely for issues associated with recovering the Landsat 7 SLC-off ETM+ image, by interested individuals, institutions or organizations. However, while much attention has been paid to the reconstruction of its multispectral bands (e.g. visible and near infrared ones), few researches have been done to recover its thermal infrared band, although this band has been used widely, such as for urban heat island (X. L. Chen et al., 2006; Rajasekar & Weng, 2009), water environment (Wloczyk et al., 2006), volcanic activity (Flynn et al., 2001). Thus, the thermal infrared band of SLC-off ETM+ image should be useful for scientific applications, provided that a suitable recovering procedure is implemented (F. Chen et al., 2011). The modified AWLHM method was preliminarily proposed for recovering the thermal infrared band of Landsat 7 SLC-off ETM+ image, while the multispectral bands of CBERS were used as auxiliary data (F. Chen et al., 2011). Results illustrated this methodology’s availability for getting brightness temperature directly from the SLC-off thermal infrared band. However, overestimation and underestimation were presented in cooler and warmer areas, respectively.

Moreover, in 2008, the USGS made a decision to make the Landsat standard data products freely available through the internet, which was served as a watershed event in the history of the Landsat program. This policy has triggered more and more studies by using Landsat series data, which had been limited previously by data costs and processing requirements. As stated by Wulder & Masek (2012), a special emphasis is placed on the burgeoning scientific and applications opportunities enabled by free access to the US archive, such as use of dense time-series data to characterize inter- and intra-annual land cover changes, new capabilities for continental-scale mapping, and applications focused on particular information needs. Meanwhile, the LDCM program, including the approaching launch of Landsat 8, and the proposed Landsat 9 and Landsat 10, will make sure the future continuity of the Landsat series, which is critical for understanding the Earth system and for providing a scientific basis for land management in the future. Nevertheless, the data gap for valid
Landsat 7 ETM+ data caused by the SLC-off problem, between 2003 and 2013, may prevent the full usage of the Landsat legacy to some extent.

In conclusion, further researches concerned on recovering the Landsat 7 SLC-off ETM+ image is necessary, mainly due to valuable legacy of the Landsat series and the freely distributed policy. In this chapter, we intend to review and show several solutions for making full use of the Landsat 7 SLC-off ETM+ image based on actual conditions, followed by methods comparison and discussion on the usage of the recovered ETM+ images.

2. The recovering approaches

As mentioned in the introduction (Section 1), different recovering approaches have been proposed based on different assumptions. Fig. 2 shows an example of recovering SLC-off ETM+ image by using information from a precedent SLC-on ETM+ image. Under each ideal condition, recovered results were properly obtained, as illustrated in simulation studies. For practical application, these recovering approaches should be compared in depth, and thus the most suitable one should be selected at last.

2.1. GLHM and LLHM method

Soon after the SLC failed, experts from USGS initially provided two recovering methods, called GLHM and LLHM respectively, and they were served as the ‘Phase One Methodology’ (Scaramuzza et al., 2003). The basis for these methods is that linear relationship exists between one image and another, in view of band-by-band comparison, as shown in Fig.3. Generally, linear least squares estimate can be used to estimate the transformation parameters (e.g. gain and bias). For convenience, in the ‘Phase One Methodology’, a simple procedure was adopted for calculating gain and bias (see equation (1)).
Making Use of the Landsat 7 SLC-off ETM+ Image Through Different Recovering Approaches 321

Figure 3. Linear model for GLHM method (modified from document (Scaramuzza et al., 2003) for illustration)

\[ Y \approx gainX + bias \quad \text{with} \quad gain = \frac{\sigma_Y}{\sigma_X} \; ; \; bias = \bar{Y} - gain\bar{X} \quad (1) \]

Where, \( X \) (also denoted \( L(x_j,y_j,t_2,b) \) in following sections) is the pixel value in the fill image, and \( Y \) (also denoted \( L(x_j,y_j,t_1,b) \) in the following sections) is the pixel value in the target image; \( \sigma_X \) is the standard deviation of the fill image, and \( \sigma_Y \) is the standard deviation of the target image; \( \bar{X} \) is the mean of the fill image and \( \bar{Y} \) is the mean of the target image.

When this transformation is applied to the entire image, it is called GLHM. GLHM is a very simple method, and all un-scanned pixels are estimated using the same gain and bias. It may perform well over scenes with invariant terrain such as deserts and rocky areas, but obviously visible errors are possible in scenes with transient. It is desirable to consider the inter-regional heterogeneity caused by factors such as atmosphere, topography, and land cover. For greater precision, gain and bias for estimating the un-scanned pixels can be calculated in a moving window around each un-scanned pixel, which is the original basis of the LLHM.

2.2. AWLHM method

‘Phase Two Methodology’ is an enhancement of the ‘Phase One Methodology’, which allows users to choose multiple images and to combine SLC-off images (USGS & NASA, 2004). AWLHM is an improved version of the LLHM. The AWLHM is based on the same assumption as LLHM, except that the moving window size is changeable. To obtain the correlation (or gain and bias) statistically, the number of effective pixels used for estimating each un-scanned pixel should be considered. The moving window around each un-scanned pixel should be adaptive according to the distribution of commonly scanned pixels (Fig.4). Based on the AWLHM method, the recovered image can be obtained through iteration process over all un-scanned pixels in the target image. Iteration for one un-scanned pixel is described as follows:
a. Extract an \( n \times n \) window around the un-scanned pixel which is need to be recovered, in both target and fill images. Larger windows were found to yield essentially identical results but impose a significant performance penalty.

b. Find smallest square of pixels between the two blocks (windows in the fill image and the target one) which contains at least a minimum number of commonly scanned pixels (the valid data in both the target and fill images). Beginning with the smallest window size with 1x1 pixel, the number of valid common pixels between the target and the fill images is counted. If the minimum number of commonly scanned pixels is not met, the moving window size is expanded by one pixel on each side of the window. Then the commonly scanned pixels are counted again. The moving window is expanded gradually, with increasing size series, such as 3x3, 5x5, 7x7,... (Fig. 4). Iteration continues until the minimum number of common pixels is met. If the minimum number of common pixels is not met before reaching the maximum extent of the window, all commonly scanned pixels in the moving window are used for obtaining the transformation parameters, regardless of the total number. In the 'Phase Two Methodology', the minimum number is defined as 144.

c. Using the common pixels in the lastly determined moving window around the un-scanned pixel, to estimate the bias and gain by means of the least squares solution for the linear equations, established based on the values of all selected common pixels in the moving window. It is necessary to check the estimated gain to see whether it is reasonable, and to prevent outliers from having a strong effect on the transform. If it is greater than the maximum allowable gain (\( M_{gain} \)) or less than the inverse (1/\( M_{gain} \)) of
the maximum allowable gain, then calculate gain and bias using the mean and standard deviation, just like the method mentioned in the ‘Phase One Methodology’ (Scaramuzza et al., 2003).

d. Estimate the value for the un-scanned pixel in the target SLC-off ETM+ image, using the linear relationship (see equation (1)).

Complete this procedure for each un-scanned pixel in the target SLC-off ETM+ image, thus the recovered ETM+ image may be obtained. However, in some cases, there are still some un-scanned pixels remaining in the recovered image, resulted from many factors. Under this condition, the recovering procedure is need to be conducted again, and even several times, by assigning the lastly recovered image as target image and taking the other fill image or images to repeat the recovering process mentioned above.

Note that this method will encounter the problem for the proper window around the edges of the image, and expanding the extent of the image is needed thereby. In this study, a symmetrical extending procedure was implemented around the edges of the image with the length of 17 pixels, in order to make AWLHM available to estimate each un-scanned pixel including the marginal ones in the tested ETM+ SLC-off images.

Compared with GLHM and LLHM, the AWLHM is more rational; however, poor results are potentially suffered if the images being combined exhibit radical differences in target radiance. At the same time, the AWLHM can perform well when adjust the temporal changes happened in homogeneous areas such as agricultural field, airport (F. Chen et al., 2010), but it has difficulty when the size of the areas exhibiting change is too small. Generally, features smaller than the local window size (normally 19 × 19 pixels) are difficult to model if significant change occurred. That is why fill data are typically better matched in the large agricultural fields typical of the American Midwest and West than in the smaller fields of the Northeast and, presumably, much of Europe and Asia (USGS & NASA, 2004). Taking into account the shortcomings for AWLHM, especially for heterogeneous landscape, J. Chen et al (2011) proposed an improved method, referred to as the Neighbourhood Similar Pixel Interpolator (NSPI). More details on the NSPI method will be given in next section (Section 2.3).

For obtaining recovered ETM+ image with good quality, selecting proper fill images is critical. According to the report written by USGS & NASA (2004), the most important criteria for selecting fill images are those related to the similarity of image content between the target and fill images. Any new data the similar fill image provide is likely to be of high quality, which will provide valuable additional information for transforming subsequent fill images. As suggested by USGS & NASA (2004), following guidelines for fill image selection are listed in order of priority:

a. Select fill images that are as free of clouds as possible and that contain as few obvious changes in image as possible.

b. Select SLC-off images that are as close time to target image as possible to minimize changes in vegetation conditions. Failing this, select fill images that are as close to an anniversary data as possible to take advantage of the similarities in seasonal vegetation cycles.
c. Select fill images that provide good predicted un-scanned coverage based on the un-scanned phase statistics.
d. Select SLC-off fill images that are within +/-4 WRS cycles of the target image, if possible.
e. For recovering image with multiple SLC-off images, including an anniversary date SLC-on image as the final fill image is recommended as a way of ensuring good image registration performance and providing complete un-scanned coverage.

2.3. NSPI method

The NSPI method is a simple but effective one, for estimating the values of un-scanned pixels in SLC-off ETM+ image. It is based on the assumption that the same-class neighbouring pixels around the un-scanned pixel have similar spectral characteristics, and that these neighbouring pixels and un-scanned ones exhibit similar temporal change patterns. Thus, as found in the initial research based on simulated and actual experiments, the NSPI method can restore the value of un-scanned pixels very accurately, and even performs well in heterogeneous areas (J. Chen et al., 2011). Another advantage of the NSPI method is that it can work well even if there is a relatively long time interval or obvious spectral changes between the fill and target images. Compared with the original AWLHM method, the NSPI method incorporates several other sub-processes, including selection of neighbouring similar pixels, calculation of the weights for similar pixels, calculation of the target pixel value in view of temporal differences.

2.3.1. Selection of neighboring similar pixels

As same as AWLHM method, all commonly scanned pixels in the determined adaptive moving window around the un-scanned pixel (target pixel) are selected. Then, similar pixels are selected from these common pixels based on spectral similarity, which is defined as root mean square deviation (RMSD) between each common pixel and the target one.

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (L(x_i, y_i, t_i, b) - L(x, y, t, b))^2}$$  \hspace{1cm} (2)

Where, $L(x_i, y_i, t_i, b)$ are the value of common pixel $i$, which locates in $(x_i, y_i)$, in band $b$ for the fill image acquired at time $t_i$, while $L(x, y, t, b)$ is for the target pixel locating in $(x, y)$; and $n$ is the number of spectral bands. Larger RMSD value denotes larger spectral difference.

According to previous research conducted by Gao et al (2006), common pixels with lower RMSD values than the predefined threshold are identified as the similar ones:

$$RMSD_i \leq \left[ \frac{\sum_{b=1}^{n} \sigma(b) \times 2 / m}{n} \right]$$  \hspace{1cm} (3)
Where, $\sigma_b$ is the standard deviation of the entire fill image for band $b$, and $m$ is the number of classes, which is a predefined value according to the complexity of the landscape.

Requirement for the minimum number (M) of similar pixels in the moving window is also need predefinition. Then, the initial moving window size can be estimated, and the number of selected similar pixels in this initial window is counted. Expanding the window iteratively, as described in Section 2.2, until either the requirement for minimum number of similar pixels (M) is met, or the window reaches the maximum size.

2.3.2. Calculation of the weights for similar pixels

The contribution of the selected similar pixels to the prediction of target pixel might vary according to their spectral similarity, temporal change difference and geographic distance related to the target pixel. Generally, similar pixels with higher spectral similarity and smaller distance to the target pixel are more important than other ones. The geographic distance between the $j$th similar pixel $(x_j, y_j)$ and the target pixel $(x, y)$ is calculated in Euclid space:

$$D_j = \sqrt{(x_j - x)^2 + (y_j - y)^2}$$  \hspace{1cm} (4)

Then, the weight indicating the contribution of similar pixel $j$ to the prediction of the target pixel is calculated by following equation:

$$W_j = \frac{1}{\text{RMSD}_j \times D_j} \times \frac{\sum_{j=1}^{N} 1/\text{RMSD}_j \times D_j}{\sum_{j=1}^{N} 1/\text{RMSD}_j \times D_j}$$  \hspace{1cm} (5)

The weight indicator $W_j$ is a normalized one, with the range from 0 to 1. For the special situation, when some similar pixels have the same or distinct spectral characteristics as the target pixel, that is to say, their spectral similarities RMSD are 0, then the weight for the special pixels is set to an equal value, and all the information is given by these special similar pixels.

2.3.3. Calculation of the target pixel value

According to the statements by J. Chen et al (2011), two procedures can be used to estimate the target pixel value using the information provided by the similar pixels separately. However, the integrated value based on the two predictions is possibly more reliable (J. Chen et al., 2011).

The first method is based on the fact that the similar pixels have the same or at least approximate spectral value with the target pixel when they are observed at the same time, and the weighted average of all the similar pixels in the target image (SLC-off one) is calculated.
\[ L_1(x,y,t_2,b) = \sum_{j=1}^{N} W_j \times L(x_j,y_j,t_2,b) \]  

At the same time, it is rational that the value (e.g. radiance, reflectance) of the target pixel at \( t_2 \) (the acquisition time of the target SLC-off image) is the sum of its value at \( t_1 \) (the acquisition time of the fill image) and the temporal spectra difference between \( t_1 \) and \( t_2 \). It is also acceptable in most cases that the similar pixels undergo the similar temporal change as the target one. Then, the second method for estimating the target pixel value at \( t_2 \) is developed, by using the change information provided by the similar pixels:

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

Predicted values with more accuracy may be obtained by integrating these two estimated results, with a suitable weight framework, in which both the landscape heterogeneity and the temporal change are incorporated. In the work of J. Chen et al (2011), the averaged RMSD between the similar pixels and the target pixel (RMSD\(_1\), see equation (8)) was used to indicate the heterogeneity, while the averaged RMSD of similar pixels between observations at \( t_1 \) and \( t_2 \) (RMSD\(_2\), see equation(9)) is used to denote the temporal change between two images:

\[ RMSD_1 = \frac{1}{N} \sum_{j=1}^{N} \left[ \frac{1}{n} \sum_{b=1}^{n} (L(x_j,y_j,t_1,b) - L(x,y,t_1,b))^2 \right] / n \]  

\[ RMSD_2 = \frac{1}{N} \sum_{j=1}^{N} \left[ \frac{1}{n} \sum_{b=1}^{n} (L(x_j,y_j,t_1,b) - L(x_j,y_j,t_2,b))^2 \right] / n \]  

The combined weights are calculated through a normalized process, shown as follows:

\[ T_1 = \frac{1}{RMSD_1 + 1} / (1 / RMSD_1 + 1 / RMSD_2) \]  

\[ T_2 = \frac{1}{RMSD_2 + 1} / (1 / RMSD_1 + 1 / RMSD_2) \]  

Finally, the estimated value for the target pixel is an integrated result combined two initial results and two combined weights:

\[ L(x,y,t_2,b) = T_1 \times L_1(x,y,t_2,b) + T_2 \times L_2(x,y,t_2,b) \]  

However, it is worth to note that when there is no similar pixel selected, the AWLHM method (see Section 2.2) can be applied to estimate the value of the un-scanned pixel.

### 2.4. Projection transformation method

A recovering method based on projection transformation was evaluated through a simulated experiment, which called the PCT-gap-fill method (Boloorani et al., 2008b). The
principal component transformation (PCT), which can make new representative data sets of the original ones, is mathematically lossless and vigorous invertible transformation. Over the obtained new components, if transformation be carried out, the original data sets can be recovered with no loss. The PCT-gap-fill method is to recover the un-scanned areas in the target image based on the scanned pixels from the fill image. Generally, this recovering procedure is mainly based on forward and backward principle component transformations. When the target and the fill images both have been normalized, such as transformed to reflectance, the PCT-gap-fill method can be applied as follows (Boloorani et al., 2008b):

a. Extract the scanned pixels in the target SLC-off image, and calculate the needed statistics for inverted PCT based on these pixels.

b. Extract the scanned pixels in the fill image, which were un-scanned in the target image, and apply the PCT to these selected pixels to get the transformed ones (denoted TFs).

c. Inversely transform the TFs using the statistics obtained from the scanned pixels in the target image (in step a.), to make new values for those pixels (NITFs) in the fill image selected in step b, then the un-scanned pixels in the target image can be recovered by means of setting the geospatially corresponding NITFs to them.

The step a. is important, where the scanned pixels in the target image need to be selected carefully due to the crucial role they play in the recovering procedure. In fact, the pixels selected for transformation statistics can be gotten either from the whole scanned area in the fill image or some samples of it. As suggested by Boloorani et al (2008b), the selected samples must be the real representative of the physical phenomena from the land surface. The previously simulated experiment shows that the PCT-gap-fill method recovered the SLC-off image with higher accuracy than the LLHM method, and preserved the radiometric characteristics better even in the areas with sharp radiometric changes. However, the simulated result also demonstrated that the gap lines were still visible in areas with sharp radiometric differences. The findings from this analysis based on simulated SLC-off image, provide a new concept for recovering the Landsat 7 SLC-off ETM+ image by means of data transformation, such as PCT, which is able to estimate the un-scanned pixels fully using the multispectral bands (e.g. Band 1-5, 7). Therefore, it is possible that the values of one specific un-scanned pixel in all bands can be estimated simultaneously by inverted PCT. Nevertheless, some other methods just are able to recover the Landsat 7 SLC-off ETM+ image band-by-band. In view of convenience for application, this methodology based on transformation is a good choice, given that several related problems is tackled properly, such as for heterogeneous area with sharp radiometric difference, the incorporation of the moving window tech.

2.5. Geostatistical method

Initially, the GLHM, LLHM and AWLHM methods were developed based on the basic assumption that there is a linear relationship between the target image and the fill one or ones. Thus, several images (SLC-on or SLC-off) need to be selected properly as the fill images for recovering the target SLC-off ETM+ image. The most important criteria for
selecting suitable fill images are cloud free and associated with time interval. However, in fact, these criteria are difficult to meet, especially for regions with moist environment. Under this condition, the GLHM, LLHM and AWLHM methods are limited to use consequently. On the basis that the autocorrelation is one of the predominant phenomena in physical geography, pixels in a specific extent may correlate with each other to a certain degree. So, it is possible to estimate un-scanned pixels based on the pixel relationship rules within the image itself, and this methodology is served as interpolation. General interpolation methods, such as inverse distance weighted methods and triangulation methods may be not suitable for recovering un-scanned pixels in the SLC-off image, because information about spatial structure is not used fully by these interpolation methods. Understanding the magnitude and pattern in spatial variability is necessary for accurately interpolating the un-scanned pixels. Zhang et al (2007) and Pringle et al (2009) discussed the possibility that using geostatistical techniques (e.g. kriging methods) to resolve the un-scanned pixels issue, which provides unbiased estimation with minimum and known error.

As mentioned above, the fundamental idea of geostatistics is the first law of geography, that spatial data from locations close to each other are more similar than data from locations far apart. Variogram used to indicate the spatial structure can be applied to estimate values at un-sampled locations (here, are un-scanned pixels in SLC-off ETM+ image).

Variogram model the spatial dependence in a regionalized variable \( Z \) (here is reflectance, radiance or digital number (DN) in SLC-off ETM+ image) under the ‘intrinsic’ hypothesis that the increments \( Z(x_i, y_i, h) - Z(x_j, y_j) \) associated with a small distance \( h \) are weakly stationary (Zhang et al., 2007). For multispectral image, in practice, a semivariogram is computed using the values in each band, shown as follows:

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i, y_i) - Z(x_j, y_j, h)]^2
\]

(13)

Where, \( N(h) \) is the number of pairs of pixel locating with \( h \) apart, and distance \( h \) is defined in Euclidean space. The larger value of \( \gamma(h) \) indicate the pixel pairs distributing apart with \( h \) distance are less similar.

Similarly, in view of the two co-dependent variables, the joint spatial dependence is common, and is often modelled using a cross-semivariogram, shown as follows:

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z_1(x_i, y_i) - Z_2(x_j, y_j, h)] [Z_2(x_i, y_i) - Z_2(x_j, y_j, h)]
\]

(14)

Where, \( Z_1 \) and \( Z_2 \) represents values in image 1 and image 2 (here referred to the target image and the fill image), respectively. Unlike semivariogram models, the cross-semivariogram may have negative values because of a negative cross-correlation between the two images.

The calculated semivariogram (cross-semivariogram) values based on the sampled pixels according to equations mentioned above, need to be fit for its further application in kriging.
Mathematical models frequently used for fitting semivariogram include nugget effect model, spherical model, exponent model, Gaussian model and power model. At the same time, for fitting cross-semivariogram, the basic models usually used are linear, spherical and exponential models.

Kriging is a family of generalized least-square regression algorithms that take advantage of the spatial dependence information given by the variogram. Ordinary kriging is the most common and robust one, which is able to rescale locally the random function model to a different mean value in different locations (Zhang et al., 2007). For estimating an un-scanned pixel \((x_i, y_i)\) in the SLC-off ETM+ image, pixels locating in the local neighbourhood are only used, and the extent of the neighbourhood is determined by semivariogram. Therefore, the value of pixel \((x_i, y_i)\) (denoted \(Z_{\text{OK}}(x_i, y_i)\)) is estimated through a linear model:

\[
Z_{\text{OK}}(x_i, y_i) = \sum_{j=1}^{n} \lambda_{ij}^{(\text{OK})}(x_i, y_i)Z(x_j, y_j) \quad \text{with} \quad \sum_{j=1}^{n} \lambda_{ij}^{(\text{OK})}(x_i, y_i) = 1
\]  

(15)

Where, \(\lambda_{ij}^{(\text{OK})}(x_i, y_i)\) is the ordinary kriging weight for the pixel \(j\) locating in the neighbourhood of pixel \((x_i, y_i)\), of which the value needs to be estimated, and \(n\) is the number of pixels locating in the neighbourhood of pixel \((x_i, y_i)\). The ordinary kriging estimator is unbiased by forcing the kriging weights to sum to 1 (see equation (15)).

Cokriging is an extended version of kriging, which incorporates more than one variable in the estimating procedure. Thus, generally, it can improve the accuracy, particularly when the primary variable is sparse and poorly correlated in space compared to the secondary one. Both the autocorrelation in each variable and the cross-correlation between the variables are taken into consideration in cokriging method. The value of pixel \((x_i, y_i)\) (denoted \(Z_{\text{COK}}(x_i, y_i)\)) estimated by means of cokriging method is expressed as follows:

\[
Z_{\text{COK}}(x_i, y_i) = \sum_{j=1}^{n_1} \lambda_{ij}^{(\text{COK})}(x_i, y_i)Z(x_j, y_j) + \sum_{j=1}^{n_2} \lambda_{ij}^{(\text{COK})}(x_i, y_i)Z_s(x_j, y_j)
\]

\[
\text{With} \quad \sum_{j=1}^{n_1} \lambda_{ij}^{(\text{COK})}(x_i, y_i) = 1 \quad \text{and} \quad \sum_{j=1}^{n_2} \lambda_{ij}^{(\text{COK})}(x_i, y_i) = 0
\]  

(16)

Where, \(\lambda_{ij}^{(\text{COK})}(x_i, y_i)\) is the ordinary cokriging weight for the pixel \(j\) locating in the neighbourhood of pixel \((x_i, y_i)\) for the primary variable, and \(n_1\) is the number of pixels locating in the neighbourhood of pixel \((x_i, y_i)\) for the primary variable. Similarly, \(\lambda_{ij}^{(\text{COK})}(x_i, y_i)\) is the ordinary cokriging weight for the pixel \(j\) locating in the neighbourhood of pixel \((x_i, y_i)\) for the secondary variable, and \(n_2\) is the number of pixels locating in the neighbourhood of pixel \((x_i, y_i)\) for the secondary variable.

As illustrated above (see equations (15) and (16)), two procedures are needed for estimating the value of the un-scanned pixel in advance, including determining the extent of local neighbourhood and thereby the valid sampling pixels, and calculating the kriging weights.
for each neighbourhood valid sampling pixel. Ordinarily, these procedures especially the last one need much time. The simplified versions for both ordinary kriging and cokriging were proposed for tackling the SLC-off issue by Zhang et al (2007), which are faster. The extent of the local neighbourhood is mainly defined based on the variogram, which are well-structured and can be fitted well due to the abundant data available in the SLC-off ETM+ image (about 80%). The specific pattern of the un-scanned pixels makes it possible to recover the SLC-off ETM+ image accurately, which was verified by Zhang et al (2007). Mean while, the case study showed that the cokriging method provided little improvement in estimating the un-scanned pixels in the SLC-off ETM+ image compared with the ordinary kriging method(Zhang et al., 2007). That is to say, similar results were obtained by these two approaches separately. The ordinary kriging allows users to extract the maximum information from the individual scene of the SLC-off ETM+ image by considering spatial variability and dependence, and it also overcomes the inherent problems caused by radiometric differences and small georeferencing errors possibly suffered in the GLHM, LLHM and AWLHM methods.

However, according to the statement of Pringle et al (2009), ordinary kriging and cokriging can be used complementarily to interpolate the value of un-scanned pixels in the SLC-off images. In detail, cokriging can be implemented by default and kriging is used in some sub-regions where the secondary variable is under-sampled. The recovered experiment based on the geostatistical method and two fill images (both are cloud-free), of which one was collected a few weeks earlier than the target image and the other one was collected a few weeks later, demonstrated the superiority of the geostatistics to several compositing methods and the improved prediction of cokriging method in the sub-regions where the secondary variable was more numerous than the target one (Pringle et al., 2009).

These previous researches suggest that the geostatistical method, especially the ordinary kriging is a good alternative to these traditional methods (e.g. the GLHM, LLHM and AWLHM) for predicting the values of un-scanned pixels in the SLC-off ETM+ image, especially for regions where suitable fill images are difficult to obtain, mainly resulted from weather conditions, cloud and temporal interval between images acquisition. Also, as shown in the previous results, main reasons for the advantages of geostatistics are including, (1) all un-scanned pixels in the SLC-off image are able to be recovered; (2) recovered results agree more strongly with the observations, even when there is an abrupt temporal changes; (3) for each recovered pixel, geostatistical method provides an estimation variance associated with uncertainty. However, the disadvantages of the geostatistical method in recovering the SLC-off ETM+ image are associated with its relatively slow speed and the selection of neighbourhood pixels for kriging weights. To solve its disadvantage in time consuming, Pringle et al (2009) proposed a hybrid technique. Additionally, further research is needed to find how many valid pixels in local neighbour is most suitable for estimating the un-scanned pixel, considering the specific spatial distribution characteristics of the un-scanned pixels in the SLC-off ETM+ image (Zhang et al., 2007).
2.6. Using images from other sensors

As mentioned before, selecting the fill image is crucial for recovering the SLC-off ETM+ image, and that cloud free images (SLC-on or SLC-off) acquired temporally close to the target image are desirable. But, for regions with moist environment the valid fill images are few. Consequently, due to the limitation of data acquisition, e.g. temporal resolution and atmospheric condition, there are always a large number of overlapping areas filled with un-scanned pixels in two cloud-free SLC-off ETM+ images which are close in time. Ordinary kriging has been illustrated as a good tool for tackling this issue (Zhang et al., 2007). However, as an alternative to ordinary kriging, the image from a sensor other than Landsat 7 ETM+, which is observed at the close time as the target SLC-off ETM+ image, may be able to be used to estimate the un-scanned pixels. This assumption has been verified by several previous researches. For example, Reza and Ali (2008) used IRS products, Boloorani et al (2008a) and F. Chen et al (2010) adopted EO-1/ALI and CBERS data as auxiliary information, respectively, while Roy et al (2008) used the information observed by MODIS.

Our experiments conducted on exploiting CBERS data as the fill image are shown here. All provided bands (totally four) of CBERS-02B CCD sensor have a spatial resolution of approximately 20m. There is similarity more or less between CBERS-02B CCD and Landsat7 ETM+ VNIR bands (Fig.5.) as well as the medium spatial resolution of each band, which makes it possible to estimate the un-scanned pixels in the ETM+ SLC-off image considering the CBERS-02B as the fill image. After the main procedures, including image resampling, geometrical registration, radiometric correction, several methods (e.g. Simple Filling, GLHM, LLHM, and AWLHM) were applied. Findings show the availability that using CBERS image as the fill data, and the superiority of AWLHM compared to the rest methods. However, further researches should be conducted, such as recovering all multispectral bands of SLC-off ETM+, and improving the estimated accuracy in urban or sub-urban areas with obvious heterogeneity.

Generally, AWLHM along with CBERS-02B estimated un-scanned pixels well around which the neighborhood was relatively homogeneous, however, was less accurate for pixels in the context of heterogeneous land use/cover. It is clear that the procedure performed well near the International Airport (Fig.6 (B2), (B4)) but less properly near the mixture places (Fig.6 (U2), (U4)) with incompatible appearance, no matter which date is considered. Setting recovered results of ETM+ 021 as an example, Fig.7 demonstrates that the linear relationship is less proper between pixels around P1 in SLC-off ETM+ and corresponding ones in CBERS-02B, compared with the good linear relationship for the relatively homogeneous neighborhood around P2, shown in Fig.8. The imprecise relationship might impact the accuracy of estimation at P1. These findings suggest the necessity to resolve the issues associated with heterogeneity to recover the SLC-off images in urban and sub-urban regions. According to the previous researches, ideas incorporated into the NSPI method (J. Chen et al., 2011) may be worth learning.

Another problem is that the number of multispectral bands validly provided by CBERS (here is CBERS-02B) is less than that of Landsat 7 ETM+. It may limit the full usage of the
recovered results; because recovering procedures are always operated band-by-band, except the recovering method based on projection transformation (see Section 2.4). The simple method by using the intra-bands relationship is possible, in addition to the methodology based on projection transformation. Reza and Ali (2008) adopted the simple method to fully recover the multispectral bands of the SLC-off ETM+ image using IRS/1D LISS-III as the fill image through two stages.

Figure 5. Comparison of relative spectral response (RSR) between CBERS-02B and ETM+ (Original data were provided by CRESDA (www.cresda.com) and NASA respectively, and each ETM+ RSR exhibited above was resampled to a spectral resolution of 10nm so as to compare)
Figure 6. Detailed exhibition of small extracts (R: band4, G: band3, B: band2) (U1: from CBERS-02B, U2: from original ETM+ 021, U3: from recovered ETM+ 021, U4: from original ETM+ 309, U5: from recovered ETM+ 309; B1: from CBERS-02B, B2: from original ETM+ 021, B3: from recovered ETM+ 021, B4: from original ETM+ 309, B5: from recovered ETM+ 309) (in F. Chen, et al., 2010)

Figure 7. Scatter plots of corresponding bands for scanned pixels around P1 (in F. Chen, et al., 2010)
MODIS is another data resource considered as the fill image. Landsat 7 ETM+ bands and the corresponding bands of MODIS are shown in Table 1. Roy et al (2008) indicated the possibility that recovering the SLC-off ETM+ image by means of a semi-physical fusion approach, when the MODIS BRDF/Albedo land surface characterization product was used as auxiliary data. Detailed analysis demonstrated the major advantages of this fusion method, the simplicity for application and the ability in accommodating for temporal variations due to surface change (Roy et al., 2008). The main disadvantages of this method are related to the requirement on adequate data co-registration, computation of geometry and the obvious scale discrepancy between MODIS BRDF/Albedo product and the Landsat 7 ETM+ data. Nevertheless, MODIS provides one more chance for recovering the SLC-off ETM+ image, mainly due to the abundant images it provides. Therefore, further researches are required to tackle the associated issues, then to make the full advantages of MODIS.
Table 1. The multispectral bands of Landsat 7 ETM+ and the corresponding ones of MODIS

<table>
<thead>
<tr>
<th>Band number of ETM+</th>
<th>Band wavelength (nm)</th>
<th>Band number of MODIS</th>
<th>Band wavelength (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>450-520</td>
<td>3</td>
<td>459-479</td>
</tr>
<tr>
<td>2</td>
<td>530-610</td>
<td>4</td>
<td>545-565</td>
</tr>
<tr>
<td>3</td>
<td>630-690</td>
<td>1</td>
<td>620-670</td>
</tr>
<tr>
<td>4</td>
<td>780-900</td>
<td>2</td>
<td>841-876</td>
</tr>
<tr>
<td>5</td>
<td>1550-1750</td>
<td>6</td>
<td>1628-1652</td>
</tr>
<tr>
<td>7</td>
<td>2090-2350</td>
<td>7</td>
<td>2105-2155</td>
</tr>
</tbody>
</table>

2.7. Integrated method

According to previous comparisons, each method has its own relative advantages and disadvantages, compared with other ones. For example, geostatistical method is generally superior to several methods based on the linear relationship (e.g. GLHM, LLHM and AWLHM), but it consumes much more time. Mean while, data obtained by other sensors instead of ETM+ may provide useful auxiliary information for recovering the SLC-off ETM+ image, but the differences between these sensors and ETM+ should be considered, which are related to sensor geometry, spatial resolution, band settings, and so on. Pringle et al (2009) proposed a hybrid technique which incorporates the simple ones with the geostatistical method. For example, after implementing the LLHM the rest un-scanned pixels in the SLC-off ETM+ image can be estimated by means of kriging or cokriging (Pringle et al., 2009).

2.8. Recovering the thermal band of SLC-off ETM+ image

Currently, there have been few researches done to resolve the SLC-off thermal band image, although it had been used widely, as shown in Section 1. A method called the modified AWLHM (see Fig.9) was proposed to fill the gaps in the thermal band of SLC-off ETM+ image, while data obtained by other sensors are used as fill image (here is CBERS). The experiment shows the possibility that using the modified AWLHM method and suitable band combinations to recover the thermal band of SLC-off ETM+ image. While giving the priority to the accuracy, it is more rational and practicable to get brightness temperature directly than to inverse brightness temperature from the recovered thermal band. Based on the detailed analysis on the validation results, we can assume that it is practicable to implement the recovered thermal band in the study of urban thermal environment, with certainty to some extent. However, further researches need to be done to enable more scientific use of the thermal band of SLC-off ETM+ data.
3. Experiments and applications

The possible usage of the recovered ETM+ images either the multispectral bands or the thermal band is an important issue when talking about recovering the SLC-off ETM+ image. However, few researches associated have discussed the possible application of the recovered images. For example, Boloorani et al (2008b) and J. Chen et al (2011) demonstrated the possibility of applying the recovered multispectral bands in land cover/use classification, although they just conducted these studies based on the simulated data. F. Chen et al (2010) recovered two SLC-off ETM+ images acquired at different times and demonstrated the usage of the recovered images in mapping urban impervious surface. The possible usage of the recovered thermal band in urban thermal environment study was discussed (F. Chen et al., 2011a), while its application has also been illustrated (F. Chen et al., 2011b).

Fig.10 shows two recovered ETM+ images for Xiamen Island, locating in Fujian Province, China, which is highly urbanized. Urban impervious surface modelled from these recovered images indicates that similar estimation accuracy was obtained at the filled pixels and scanned ones, as shown in Fig.11 (F. Chen et al, 2010).

Fig.12 shows the simulated results of recovering the SLC-off thermal band. The recovered results are visually acceptable, even in high-contrast areas. Furthermore, the validation reveals that unbiased estimation can be obtained through our recovering procedure at most of the un-scanned pixels (about 80%) in the thermal band (Chen et al, 2011). According to these findings, we confidently assume that the recovered thermal band can be applied in the issues associated with urban thermal environment. Several recovered thermal bands were applied in urban thermal environment studies (F. Chen et al., 2011b). Fig.13 demonstrates the land surface temperature results obtained from four recovered SLC-off thermal bands.
Figure 10. RGB colour composite of the recovered and original ETM+ images and CBERS-02B ((A) CBERS-02B; (B) original ETM+021; (C) original ETM+309; (D) recovered ETM+309 using ETM+ 021 as auxiliary data; (E) recovered ETM+021 using CBERS-02B as auxiliary data; (F) recovered ETM+309 using CBERS-02B as auxiliary data) (in F. Chen, et al., 2010)

Figure 11. Results of UIS estimation accuracy assessment for recovered ETM+ 021: (A) accuracy assessment for filled pixels; (B) accuracy assessment for scanned pixels; (C) residual analysis for filled pixels; (D) residual analysis for scanned pixels (in F. Chen, et al., 2010)
4. Discussion and conclusion

Since the launch of its first satellite in 1972, Landsat mission has continuously and consistently collected images of Earth with global coverage, due to the continuous series (see Fig.1). The moderate spatial resolution of the Landsat sensors is important for scientists or common users, because it is either coarse enough for observing global environment or detailed enough for monitoring human-scale processes. This specifically spatial resolution is sensitive to both natural and human changes, although abundant information is missing for some individual objects. The special characteristics, both data archive with long history and suitable spatial resolution, make the Landsat an important tool, which gives scientist the ability to assess environment changes
happened or is happening at local, regional and global scales. Therefore, as NASA administrator Dr. James Fletcher predicted in 1975, if one space age development would save the world, it would be Landsat and its successor satellites (http://landsat.gsfc.nasa.gov/about/).

In addition, based on the free data policy announced in 2008, all new and archived Landsat data held by the USGS have been made freely obtainable over the internet for any user (Woodcock, et al., 2008). In the short time since this policy was put into effect, significant benefits are already being realized within the Landsat data user community (Loveland & Dwyer, 2012). So, it is highly possible that Landsat data will experience more widespread usage especially for long-term and large-area environmental issues. Furthermore, the subsequent Landsat satellites (e.g. Landsat 8, Landsat 9, and Landsat 10 and so on) make the possibility to be realized. As Loveland & Dwyer states, the next few years are quite possibly going to be Landsat’ s ‘golden years’ and could be the time in which the Landsat program achieves its full potential (Loveland & Dwyer, 2012).

However, the SLC failure of Landsat 7 ETM+ sensor occurred in 2003 may impact the Landsat data archive to some extent, and even results in a data gap period, especially under the condition that Landsat 5 has also suffered problems with its solar array drive sometimes and the failure of its electronic component currently. Therefore, suitable ways are desirable to resolve the SLC-off problem accurately and robustly, and then enable the usage of SLC-off ETM+. Much effort has been done to meet this goal, widely by different institutions and individuals. Currently, methods based on different assumptions have been proposed and demonstrated, some of which were also discussed in this chapter. Generally, for areas with spatially homogeneous property and less temporal fluctuation, most of them can perform well and the un-scanned pixels in the SLC-off ETM+ images are able to be recovered accurately. Meanwhile, in practice, the un-scanned pixels, locating in heterogeneous areas with obvious temporal changes, are difficult to be estimated. Proper selection of the fill images is important for the recovering procedure; nevertheless, the important criteria suggested by USGS & NASA (2004) are difficult to meet usually. In practice, for a specific application, method and recovering procedure should be determined by balancing its performance and efficiency, besides taking into account the availability of the fill image and land surface properties in SLC-off image. In some cases, images available from other sensors may provide valuable auxiliary information for tackling the SLC-off issue, after being properly processed.

It is worth to note here that, much attention should be paid to recover the SLC-off ETM+ images acquired over urban and suburban areas, although few researches have been conducted. Making full use of the recovered SLC-off ETM+ images is desirable for investigating changes in regions, which have experienced rapid urbanization and consequently suffered environmental deterioration at present. This is a very important and valuable issue, especially for many developing countries or regions, for whom financial
supports given to afford abundant satellite images are limited. The recovered SLC-off ETM+ images, along with other Landsat data will provide an opportunity for these countries and regions to monitor, study and manage their own urbanized issues.

**Author details**

Feng Chen, Xiaofeng Zhao and Hong Ye  
Key Lab of Urban Environment and Health,  
Institute of Urban Environment, CAS, The People’s Republic of China  
Xiamen Key Lab of Urban Metabolism, The People’s Republic of China

**Acknowledgement**

This work was supported by the Knowledge Innovation Program of Chinese Academy of Sciences, CAS (Grant No. 09L4401D10), the National Science Foundation of China (40901218) and (41101551). Thanks are given to data providers (e.g. CRESDA, USGS and NASA) for providing the images used in this work. The authors would like to acknowledge all who have contributed to this work, practically Dr. Jingzhu Zhao, Dr. Lina Tang and Dr. Quanyi Qiu, with the Institute of Urban Environment, Chinese Academy of Sciences. Many thanks to all authors of whose scientific publications or reports their work has profited in our manuscript.

**5. References**

Chen, F.; Tang, L. & Qiu, Q. (2010). Exploitation of CBERS-02B As Auxiliary Data in Recovering the Landsat7 ETM+ SLC-off Image, In: *Proceedings of the 18 International Conference on GeoInformatics*, 18-20 June, Peking University, Beijing, China  


