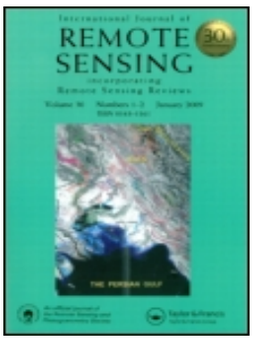


International Journal of Remote Sensing

ISSN: 0143-1161 (Print) 1366-5901 (Online) Journal homepage: http://www.tandfonline.com/loi/tres20

Normalization of medium-resolution NDVI by the use of coarser reference data: method and evaluation

Wenxia Gan, Huanfeng Shen, Liangpei Zhang & Wei Gong

To cite this article: Wenxia Gan, Huanfeng Shen, Liangpei Zhang & Wei Gong (2014) Normalization of medium-resolution NDVI by the use of coarser reference data: method and evaluation, International Journal of Remote Sensing, 35:21, 7400-7429, DOI: 10.1080/01431161.2014.968684

To link to this article: http://dx.doi.org/10.1080/01431161.2014.968684

International Journal of Remote Sensing, 2014
Vol. 35, No. 21, 7400–7429, http://dx.doi.org/10.1080/01431161.2014.968684

Taylor & Francis
Taylor & Francis Group

Normalization of medium-resolution NDVI by the use of coarser reference data: method and evaluation

Wenxia Gan[a], Huanfeng Shen[b]*, Liangpei Zhang[a]*, and Wei Gong[a]

[a]*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei, 430079, PR China;* [b]*School of Resource and Environmental Sciences, Wuhan University, Wuhan, Hubei, 430079, PR China*

(*Received 14 January 2014; accepted 4 September 2014*)

Medium-resolution remote-sensing images with tens of metre spatial resolutions have spatial and spectral characteristics that are suited for mapping a range of structural and compositional properties of vegetation. However, many factors, such as the long revisit cycles and frequent cloud contamination, limit the availability of images for the monitoring and time-series analysis of vegetation. Thus, there is a strong incentive to combine data from more than one observation system in order to fill the gaps in observation and enhance the capability of remote sensing to monitor dynamics. In this paper, we introduce a framework for the normalization of the normalized difference vegetation index (NDVI) from different sensor systems by the use of synchronous coarse-resolution NDVI data. A new model called the Local Cluster-specific Linear Model (LCLM) is proposed. This model is designed to build the specific relationships for different clusters, block by block, considering the spatial heterogeneity of the influencing factors. To improve the stability of the parameter estimation, an M-estimation method is utilized to solve the coefficients. Based on an analysis of the previous evaluation methods, new schemes are designed for evaluating the accuracy of the parameter normalization. Different assessment experiments were undertaken with the new evaluation schemes, to validate the performance of the LCLM method. The results indicate that the LCLM method performs better than the existing methods. An application experiment was also undertaken, in which synchronous NDVI from Landsat ETM+ and Terra ASTER sensors were normalized by the use of a coarse-resolution MODIS product.

1. Introduction

The normalized difference vegetation index (NDVI) is calculated from the spectral reflectance measured in the visible and infrared bands of a satellite sensor and provides an indication of photosynthetically active vegetation (Tucker and Sellers 1986). Among large numbers of remote-sensing products, coarse-resolution NDVI from sensors such as the NOAA Advanced Very High Resolution Radiometer (AVHRR), Terra/Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) and SPOT VEGETATION usually has the advantage of frequent observations, but lacks enough spatial resolution for a fine-scale (the term scale in this paper refers to spatial resolution) regional study. In contrast, medium-resolution remote-sensing images, with a spatial resolution from 10 m to 100 m, acquired from sensor systems such as the Landsat Thematic Mapper, and Enhanced Thematic Mapper Plus (TM/ETM+), Terra Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), China–Brazil Earth Resources Satellite

*Corresponding authors. Email: shenhf@whu.edu.cn (H. Shen); zlp62@whu.edu.cn (L. Zhang)

(CBERS) CCD, and HJ-1A/1B CCD have spatial and spectral characteristics that are suited for mapping a range of vegetation structural and compositional properties. However, frequent cloud contamination, long revisit cycles of the medium-resolution satellite observation systems, and a number of other factors sometimes make it difficult to acquire remotely sensed data with a continuous spatial coverage. In addition, these factors also limit the availability of images at intervals and durations that are suitable for the time-series analysis of vegetation, especially for the period of active growth (Price 1994). Thus, there is a strong incentive to combine data from multiple observation systems, to fill the gaps in the observation and enhance the monitoring of the land surface.

Integration of NDVI from different sensors is complicated as various factors are known to affect the consistency and continuity of NDVI data sets (Gallo and Eidenshink 1988; Guyot and Gu 1994; Holben 1986). It is known that the spectral characteristics of surface vegetation and soil components, sensor-specific spectral band characteristics, and atmospheric conditions can all result in inconsistencies in NDVI between different sensors (Trishchenko, Cihlar, and Li 2002; Jonas and Menz 2004; Steven et al. 2003; Miura, Huete, and Yoshioka 2006; Van Leeuwen et al. 2006; Franke, Heinzel, and Menz 2006). And it was found in a previous study that the use of equivalent surface reflectance data in computation of a vegetation index eliminates many of the errors and provides a sound basis for the comparison of indices measured over time (Guyot and Gu 1994). However, vegetation indices measured by different sensors may not match, even with precise calibration and atmospheric correction. This is because the sensor-specific bands (position and width) and spectral response functions may result in differences in NDVI between different sensors because the bands receive slightly different components of the reflectance spectra of vegetation and soil (Guyot and Gu 1994; Gallo and Daughtry 1987). These inconsistencies can be problematic when different monitoring systems are used in combination, or when one system takes over from another in a long time series. Thus, direct comparison or synergistic use of NDVI from different sensor systems is crucial. This all demonstrates the need to standardize/normalize the vegetation indices from different sensor systems (Steven et al. 2003).

Several empirical methods have been proposed for reconciling NDVI data produced from different sensors. One popular approach for normalization of NDVI from different sensor systems, which has been the subject of many studies, is building inter-sensor relationships by simulating the reflectance data of different sensors. Using this approach, equations have been provided to convert the NDVI calculated by surface reflectance from multiple sensors to MODIS NDVI (Steven et al. 2003; Van Leeuwen et al. 2006; Jiang et al. 2006). Furthermore, a series of quadratic 'spectral correction' functions have been developed to translate the NDVI of selected sensors to NOAA-9 AVHRR-equivalents through hyperion-based simulation of different land-cover types (Yoshioka, Miura, and Huete 2003). However, it should be noted that these methods build relationships by simulated NDVI samples; that is, they are produced by simulating the responses of a series of satellite instruments from the same data set of spectroradiometric measurements, or hyper-spectral sensors over canopies, by convolving the spectral response functions. This approach might, however, be inapplicable for data from other surface targets or data acquired at other times, which could be influenced by much more varied conditions. Moreover, the difficulty of atmospheric corrections means that it is not easy to obtain an accurate surface reflectance. Therefore, it might be problematic when adopting these specific relationships based on simulated NDVI data for cross-sensor NDVI transformation.

Another kind of approach is building an inter-sensor transformation equation based on near-synchronous images, and then undertaking the normalization of the NDVI from the different sensor systems. A relationship was built between the NDVIs from synchronous ASTER and Landsat ETM+ images (Zhang 2011). A similar relationship was also built between the IKONOS and ETM+ NDVIs by the use of approximately synchronous images from the two sensors, with a highly significant R^2 value of 0.67 or 0.68 (Thenkabail 2004). The inter-sensor relationships between the Landsat-7 ETM+ NDVI and the NDVIs from Landsat-5 TM, QuickBird, LISS-III IRS 1C/D, and ASTER were also built based on synchronous image pairs (Martínez-Beltrán et al. 2009). The equations obtained in these studies, based on the synchronous imagery, correspond more with the real situation; however, as the equations were built based on specific data, they are highly efficient for the data used for the modelling, but they might not be appropriate for other imagery with different air conditions, different solar illumination, different viewing angles, and so on. It is therefore difficult to generate these inter-sensor relationships and carry out inter-sensor transformations when there is a lack of synchronous data.

As opposed to a cross-sensor transformation equation built based on simulated samples or synchronous imagery with a similar spatial resolution, normalization of the data using synchronous coarse-resolution data as a reference is faithful in the real situation, and is not limited by the acquisition of reference data. In contrast to the data from medium-resolution sensors, data from coarse-resolution sensors with shorter revisit times, and thus with a high temporal frequency, are more widely available. Adopting the coarse-resolution data to be the reference could therefore avoid the problem of a shortage of synchronous reference data. The normalization of medium-resolution data using coarse-resolution data was earlier tested for the normalization of Landsat ETM+ reflectance data with coarse-resolution 10-day composite SPOT VEGETATION (VGT) data as the reference (Olthof et al. 2005). This work was then expanded to combine medium-resolution reflectance data from different sources, including Landsat TM/ETM+, AWiFS and CEBRS, by using MODIS land products as reference (Gao et al. 2010). The normalization results of these methods were shown to be consistent and comparable, both spatially and temporally. This approach was also introduced and tested for the normalization of Landsat ETM+ NDVI, with a MODIS product as a reference, by using a linear model built for different clusters (Gan et al. 2013). This method considers the various inter-sensor relationship differences between different clusters, but some important factors are not taken into account. For example, the spatial heterogeneity of the atmospheric conditions will influence the $NDVI_{DN}$-$NDVI_{SR}$ (NDVI calculated by the digital number (DN) and surface reflectance) relationship, which is important for the normalization.

Meanwhile, there is also a requirement for an effective evaluation scheme for normalization methods. In a previous study of the normalization of Landsat ETM+ reflectance data using coarse-resolution data as reference, evaluation was implemented by applying the normalization coefficients to a part of the resampled ETM+ data, and then comparing the residual statistics with the reference (Olthof et al. 2005). Evaluation can also be implemented by checking the difference between the reference MODIS surface reflectance and upscaled normalized result (Gao et al. 2010). Evaluation has also been carried out by comparing the normalized results of different sensors in the overlapping regions, and by comparing the statistics generated from the overlapping regions (Gao et al. 2010; Olthof et al. 2005). However, as NDVI is a scale-dependent parameter, and bias can be brought about by the scaling of the NDVI (Huete, Kim, and Miura 2005), some of the above methods are not appropriate for the evaluation of NDVI normalization.

In this paper, we first introduce a framework for the normalization of medium-resolution NDVI from different sensors by using coarse-resolution data as a reference. In addition, a new model, the Local Cluster-specific Linear Model (LCLM), used under the framework of a reference-based method, is proposed. Unlike the previous studies (Gan et al. 2013; Gao et al. 2010), which used relationships built for different clusters based on samples from the whole image, we separate the whole image into several blocks and build models for the different clusters, block by block. This method can consider the spatial characteristics of the factors that influence the inter-sensor relationships, such as aerosol optical thickness, water vapour, solar viewing angles, and so on, which are known to be spatially heterogeneous. In this way, we can improve the accuracy of the inter-sensor NDVI transformation.

We summarize the existing evaluation methods and analyse their applicability for NDVI normalization in Section 3. As the issue of evaluation needs to be further developed, several schemes are designed to evaluate the accuracy of the methods, according to the principle that the normalized result should be consistent with the reference data, but with its own original spatial resolution. Following the detailed description of the framework of the reference-based method, the building of the model used under this framework is introduced in Section 2. Section 3 reviews the previous evaluation schemes and presents new evaluation schemes designed for the normalization of medium-resolution data by using the coarse-resolution data as a reference. The evaluation experiments and application experiment are then described in Section 4. In Section 5, the advantages and limitations of the proposed approach are discussed, and our work is concluded.

2. Normalization method

2.1. Framework

The basic idea behind the normalization of medium-resolution data from different sensors by using the coarse-resolution NDVI is to make use of the spatial and temporal consistency and comparability of the reference (refer to Figure 1). We make use of the spatial

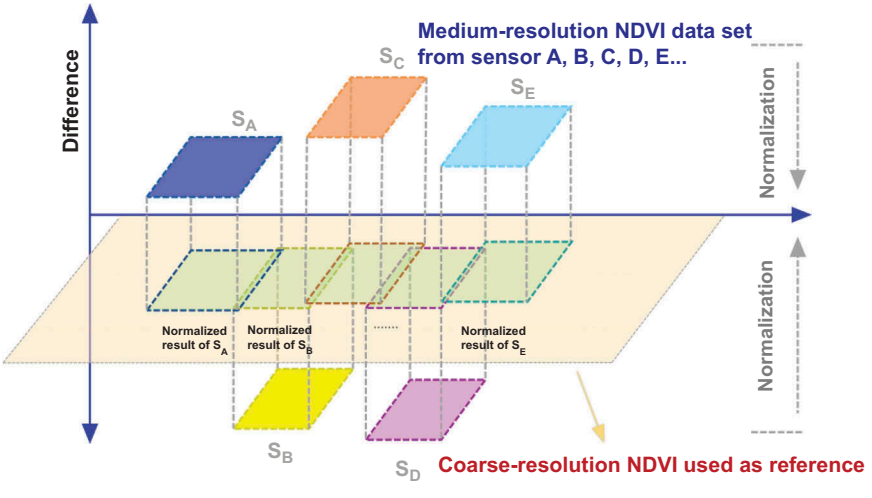


Figure 1. Schematic diagram of the framework for the normalization of the NDVI by the use of coarse-resolution NDVI as the reference.

consistency of the reference NDVI to quantify the effective spatial comparability of the NDVIs from different sources, thus to make the difference among the NDVIs over different areas relate only to the difference of the land surface and not to be disturbed by the bias brought in during the observing process. Differences in the NDVIs among different medium-resolution sensors (sensor A, B, C, …) are eliminated through normalization by the use of the same coarse-resolution NDVI data as the reference. First, a model is built at a coarse resolution at the same scale as the NDVI_R (the coarse-resolution NDVI used as the reference) between the resampled NDVI_O (the original medium-resolution NDVI needing normalization) and the corresponding NDVI_R by an area-averaging aggregation method. The established model is then applied to the NDVI_O to obtain the normalized result, the medium-resolution NDVI that is consistent with the reference data, which is marked as NDVI_N (NDVI after normalization). This process can also be expanded upon to achieve temporal normalization by taking advantage of the temporal consistency of the reference data set to assure that the multi-temporal NDVI from the different sources is consistent over time, and able to reflect the real temporal change of the land surface without being disturbed by the bias caused by sensor difference, viewing angle, atmospheric condition and so on. It is indeed useful to be able to increase the frequency of the medium-resolution NDVI record, for either local or regional analysis.

However, as the coarser-resolution NDVI tends to contain bias caused by the scaling effect (Yoshioka et al. 2008; Obata, Wada et al. 2012; Obata, Miura, and Yoshioka 2012; Obata et al. 2013), this can lower the accuracy of the medium-resolution normalization result, which is a notable problem when using coarse-resolution NDVI data to normalize medium-resolution NDVI.

Meanwhile, there are no strict requirements for the source of the original NDVI. It is possible to standardize the NDVIs from different medium-resolution sensors to one standard data set directly, without any other process, no matter whether the NDVI is calculated by DN, top of atmosphere (TOA) reflectance or other sources. However, the requirements for the coarse-resolution NDVI used as a reference are strict. The data should comply with the following principles: for spatial normalization, when undertaking normalization of the NDVI cover for different regions which might be with different land cover and different atmospheric conditions, the reference NDVI should have a coverage that is wide enough to cover all the NDVIs that need normalization, and the NDVIs of the different regions should be comparable. For temporal normalization, when undertaking normalization of multi-temporal NDVIs from different sensor systems of a given region, the reference NDVI data set needs to be supplied by a sensor with a short revisit time, so as to provide a record that is frequent enough to cover each time interval. Furthermore, similar equator crossing times help to ensure the temporal consistency and to reduce the bias caused by different imaging times. Sensors with a wide coverage and short revisit time, such as AVHRR, MODIS and SPOT VEGETATION, are all good sources for reference data. Other data could also be chosen, as long as they can meet the requirements for the reference data.

During the normalization, the model built at a coarse resolution is applied to the medium-resolution NDVI_O. Since the NDVI is a scale-dependent parameter (Jiang et al. 2006), how to build a scale-invariant model is important for the normalization methods of medium-resolution NDVI from different sensors. The strategy adopted here is to build a model based on homogeneous coarse-resolution pixels to reduce the scale-induced bias. It has been pointed out in previous studies that the difference between the two ways of upscaling NDVI will become zero when one of the fractional amounts are close to zero

(homogeneous case). The first one is computing NDVI from fine-resolution bands and then aggregating the results to some coarser level, and the second one is aggregating the bands and then computing the index (Jiang et al. 2006; Chen 1999; Huete, Kim, and Miura 2005; Quattrochi and Goodchild 1997). These two ways for upscaling are respectively marked as UP_N and UP_R in later sections. As a supplement, we analyse the relationship between homogeneity and scale-induced deviation of the NDVI based on real data. Scale-induced deviation of the NDVI (δ_{NDVI}) is defined as (Wu, Tang, and Li 2013):

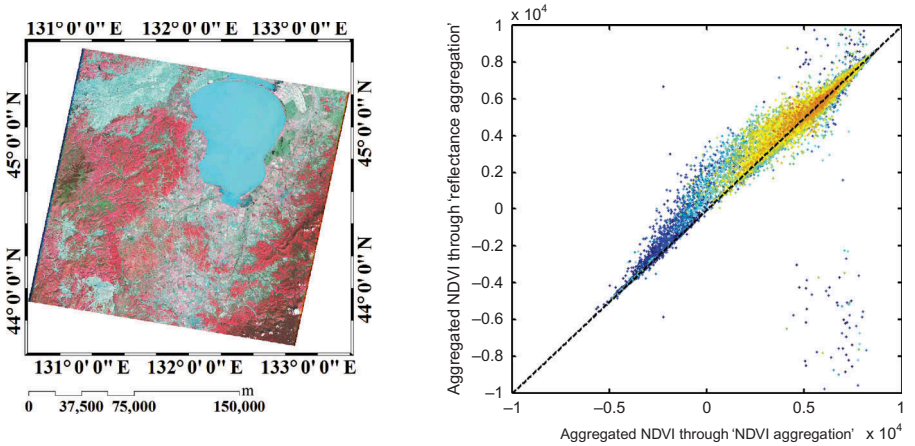
$$\delta_{\mathrm{NDVI}} = \mathrm{NDVI}' - \frac{\rho'_{\mathrm{Nir}} - \rho'_{\mathrm{Red}}}{\rho'_{\mathrm{Nir}} - \rho'_{\mathrm{Red}}}. \tag{1}$$

NDVI' represents the NDVI in a coarse resolution through the pixel aggregation of the fine-resolution NDVI, and ρ'_{Nir} and ρ'_{Red} represent the upscaled near-infrared and red reflectances. The homogeneity of each coarse pixel is measured by the percentage of pixels belonging to the majority cluster within the extent of the current coarse pixel, based on the classification map in the medium resolution. It is defined as:

$$r = \frac{k_c}{m^2}, \tag{2}$$

where r represents the ratio of the majority cluster of the coarse-resolution pixels, while k_c represents the number of pixels belonging to cluster c, which is the major cluster. m represents the scale ratio between the fine resolution and the coarse resolution.

A Landsat ETM+ image is taken here as an example. Different types of land covers are included in the image (Figure 2(*a*)), and its NDVI ranges from about −0.5



(*a*) (*b*)

Figure 2. (*a*) False-colour composite ETM+ image; (*b*) scatterplot of the NDVI calculated by coarse-resolution surface reflectance and pixel-aggregated coarse-resolution NDVI (*10,000), taking six clusters as example.

Table 1. The theoretical minimum homogeneity of a coarse resolution pixel with different numbers of clusters.

Number of clusters	2	4	6	8	10	14
Minimum homogeneity (%)	50	25	16.67	12.5	10	7.143

to 0.9, providing a good representation of most of the conditions. Taking the coarse-resolution pixels with 250 m resolution for example, every pixel is corresponding to about 64 ETM+ pixels (30 m resolution). The theoretical minimum homogeneity of a coarse-resolution pixel is shown in the Table 1. As shown in Figure 3, the average scale-induced error decreases when the homogeneity increases. When the percentage of the majority cluster reaches about 60%, the difference is stable and small. The same conclusion can also be reached from the scatterplot (Figure 2(*b*)) between 'NDVI aggregation' and 'reflectance aggregation'. A warm colour in the scatterplot represents pixels with a high homogeneity, while a cool colour represents pixels with a low homogeneity. As the warm-coloured points with high homogeneity are tighter to the 1:1 line than the cool-coloured points, this means that the pixels with a high homogeneity have a lower scale-induced error. In summary, it is reasonable to conclude that pixels with a high homogeneity are only slightly influenced by scale, and thus a model built based on homogenous coarse-resolution pixels should be efficient for the normalization of medium-resolution NDVI. In our work, the homogeneous samples selected for the modelling are determined by the predefined threshold of homogeneity.

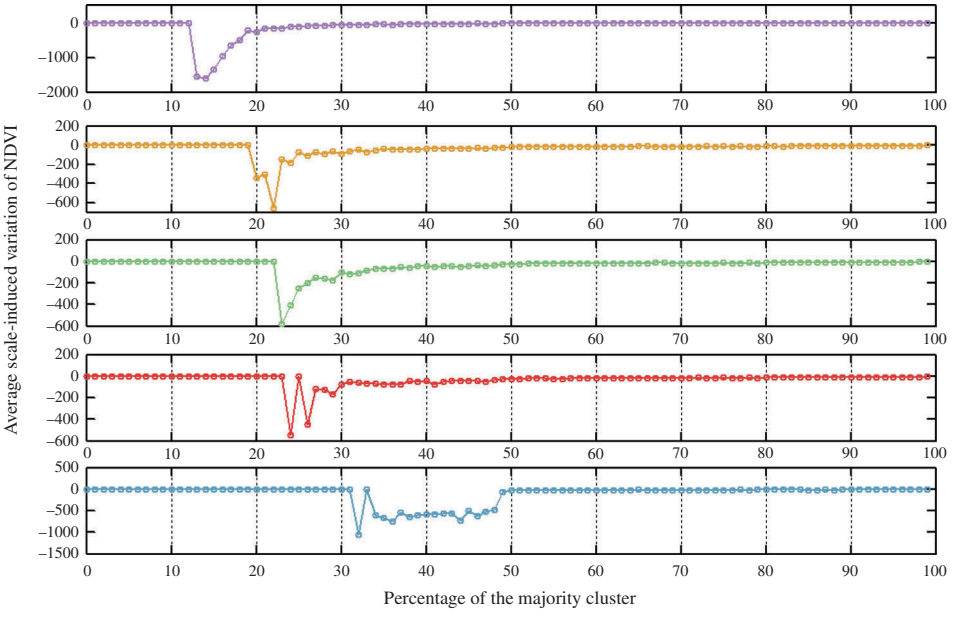


Figure 3. Plot of the average scale-induced error with different numbers of clusters (from the top to bottom, the cluster numbers are, 14, 10, 8, 6, and 4).

2.2. The Global Linear Model

The linear model is the typical model which is most widely adopted for cross-sensor NDVI transformation (Martínez-Beltrán et al. 2009; Thenkabail 2004), and it can also be used for the normalization of the NDVI from different sensors by the use of coarser-resolution NDVI as the reference. It is referred to as the Global Linear Model (GloLM) in this paper. 'Global' means that the model is built based on samples from the whole image. The main workflow of the GloLM is to first build a global linear model at a coarse resolution, and then to apply this to the medium-resolution NDVI_O.

Although the linear model is the most widely used relationship, it may not hold for the comprehensive inter-sensor relationships of the NDVI, especially for the NDVIs calculated by different parameters, such as DN, TOA reflectance, or surface reflectance. Furthermore, the inter-sensor differences in the data show different patterns according to the different land-cover types (Jiang et al. 2006), and they are influenced by many different factors.

2.3. The Global Cluster-specific Linear Model (GCLM)

The GCLM is used to build different linear models for different clusters. The idea of the GCLM has been used for the normalization of reflectance data, assuming that the inter-sensor relationship can be described by a cluster-specific linear relationship (Gao et al. 2010). It has also been used for the normalization of the NDVI from different sensors, and has obtained a good performance. It has been found that the non-linear relationship between $NDVI_{DN}$/$NDVI_{TOA}$ (the NDVI calculated by the DN or TOA reflectance) and the $NDVI_{SR}$ can be handled well by the cluster-specific linear model (Gan et al. 2013).

The main workflow of the GCLM is to build different linear models for different clusters, and to obtain the normalization coefficients at the same coarse resolution as the reference data, based on the homogenous pixels from the NDVI_R and the resampled NDVI_O. The coefficients are then applied to the NDVI_O data, according to their classification map. The cluster type of each pixel is determined by the classification map.

2.4. The Local Cluster-specific Linear Model (LCLM)

Under the framework of normalization by the use of coarse-resolution data as the reference, the factors that will influence the $NDVI_{DN}$–$NDVI_{SR}$ relationship and the cross-sensor $NDVI_{SR}$–$NDVI_{SR}$ relationship should all be taken into account.

Meanwhile, these factors, which include the air conditions, solar viewing angles, topography, elevation, and so on (Martínez-Beltrán et al. 2009), are known to be spatially heterogeneous. As we need to build the relationship between multi-sensor $NDVI_{DN}$/$NDVI_{TOA}$ and the $NDVI_{SR}$, it is therefore important to consider the characteristics of the factors which might affect this relationship. It has already been proved in related research that it is important to consider the spatial heterogeneity of such factors (Obata et al. 2013). With this in mind, we propose a method that introduces the idea of 'local' to the GCLM, where the whole image is processed block by block. The main workflow is to build a local linear model between the resampled NDVI_O and NDVI_R for each cluster at the coarse resolution and then to apply them to the NDVI_O to obtain the normalized result. Both of these steps are implemented block by block.

The workflow of the LCLM (Figure 4) can be described as follows. First of all is the acquisition and pre-processing of the material, including the original NDVI_O, NDVI_R,

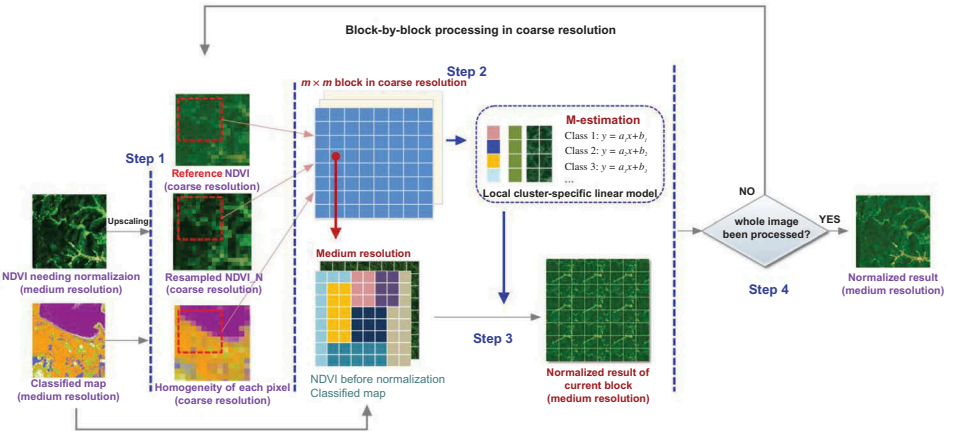


Figure 4. Flow diagram of the LCLM being used under the framework of normalization of medium-resolution NDVI by the use of coarse-resolution NDVI as the reference.

and the medium-resolution classification map. Both the NDVI_O and the classification map are registered and re-projected to be consistent with the NDVI_R. A cloud mask is also applied to both the NDVI_O and NDVI_R. The remaining tasks can be separated into the following steps.

Step 1. Based on the medium-resolution classification map, the majority spectral cluster type of each coarse-resolution pixel of the NDVI_R is computed. The classification map can be obtained from existing data or through the unsupervised classification of the corresponding medium-resolution imagery. The homogeneous pixels are used in the next step to build the inter-sensor NDVI transformation model.

Step 2. For current block j_{temp}, the models are built based on the homogenous pixels of each cluster belonging to an $m \times m$ (block size) block around the current block centre. This can be expressed as:

$$y_{c,i} = a_{i,j_{\text{temp}}} \times x_{c,i} + b_{i,j_{\text{temp}}}, i = 1, 2, 3, \ldots, nc. \tag{3}$$

Here, $x_{c,i}$ and $y_{c,i}$ represent the resampled NDVI_O and NDVI_R at a coarse resolution for cluster i in block j_{temp}. $a_{i,j_{\text{temp}}}$ and $b_{i,j_{\text{temp}}}$ represent the coefficients of the linear model for cluster i. The lowest number of samples required for the normalization is set in order to quantify the accuracy of the regressions.

Step 3. Apply the models to the original medium-resolution NDVI_O and produce an output with a medium resolution with the transformation coefficients being chosen according to their related clusters. For the small clusters that lack enough homogenous samples for model building, the GCLM built in the whole image is used instead.

$$y'_m = a_{i,j_{\text{temp}}} \times x_m + b_{i,j_{\text{temp}}}, i = 1, 2, 3, \ldots, nc. \tag{4}$$

Here, x_m represents a pixel of the medium-resolution NDVI_O, and it belongs to cluster i in current block j_{temp}. y'_m represents the output value for the pixel, NDVI_N, with a medium resolution.

Step 4. Turn to the next block with a step of s after the processing of the current block. The same processes of Steps 2 and 3 are carried out until the whole image is processed. The whole image is processed block by block in raster-scan order, from left to right and top to bottom. s can be set between 1 to m; if it is set as 1, this means that the block centre is moved pixel by pixel at the coarse resolution, while m means that each coarse pixel belongs to only one block. As pixels might belong to different blocks, the final normalization result is set as the average of all the possible output values.

Totally, parameters needed during the normalization using LCLM include: (1) The classification map, of which the essence is the number of clusters; (2) The threshold of homogeneity for determination of the pure samples; (3) The minimum number of samples required for the establishment of the linear models; (4) The size of the block; (5) Step of the block. Among these, both the parameters (2) and (3) could be set respectively during the global and the local process. We have found that a cluster number around 6–10 is appropriate. The threshold of homogeneity is suggested to be greater than 50% to exclude pixels with a large scale-induced error. Meanwhile, the block size is not very sensitive, and a good result could be obtained with a block size ranging from 100 to 300. These two parameters can be set a little smaller when the cluster number is increased. Lastly, the step of block can be set small to avoid possible blockiness brought by the blockly process.

During the workflow, the linear regression coefficients for each cover type in Step 2 for each cluster can be solved by various optimization estimation methods. The most widely used method is the least-squares estimator. However, the least-squares estimator is sensitive to both the distribution assumption of the noise and the existence of outliers, and can be easily affected by bad data. Thus, M-estimation is utilized here to solve the coefficients. This estimator is able to identify and remove outliers, so as to weaken their effect on the parameter estimation and improve the accuracy of the estimated parameters (Zhang 1997; Jiang et al. 2006). In addition, the Huber cost function is chosen to determine the associated weights for each point (Huber 1964).

During the solving of coefficients $a_{i,j_{\text{temp}}}$ and $b_{i,j_{\text{temp}}}$ for cluster i in block j_{temp} in Equation (3), let e_{uv} be the residual of the uvth data point, which is one of the samples used for the solution:

$$e_{uv} = a \times x_{uv} + b - y_{uv}. \tag{5}$$

The standard least-squares method attempts to minimize $\min \sum e_{uv}^2$, but the M-estimator tries to reduce the effect of outliers by replacing the squared residuals e_{uv}^2 by another function of the residuals (Obata et al. 2012):

$$\min \sum_{uv} \rho(e_{uv}), \tag{6}$$

where $\rho(\bullet)$ is a symmetric, positive-definite cost function with a unique minimum of zero. There are many different kinds of cost function, but here we use the Huber cost function:

$$\rho(e_{uv}) = \begin{cases} e_{uv}^2/2, & |e_{uv}| \le k \\ k|e_{uv}| - k^2/2, & |e_{uv}| > k \end{cases}, \tag{7}$$

where k is the Huber parameter. In order to facilitate the optimization process, the minimization problem of Equation (8) can be equivalently converted into an iteratively reweighted least-squares problem:

$$\min \sum (w(e_{uv})e_{uv})^2, \tag{8}$$

where

$$w(e_{uv}) = \frac{\rho'(e_{uv})}{e_{uv}} \tag{9}$$

represents the weight, and $\rho'(e_{uv})$ is the derivative of the Huber function $\rho(e_{uv})$. By this means, the coefficients can be solved by an iterative procedure (Fox and Weisberg 2010).

3. Evaluation methods

The reference-based method is designed to normalize the NDVI by making it consistent with the reference, while keeping the original spatial resolution. Considering this goal, the protocol of the evaluation could be described as: the normalized result of sensor_A (spatial resolution R_A) by the use of coarse-resolution reference NDVI from sensor_B (spatial resolution R_B) should be consistent with the reference but with the original spatial resolution R_A.

However, no real data with this feature exist, which makes it difficult to undertake an evaluation. Previous studies have described various approaches for the evaluation of the normalization of medium-resolution data by using the reference data (Olthof et al. 2005; Gao et al. 2010; Gan et al. 2013). However, we believe that the problem of how to give a quantitative accuracy assessment for a reference-based normalization method is a problem that still needs further study.

In following section, we first summarize the existing evaluation methods, and we then give an analysis of their applicability for NDVI normalization. Three new schemes designed in accordance with the above principle are then described in detail. Finally, a discussion about the given schemes, and advice for the choice of these methods is given.

3.1. Existing methods

3.1.1. Evaluation method based on the idea of cross-validation (emCV)

Evaluation has been undertaken by making use of the idea of cross-validation, which removes part of the samples and predicts their value by the use of other samples. It is implemented by applying the inter-sensor transformation coefficients to a reserved subset of the upscaled data and checking their residual with the reference data. The coefficients are obtained by using another part of the upscaled data. In the earlier case of normalization of medium-resolution Landsat ETM+ data using the coarse-resolution data from the SPOT VEGETATION (VGT) sensor as reference (Olthof et al. 2005), the workflow of emCV is shown in Figure 5:

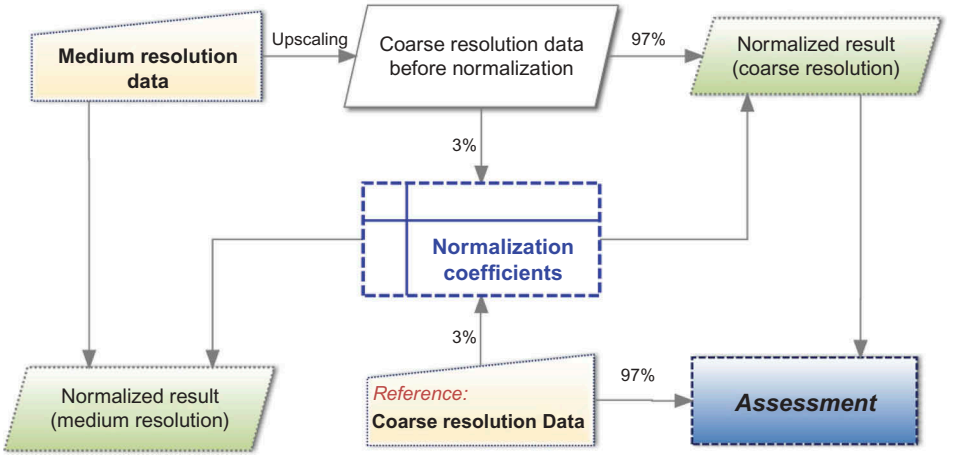


Figure 5. Flowchart of emCV, which applies normalization coefficients to part of the resampled data.

What this method evaluates is the consistency between the normalized results of the resampled data and reference data. It lacks direct evaluation of the medium-resolution normalized result, but evaluates the effectiveness of the normalization coefficients by applying them to upscaled data. For scale-dependent parameters such as NDVI, it is not reasonable to apply the normalization coefficients to the heterogeneous resampled NDVI, as it might induce bias, particularly for the heterogeneous pixels. Furthermore, we might be confronted by a lack of samples when applying the coefficients to homogeneous resampled NDVI, as most of the homogeneous samples will have already been used for building the models, especially for LCLM which needs to build the local models.

3.1.2. Evaluation method based on a comparison with atmospherically corrected data (emCAC)

The emCAC method is based on a comparison between the normalized result and data acquired from a physical atmospheric correction (comparison I in Figure 6). However, intrinsic differences exist between these two types of data, and thus the comparison is not sufficient to evaluate the ability to achieve consistency between the normalized result and the reference data. The evaluation method can, however, be improved with the assistance of analysing the difference between the reference data and the data from the physical method. This evaluation method was utilised by Gao et al. (2010) in a study of the normalization of Landsat ETM+ medium-resolution reflectance data by taking a coarse-resolution MODIS product as the reference. Gao et al. (2010) concluded that the difference between the normalized result and the data from the physical method was mainly due to the difference between the reference data and the data from the physical method.

In summary, this method can give a general evaluation of the effectiveness of eliminating the influence of the atmosphere. However, as a result of the intrinsic differences existing between the reference data and the data from the physical method, the method is inevitably disturbed by the differences. Thus, this method should be used as a supplement for other evaluation methods, instead of being utilized individually.

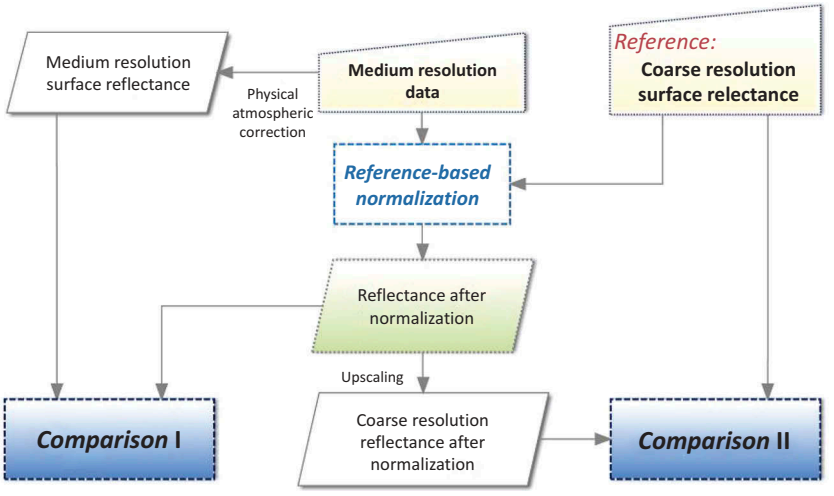


Figure 6. Flowchart of emCAC and emCUS, which compare the reflectances from different approaches.

3.1.3. *Evaluation method based on a comparison between the upscaled normalized result and the reference data (emCUS)*

Evaluation can also be undertaken by a comparison between the resampled normalized result and the reference data. The workflow of this approach is shown as comparison II in Figure 6. The normalized result is upscaled to the same resolution as the reference. This method was also utilized for the evaluation of the work introduced in Section 3.1.2 by Gao et al. (2010).

This method can give an evaluation of the consistency between the normalized result and the reference data. It can be used for real conditions, and does not require any additional data. Such an idea is common for the validation of research when there is a lack of 'standard' data, such as an image fusion method performed on real observed images (Zhang et al. 2012; Fox and Weisberg 2010). However, it is not good practice to utilize the resampled reflectance pixels in both the model establishment and the comparison. And as the comparison is undertaken in the coarse-resolution through upscaling the normalized result, scale-induced error will inevitably influence the evaluation accuracy when applying this method to scale-dependent parameters.

3.1.4. *Evaluation method based on a comparison of the normalized results of multiple sensors (emCMS)*

Another important evaluation scheme that has been widely used is to utilize the statistics that are generated from overlapping regions of the normalized result from different sensor systems, which are assumed to be the same after spatial normalization. Taking the normalized data from two sensors as an example, the flowchart of emCMS is shown in Figure 7.

The data from synchronous images of two sensors are first normalized respectively, by using the same coarse-resolution reference data. The normalized results in the overlapping region of the two images are then compared. The statistics are calculated on a pixel-by-

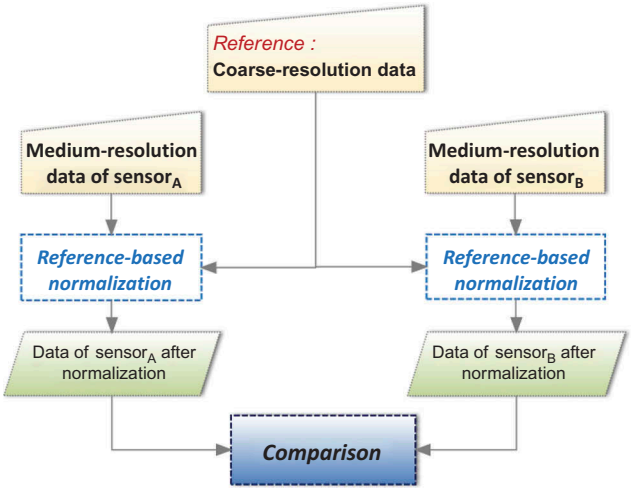


Figure 7. Flowchart of emCMS, which evaluates the performance of normalized multi-sensor data.

pixel basis, including the mean and standard deviation of the differences and absolute differences.

This scheme can effectively evaluate the consistency between the normalized results from different sensors, and is suitable for the evaluation of real spatial normalization applications such as multi-sensor data mosaicking. It has therefore been a widely used method in similar research (Gao et al. 2010; Obata, Miura, and Yoshioka 2012). It should, however, be noted that this method is not able to confirm whether the normalized result is consistent with the reference data. Thus, this method is best used by combining it with an evaluation method which can remedy this defect, to allow a full evaluation of the normalization method.

3.2. *The proposed methods*

3.2.1. *Evaluation method based on the use of only one medium-resolution image (emOMRI)*

The emOMRI scheme introduced in this section is designed to evaluate the effectiveness of the normalization in eliminating the influence of the atmosphere. It is implemented by the use of the synthetic reference data produced through UP_R (refer to Section 2.1) using medium-resolution reflectance data from the physical atmospheric correction. Instead of upscaling by UP_N, upscaling by UP_R can avoid the scale-induced bias of the coarse-resolution NDVI brought by the upscaling operation. A flowchart of the scheme is shown in Figure 8. The workflow can be decomposed to the following steps.

- First, carry out atmospheric correction of the raw data using a physical atmospheric correction model such as 6S (Vermote et al. 1997) or MORTRAN, and obtain the surface reflectance data of different bands.
- Coarse-resolution $NDVI_{SR}$ is obtained through UP_R based on the surface reflectance data and then it was used as the reference.

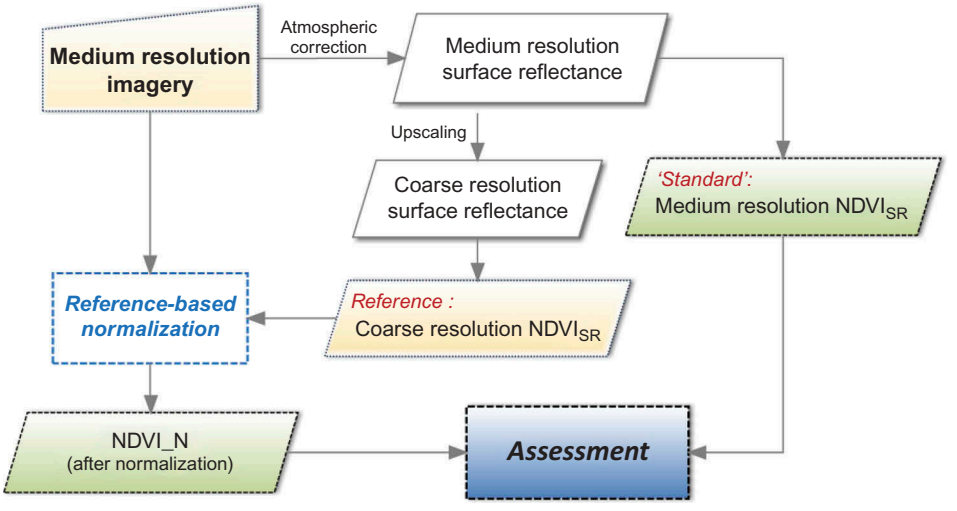


Figure 8. Flowchart of the proposed emOMRI evaluation scheme to evaluate the effectiveness of eliminating the influence of the atmosphere.

- Through the reference-based method, the raw medium-resolution NDVI_O is normalized, and the NDVI_N is obtained.
- The medium-resolution NDVI$_{SR}$ is calculated based on the surface reflectance and used as the 'standard' data for the evaluation. The difference between the NDVI_N and NDVI$_{SR}$ is then checked to assess the accuracy of the method.

This scheme is able to assess the performance of an empirical model in eliminating atmosphere-induced variance, avoiding being disturbed by the bias introduced by other factors, such as spectral response function differences and geo-rectification residuals. In addition, the method is easily implemented and only requires a small amount of data. However, the method is not comprehensive enough to reflect the performance of a normalization method when sensor characteristic differences exist between the reference data and the data needing normalization, such as band position, bandwidth, and spectral response function.

3.2.2. Evaluation method based on the use of synchronous medium-resolution images (emSMRI)

The evaluation scheme proposed in this section is designed to evaluate the effectiveness of the normalization when the sensors have different characteristics. This scheme is applied based on a pair of synchronous images from sensors with a similar spatial resolution. The coarse-resolution NDVI$_{SR}$ is produced by UP_R using data from one of the two sensors, and data from one of the two images is utilized to produce the coarse-resolution NDVI$_{SR}$ by UP_R, to provide a synthetic reference for the normalized NDVI from the other sensor. The NDVI$_{SR}$ calculated by its original medium-resolution surface reflectance is used as the 'standard' for the evaluation. The difference between the normalized result and the 'standard' is evaluated to assess the accuracy of the method. A flowchart of the scheme is shown in Figure 9.

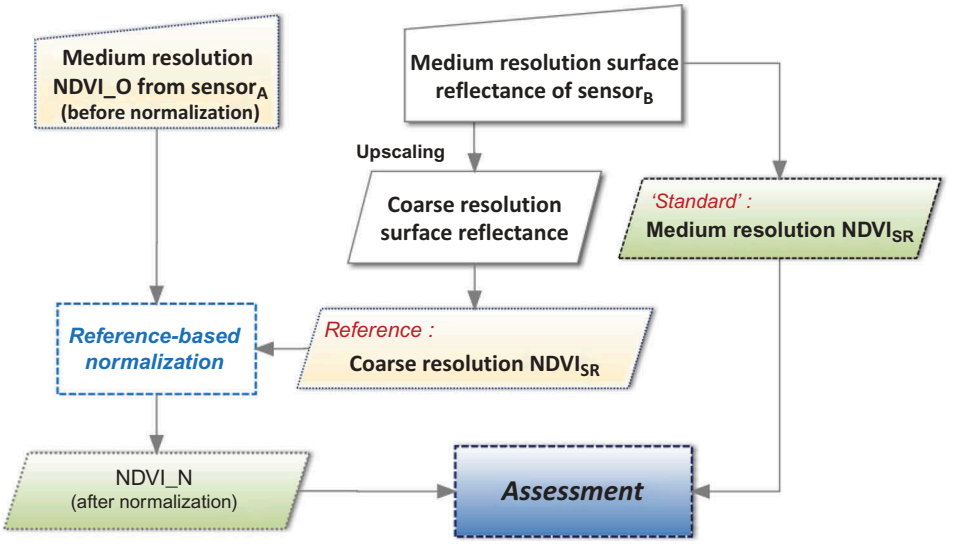


Figure 9. Flowchart of the proposed emSMRI evaluation scheme using synthetic reference data produced by synchronous medium-resolution images from sensors with different characteristics.

Sensor characteristic differences exist between the sensors providing the NDVI_R and NDVI_O in this evaluation method. Thus, this scheme is more efficient than the first scheme, due to the additional consideration of the differences in the sensor characteristics, and it could be used as a supplement to emOMRI, as introduced in Section 3.2.1. However, this scheme is strict in its data requirements, and synchronous medium-resolution data from another sensor are essential, which can sometimes be difficult to obtain.

3.2.3. Evaluation method based on the use of synchronous coarse-resolution reference data through upscaling (emSCU)

The emSCU scheme introduced in this section also considers the differences in sensor characteristics. In brief, the scheme is applied with the assistance of an 'upscale' process and those easily available synchronous coarse-resolution data, such as MODIS or AVHRR. Thus it can avoid the shortcoming of the data requirement for the method in Section 3.2.2. A flowchart of the scheme is shown in Figure 10. The scheme can be separated into the following steps.

- Upscale the original medium-resolution NDVI_O from R_M (the original resolution of NDVI_O) to R_C (the resolution of the synchronous coarse-resolution data). Meanwhile, the synchronous coarse-resolution reflectance data is upscaled from R_C to $R_C \times rr$, where $rr = R_C/R_M$.
- Calculate NDVI based on the upscaled synchronous reflectance data with resolution $R_C \times rr$, and the NDVI then serves as NDVI_R, the reference data for following normalization process.
- Normalize the upscaled NDVI_O and obtain the NDVI_N, of which the resolution is R_C.

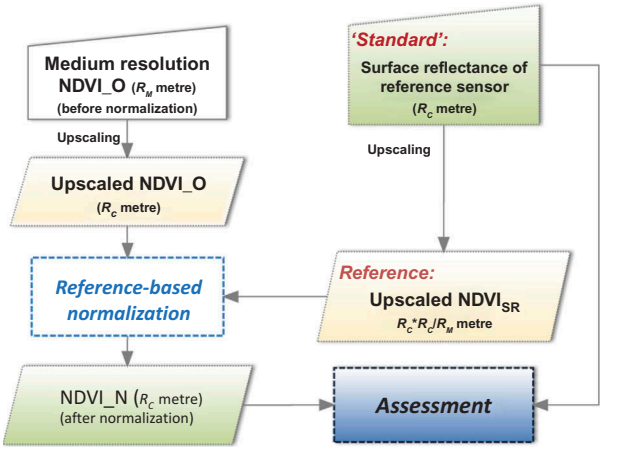


Figure 10. Flowchart of the proposed emSCU evaluation scheme using synchronous coarse-resolution reference data through upscaling.

- Calculate the $NDVI_{SR}$ by the previous synchronous reference surface reflectance data. It is then used as the 'standard' data for evaluation.

This method is able to evaluate the performance of normalization method when sensor-characteristic differences exist between the NDVI_R and NDVI_O. It is implemented with the assistance of the upscaling with the assumption that the normalization method will be effective at a coarse resolution if it performs well at the higher resolution. It should be noted that the image after upscaling should be wide enough to provide enough points to obtain robust normalization coefficients. Furthermore, the upscaling may introduce errors as a result of the pixel mixture, which can affect the accuracy of the evaluation.

3.3. *Choice of evaluation method*

Table 2 gives the summarization of all of the above evaluation method. Of these evaluation methods, emCV, emCAC, emCUS, emOMRI, emSMRI, and emSCU are useful for evaluating the accuracy of the normalization by checking whether the difference between the data needing normalization and the data used as the reference is eliminated. Differing from these six methods, emCMS is able to measure the consistency of multiple data after spatial normalization and therefore is an indispensable evaluation method.

These schemes can be utilized individually or together, according to the specific conditions in practice, to give an overall evaluation of the normalization method. The methods proposed in this paper are suitable for the NDVI and can provide accurate quantitative evaluations for reference-based normalization methods. Meanwhile, they can also provide a reference for the evaluation of this kind of reference-based normalization method, for the NDVI or other parameters.

It should be pointed out that although we have tried to eliminate the scale effect in our normalization method, based on utilizing pixels with a high homogeneity, the inclusion of samples with a small scale-induced error in the regression means that the normalized

Table 2. Information about the evaluation methods, taking TM and MODIS NDVI data as example.

		Normalization				Evaluation				
		Original Data	Reso.*	Reference data	Reso.	Result for evaluation	Reso.	'Standard' data	Reso.	Statistic
Existing methods	emCV	TM data	30 m	Part of the MODIS data	250 m	Normalized result of resampled TM data	250 m	The rest of the MODIS data	250 m	Statistic of difference between the result and the 'standard' data: R^2 MRE MAE MSE
	emCAC	TM data	30 m	MODIS data	250 m	Normalized result	30 m	Data from physical correction method	30 m	
	emCUS	TM data	30 m	MODIS data	250 m	Resampled normalized result	250 m	The MODIS data	250 m	
	emCMS	Multiple TM data	30 m	MODIS data	250 m	Normalized result of one TM data	30 m	Normalized result of another TM data	30 m	
Proposed methods	emOMRI	TM data	30 m	Resampled TM data after physical atmospheric correction	250 m	Normalized result	30 m	TM data after physical atmospheric correction	30 m	
	emSMRI	TM data	30 m	Resampled synchronous data from another sensor after physical atmospheric correction	250 m	Normalized result	30 m	Synchronous data from another sensor after physical atmospheric correction	30 m	
	emSCU	Resampled TM data	250 m	Resampled MODIS data	2000 m	Normalized result	250 m	MODIS data	250 m	

*Resolution

result is still influenced by the scaling effect. Therefore, in practical applications, it is difficult to obtain a normalized result that is exactly the same as the 'standard' data.

In the following section, emOMRI, emSCU, emSMRI, and emCMS are chosen. The first three methods focus on an accuracy evaluation of the normalized result to measure the effectiveness of the normalization in eliminating the effects of the atmosphere, sensor characteristic differences, solar and viewing angle differences, and so on. Finally, emCMS is based on evaluating the effect of the real normalization application of the NDVI from different sensors.

4. Experiments and analysis

A number of experiments were undertaken to evaluate the accuracy of the LCLM and to evaluate its performance in the normalization of the NDVI from different sensors. The GloLM and the GCLM were also tested, to examine whether the LCLM is superior for the normalization of medium-resolution NDVI by the use of coarse-resolution NDVI products as the reference. In addition, an evaluation experiment for a real application, combining the NDVI from Terra ASTER and Landsat ETM+ sensors, was also undertaken to check the effectiveness of the normalization of the NDVI from different sensor systems. During these experiments, the statistical values of the differences were calculated, including R^2 (coefficient of determination), MAD (mean absolute difference), and MRD (mean relative difference), to evaluate the accuracy of the normalization method and to check the consistency between the normalized result and the reference data. The information of the following experiments is stated in Table 3, including the parameters setting and the percentage of homogeneous pixel.

4.1. Evaluation experiment using emOMRI

A scene of Landsat-7 ETM+ data from around Khanka Lake was chosen for this experiment (Figure 11(*a*)). Khanka Lake is located on the border between Primorsky Krai, Russia and Heilongjiang province, Northeast China (WRS-2 path 114 and row 29), and the imagery was acquired on 25 September 2001.

First, the imagery was atmospherically corrected to acquire the band surface reflectance. It was then aggregated to a coarse resolution, based on which the NDVI was calculated to provide the synthetic reference. At the same time, the NDVI in the original resolution was calculated by surface reflectance to provide the 'standard' for the evaluation (Figure 11(*c*)). Finally, the difference between the NDVI_N and $NDVI_{SR}$ at the

Table 3. Setting of the parameters for the experiments in Section 4.

	P. 1*	P. 2	P. 3	P. 4	P. 5	Percentage**
4.1 (250 m)	6	0.95	40	100	50	0.3972
4.1 (1000 m)	6	0.90	40	100	50	0.1933
4.2	6	0.95	40	200	100	0.1918
4.3	8	0.94	20	60	30	0.2097
4.4 (ASTER)	6	0.60	20	60	30	0.5373
4.4 (ETM+)	6	0.60	20	60	30	0.7371

* The number of parameter (refer to Section 2.4).
** The percentage of homogeneous pixels (pixels with homogeneity larger than the given threshold).

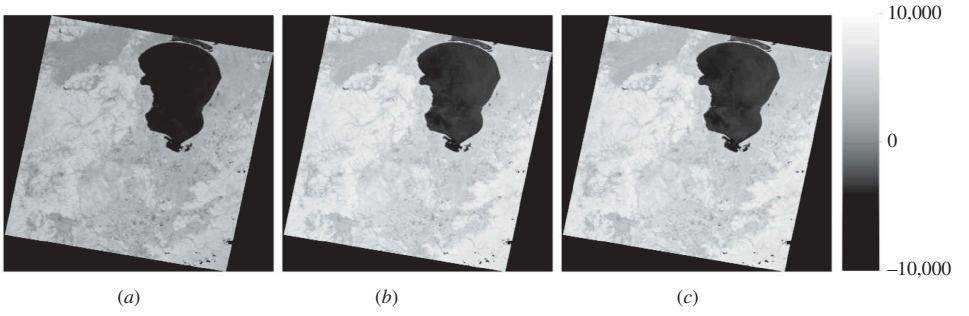


Figure 11. (*a*) The NDVI calculated by DN before normalization; (*b*) normalized result of the NDVI by the use of the LCLM; (*c*) the $NDVI_{SR}$ used as the 'standard' for evaluation, for the Khanka Lake area (Landsat WRS-2 path 114 and row 29).

original resolution was examined. In order to assess the performance at different scales, the synthetic reference NDVI was obtained at two scales: 250 and 1000 m.

The LEDAPS (Landsat Ecosystem Disturbance Adaptive Processing System) project was developed to create a Landsat-based surface reflectance product using the MODIS 6S approach for the atmospheric correction of Landsat TM and ETM+ data. With the aerosol thickness derived from the imagery itself, LEDAPS has been successfully used for atmospheric correction (Masek et al. 2006; Vermote, El Saleous, and Justice 2002), and the accuracy of the LEDAPS approach has been found to be only slightly worse than the MODIS-based atmospheric correction method (Ju et al. 2012).

The normalized result (Figure 11(*b*)) shows a similar spatial pattern and distribution to the 'standard', the ETM+ $NDVI_{SR}$ (Figure 11(*c*)), but changes greatly when compared to its own $NDVI_{DN}$. As shown in Figure 12(*b*), a strong linear relationship between the NDVI_N by LCLM (by the use of 250-m synthetic NDVI as the reference) and the

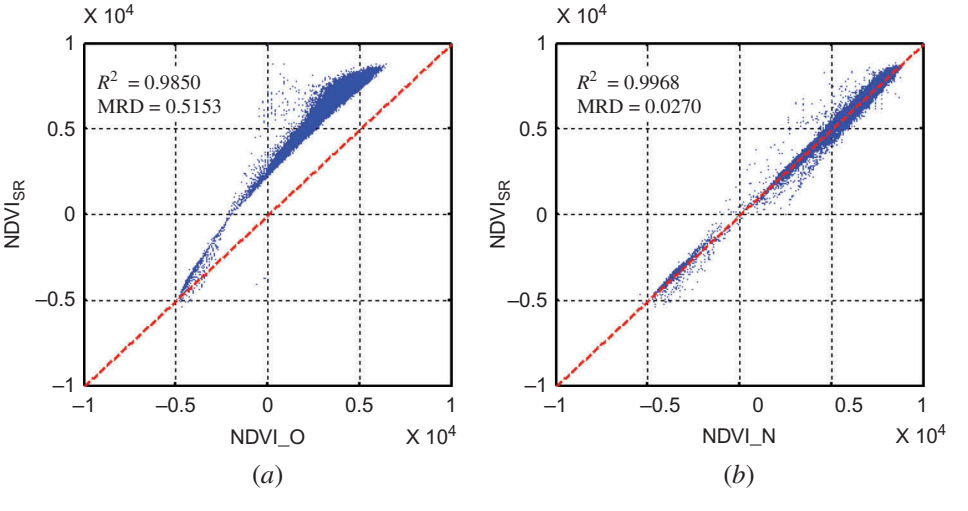


Figure 12. (*a*) Scatterplot of the NDVI before normalization and the 'standard' NDVI calculated by surface reflectance (*10,000); (*b*) scatterplot of the NDVI after normalization and the 'standard' NDVI calculated by surface reflectance (*10,000).

Table 4. R^2, MAD and MRD between the 'standard' NDVI calculated by surface reflectance and the NDVI after normalization, at 1000 and 250 m resolutions, respectively.

		1000 m			250 m		
		R^2	MAD	MRD	R^2	MAD	MRD
Before normalization		0.9850	0.2635	0.5153	0.9850	0.2635	0.5153
After normalization	GloLM	0.9852	0.0319	0.0647	0.9852	0.0322	0.0665
	GCLM	0.9930	0.0161	0.0341	0.9963	0.0141	0.0281
	LCLM	0.9943	0.0150	0.0322	0.9968	0.0126	0.0270

'standard' NDVI can be observed from the scatterplots which are tight to the 1:1 line. The difference between the data needing normalization and the reference is almost eliminated through the normalization, no matter which method is used, and the results are consistent with those of the physical atmospheric correction method. The R^2 value is as high as 0.9968 (Table 4), while the MRD is less than 0.06, compared to more than 0.5 before normalization. This illustrates that the normalization method using coarse-resolution reference data can effectively eliminate the influence of the atmosphere. For the 1000 m resolution, the normalization also works well, but the results are slightly poorer than for the 250 m resolution, which may be due to the mixture induced by upscaling.

Moreover, from Table 4 it is clear that of the three methods, the GCLM and LCLM can get a better result than the GloLM, which indicates that the 'cluster-specific' strategy can indeed bring about an improvement. The LCLM shows a slight improvement when compared to the GCLM (MRD: 0.0270 vs. 0.0281 at the 250 m resolution; 0.0322 vs. 0.0341 at the 1000 m resolution). In addition, the normalized error of the LCLM (MAD = 0.0126) is less than that of the GCLM (MAD = 0.0141) and GloLM (MAD = 0.0322), which illustrates the superiority of the 'block-by-block' process adopted by the LCLM.

However, as Figure 12(*b*) shows, the relationship deviates from the exact 1:1 line. This implies that the NDVI_N from the empirical approach may not precisely duplicate the standard $NDVI_{SR}$. This might be due to the difference caused by the mixture of different scales between the original $NDVI_{SR}$ at a 30 m × 30 m resolution and the $NDVI_{SR}$ at a 250 m × 250 m resolution. Alternatively, it may be because the cluster-specific linear models only approximately represent the non-linear relationship between the $NDVI_{DN}$ and $NDVI_{SR}$.

4.2. *Evaluation experiment using emSMRI*

Synchronous images from Landsat ETM+ and Terra ASTER sensors were used for the evaluation experiment using emSMRI. These scenes were acquired on 16 August 2000, around Hopkinsville in Kentucky, USA. The coverage of the two images is shown in Figures 13(*a*) and (*b*). Since this pair of images observe the same land surface at the same time, their NDVIs are theoretically required to be the same for accurate earth science community downstream research. However, because of the difference in the sensor characteristics, and some other factors, they are not equivalent. The normalization method proposed in this paper should, however, be able to eliminate the difference between them.

First of all, the ASTER image was resampled to 30 m, the same scale as the ETM+ image, and it was then registered to the ETM+. The ETM+ data were then atmospherically

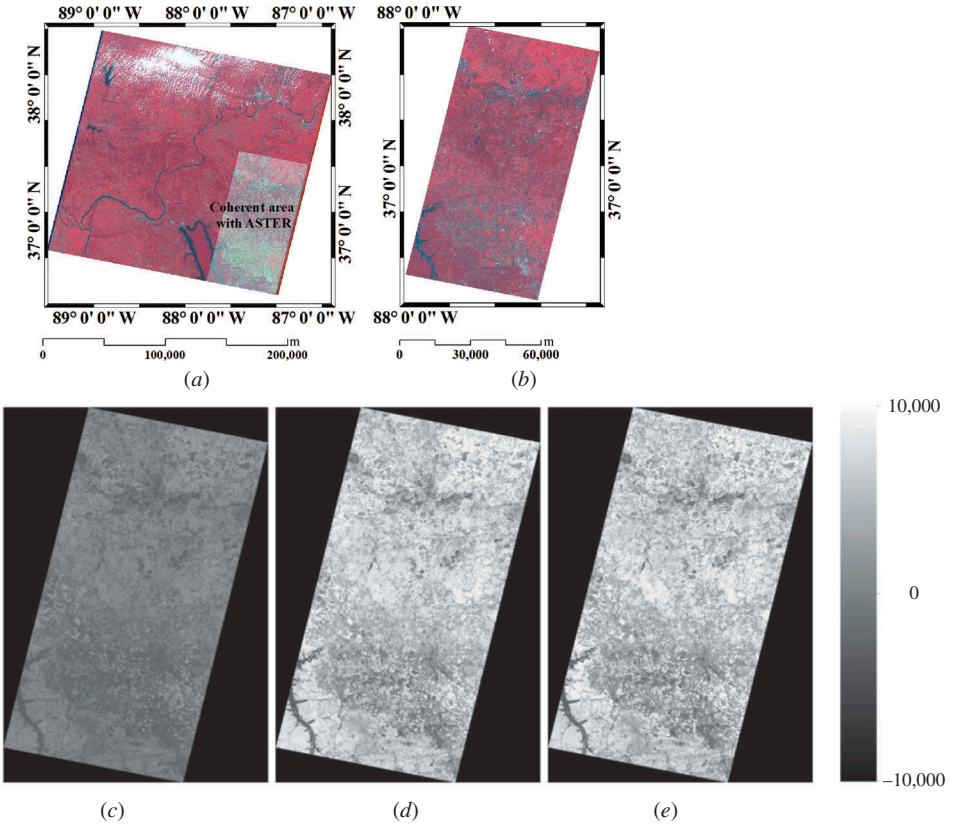


Figure 13. (*a*) False-colour composite ETM+ image; (*b*) false-colour composite of the synchronous ASTER image; (*c*) the NDVI calculated by DN, which needed normalization; (*d*) normalized result of the ASTER NDVI produced by the use of the LCLM; (*e*) the $NDVI_{SR}$ of ETM+, used as the 'standard' for the Hopkinsville area (Landsat WRS-2 path 22 and row 34).

corrected by the use of LEDAPS (refer to Section 4.1.1), and the surface reflectance was upscaled to a 250 m resolution by pixel aggregation. This was then used to calculate the NDVI, which was used as the reference. The NDVI calculated by DN value from the ASTER data was normalized to obtain the NDVI_N using the synthetic reference. Finally, the NDVI_N of the ASTER scene was compared with the ETM+ NDVI calculated by surface reflectance.

The NDVI_N (Figure 13(*e*)) shows a similar spatial pattern and distribution to the ETM+ $NDVI_{SR}$ (Figure 13(*d*)), but changes greatly when compared to its own $NDVI_{DN}$ before normalization (Figure 13(*c*)). The difference between the data needing normalization and the reference was almost eliminated through the normalization. As shown in Figure 14(*b*), there is a strong linear relationship between the NDVI_N of the ASTER imagery using the LCLM and the 'standard', the $NDVI_{SR}$ of the ETM+ imagery, with the points being tight around the 1:1 line, which is in contrast with the points distributed away from the 1:1 line before normalization (Figure 14(*b*)). In addition, the R^2 increased to 0.8351 (Table 5), while the MRD was reduced to 0.0591 and the MAD to 0.0350, which is in great contrast with the values before normalization.

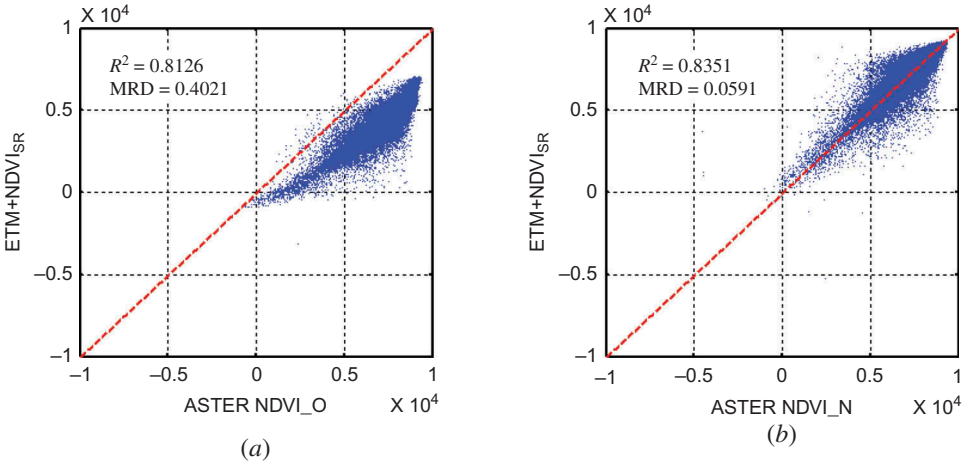


Figure 14. (*a*) Scatterplot of the ASTER NDVI before normalization and the ETM+ NDVI calculated by surface reflectance, used as the 'standard' for evaluation (*10,000); (*b*) scatterplot of the normalized result of the ASTER NDVI produced by the use of the LCLM and ETM+ NDVI calculated by surface reflectance, used as the 'standard' for evaluation (*10,000).

Table 5. R^2, MAD and MRD between the $NDVI_{SR}$ of ETM+ used as the 'standard' and the normalized result of the ASTER NDVI produced using the LCLM.

		R^2	MAD	MRD
Before normalization		0.8126	0.2932	0.4021
After normalization	GloLM	0.8126	0.0405	0.0762
	GCLM	0.8314	0.0362	0.0606
	LCLM	0.8351	0.0350	0.0591

Furthermore, a more consistent result was obtained through the LCLM (R^2: 0.8351) than the GCLM (0.8314) and GloLM (0.8126). The normalized error of the LCLM (MAD = 0.0350) was less than those of the GCLM (MAD = 0.0362) and GloLM (MAD = 0.0405), and the relative normalized error of the LCLM was also less than those of the GCLM and GloLM (MRD = 0.0591 vs. 0.0606 vs. 0.0762). This implies that the LCLM performed better than the other two models.

4.3. *Evaluation experiment using emSCU*

An evaluation experiment using emSCU was undertaken to evaluate the effectiveness of the normalization methods in eliminating the variance between the data needing normalization and the data used as the reference. The ETM+ imagery used in Section 4.1 was again utilized. A MODIS daily surface reflectance product (MOD09) acquired on the same day was used to provide the reference data (Figure 15(*a*)).

The ETM+ image and the MODIS surface reflectance product were both first upscaled by a scale factor of about 1/8. The ETM+ image was upscaled to the same resolution as the original MOD09 product. The NDVI calculated by the upscaled ETM+ DN (Figure 15(*b*)) was then normalized by the use of the NDVI calculated by the upscaled MODIS

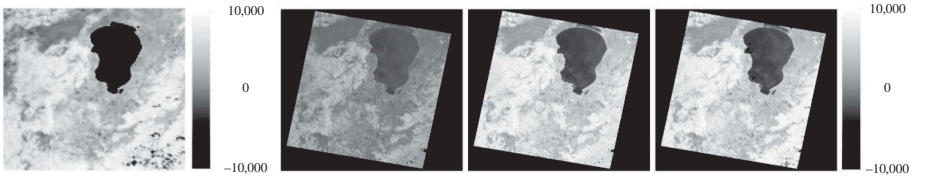


Figure 15. (*a*) The upscaled MODIS $NDVI_{SR}$ used as the reference for the normalization; (*b*) the upscaled $NDVI_{DN}$ before normalization; (*c*) normalized result of the upscaled $NDVI_{DN}$ produced by the use of the LCLM; (*d*) the MODIS $NDVI_{SR}$ at the original spatial resolution, used as the 'standard' for evaluation for the Khanka Lake area (Landsat WRS-2 path 114 and row 29).

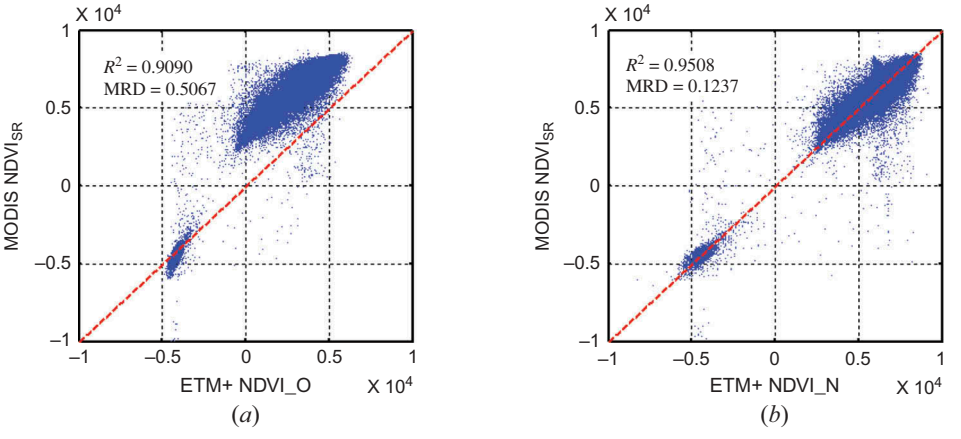


Figure 16. (*a*) Scatterplot of the ETM+ NDVI before normalization and the MODIS NDVI calculated by surface reflectance, used as the 'standard' for evaluation (*10,000); (*b*) scatterplot of the ETM+ NDVI normalized result produced by the use of the LCLM and MODIS NDVI calculated by surface reflectance, used as the 'standard' for the evaluation (*10,000).

surface reflectance product as the reference (Figure 15(*a*)). Finally, the normalized result (Figure 15(*c*)) of the upscaled ETM+ $NDVI_{DN}$ was evaluated by the use of the 'standard' NDVI calculated by the original MOD09 product (Figure 15(*d*)).

The normalized result (Figure 15(*c*)) changed greatly when compared to its original $NDVI_{DN}$ before normalization, with a similar spatial pattern and distribution to the 'standard' (Figure 15(*d*)). The difference between the data needing normalization and the reference was eliminated by the normalization. The scatterplot in Figure 16 illustrates that there is strong linear correlation between the NDVI_N and the 'standard', with the points being tight around the 1:1 line. In addition, the R^2 is high at 0.9508 (Table 6), which means that there is great consistency between the NDVI after normalization.

As shown in Table 6, it is clear that the performance of the LCLM was better than the GCLM and GloLM. R^2 increased from 0.9489 (GCLM) to 0.9508 (LCLM). In addition, MAD decreased to 0.0508 (LCLM), compared to 0.0522 (GCLM) and 0.0838 (GloLM), while MRD decreased to 0.1237, compared to 0.1257 (GCLM) and 0.1795 (GloLM). These figures indicate that the LCLM performed better than the other two methods. The LCLM is able to capture the difference between ETM+ and MODIS well, and thus can

Table 6. R^2, MAD and MRD between the MODIS $NDVI_{SR}$ and the NDVI_N from the LCLM.

		R^2	MAD	MRD
Before normalization		0.9090	0.2745	0.5067
After normalization	GloLM	0.9092	0.0838	0.1795
	GCLM	0.9489	0.0522	0.1257
	LCLM	0.9508	0.0508	0.1237

successfully produce an NDVI that is much more consistent with the reference, while keeping the original high resolution.

4.4. *Evaluation experiment using emCMS for a real application*

An application experiment involving the normalization of the NDVI from different sensors was also undertaken. Synchronous Landsat ETM+ and ASTER images were chosen for the experiment. A scene of Landsat-7 ETM+ data (WRS-2 path 19 and row 33) acquired on 21 May 2005 and three scenes of ASTER data acquired on the same day were used (Figure 17(*a*)). Their overlapping region was around Portsmouth, located on the border between Kentucky, Ohio and West Virginia, USA. A MODIS daily surface reflectance product (MOD09) acquired on the same day was chosen to produce the reference NDVI data.

The MODIS product was first re-projected and resampled to the same projection as the ETM+ scene. A pixel-based quality file was then used to filter out clouds and cloud shadows. Only high-quality and clear pixels were used for establishing the linear model. The ASTER scenes were mosaicked and geo-rectified to the ETM+ image. Both the ASTER and ETM+ data were also resampled to 30 m × 30 m for consistency. The $NDVI_{DN}$ from the ETM+ (Figure 17(*c*)) and ASTER (Figure 17(*d*)) images were then normalized, respectively using the LCLM with the same MODIS NDVI data as the reference.

The two normalized results were compared to assess the effectiveness of the normalization by emCMS, as introduced in Section 3.2.3. After the normalization, the difference between the NDVI from the two images was slight. The density scatterplot (Figure 18) between the two sets of NDVI_N shows that they agree closely, and the difference between the two sets of NDVI_N is very small. In addition, the bright colour, which means a high density, is tight to the 1:1 line, illustrating that the difference between most of the pixels is very low. We also note that pixels with a higher bias are mainly distributed in the edge of the features of the image, which suggests that the differences might be mainly due to the mismatch in the pixels' footprint between the two sensors. As Table 7 shows, the R^2 of the two sets of NDVI is high at 0.8754. The difference between the $NDVI\_N_{ASTER}$ and $NDVI\_N_{ETM+}$ is 0.0235, and the MRD decreased from 0.6294 before normalization to about 0.0375 after normalization, which means that the relative difference between the two is only 3.7% after normalization.

Furthermore, in order to allow us to closely examine the normalization effect, we mosaicked the normalized results, to fill the stripes of $NDVI\_N_{ETM+}$ with $NDVI\_N_{ASTER}$. In addition, the $NDVI_{DN}$ before normalization was also mosaicked with the same process for contrast. In order to closely examine the differences, a subarea of the mosaicked result is extracted and displayed. Figure 17(*e*) shows the $NDVI_{DN}$ of ETM+ before normalization, with its striped area set as the fill value. Figure 17(*f*) shows a subarea of the mosaicked result of the NDVI_N. Although some differences exist, the pattern of the

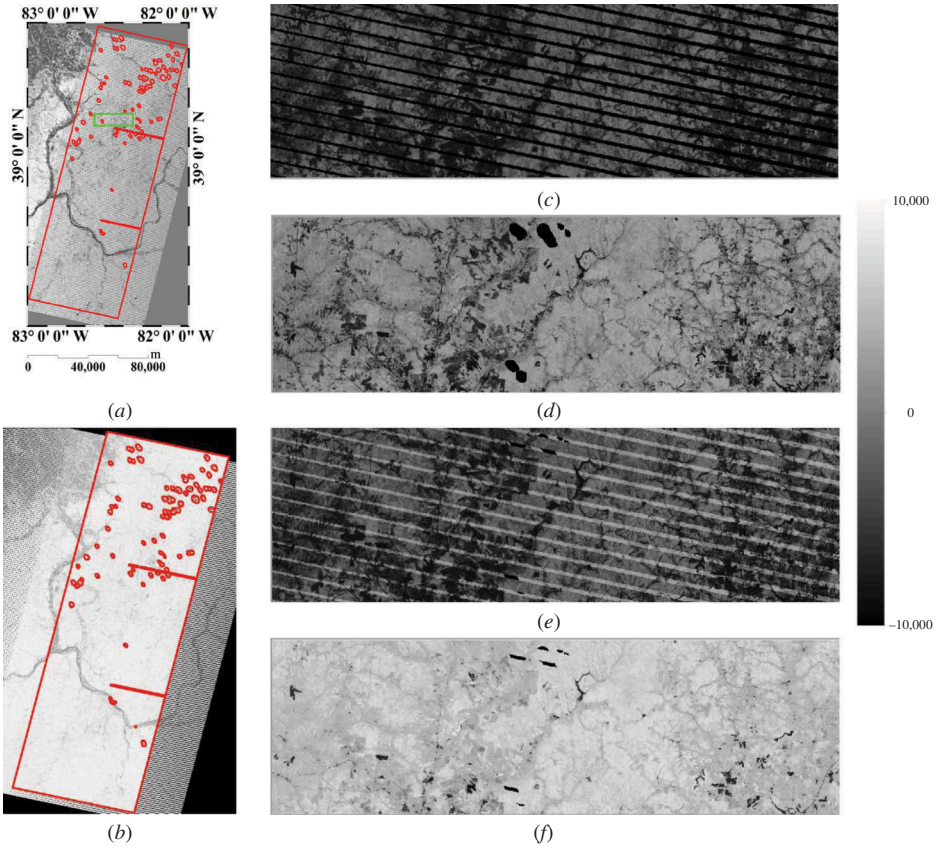


Figure 17. (*a*) The $NDVI_{DN}$ of ETM+ before normalization, in which red illustrates the extent of the ASTER NDVI, and the green box represents the subarea displayed in (*c*)–(*f*). (*b*) The normalized result of the ETM+ NDVI. (*c*) The subarea of the ETM+ $NDVI_{DN}$ before normalization. (*d*) The subarea of the ASTER $NDVI_{DN}$ before normalization. (*e*) The mosaicking result filling the gaps of the ETM+ $NDVI_{DN}$ with ASTER $NDVI_{DN}$. (*f*) The mosaicking result filling the gaps of the ETM+ NDVI after normalization by the use of the ASTER NDVI, in the area of Portsmouth, located on the border between Kentucky, Ohio and West Virginia, USA.

mosaicked NDVI_N is spatially seamless and continuous. In contrast, the mosaicked result before normalization is shown to be spatially discontinuous (Figure 17(*e*)). In addition, no stripes can be found on the mosaicked result, compared with the obvious difference in the mosaicked NDVI before normalization. This illustrates that the NDVI_N of ETM+ is almost the same as the NDVI_N of ASTER after normalization. This experiment can be considered as a successful example of the integration of multi-sensor NDVIs.

5. Conclusion

Medium-resolution NDVI has been widely used in both regional and local studies. However, being affected by many different factors, it is necessary to integrate data from different sensors to form a spatially seamless NDVI product and a consistent and continuous NDVI data set for the downstream analysis and application. However, the

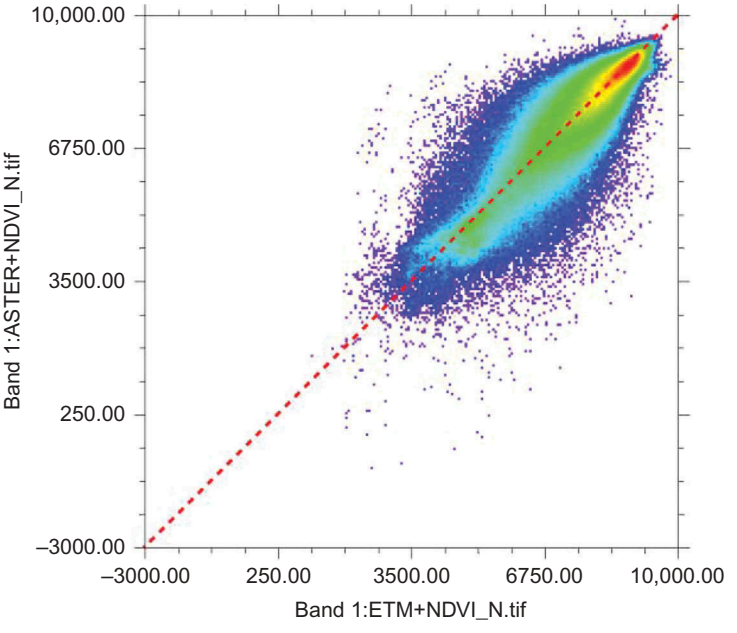


Figure 18. Density scatterplot of the normalized result of the ETM+ NDVI and that of ASTER, both produced by the use of the LCLM (*10,000).

Table 7. Statistics between the ASTER NDVI and the ETM+ NDVI, before and after normalization.

		R^2	MSE	MAD	MRD
Before normalization		0.8982	0.0385	0.1875	0.6294
After normalization	GloLM	0.8982	0.0020	0.0280	0.0412
	GCLM	0.8706	0.0020	0.0284	0.0396
	LCLM	0.8754	0.0016	0.0235	0.0375

NDVIs from different sensor systems vary and thus are difficult to integrate. It is therefore important to normalize the NDVIs from different sensors. In this paper, a reference-based method is chosen for this aim, which normalizes the medium-resolution NDVI from different sensors by the use of stable coarse-resolution data as the reference. The framework for NDVI normalization is described in this paper, and under this framework with consideration of the spatial heterogeneity of the influencing factors, a novel, local cluster-specific linear model, namely LCLM, is developed. This approach can normalize the NDVI from different sources and can produce an NDVI that is more consistent with the reference NDVI, but in its original resolution.

Differing from the previous studies (Jiang et al. 2006; Martínez-Beltrán et al. 2009; Steven et al. 2003; Thenkabail 2004; Van Leeuwen et al. 2006), the proposed method is more efficient for real applications. The method uses the relationship differences from image to image, instead of a sensor-identical cross-sensor transformation equation developed based on simulated data. By making use of the high-frequency, coarse-resolution

image to provide the synchronous reference data, it can build a relationship that is closer to the actual situation of the data.

Moreover, in this work, evaluation schemes appropriate for NDVI normalization by using reference data are presented, with a review of the existing methods. Three new schemes are designed, namely emOMRI, emSMRI and emSCU. These proposed schemes can measure the consistency between the normalized result and the reference data, and thus give an evaluation of the normalization method. emOMRI is a powerful tool for evaluating the effectiveness of normalization in eliminating the difference caused by the atmosphere. emSMRI and emSCU have the potential to evaluate the effectiveness of eliminating NDVI differences when the sensor characteristics differ. In addition, the emCMS method is confirmed to be efficient for assessing the consistency between the NDVI normalization results from different sensors. Multiples or all of these methods could be used together to give an exhaustive evaluation of a normalization method. The evaluation schemes systematically introduced in the paper could help to improve and perfect the framework of NDVI normalization methods, and will be valuable for related research.

The evaluation experiments following these proposed schemes indicate that the LCLM can produce more accurate results than the other two models, namely the GCLM and GloLM. The implementation experiment results for the normalized NDVI from Landsat ETM+ and Terra ASTER sensors are highly consistent with each other, illustrating that this reference-based method has good potential, and thus will be useful for the integration of medium-resolution NDVI from different sensors.

However, when applying the reference-based method with the LCLM to normalize the NDVI from different sources, several factors need to be considered. Firstly, this method is highly dependent on the consistency of the reference NDVI data set. The reference data need to be comparable, spatially or even temporally. In our study, MODIS products were chosen as the reference. If another data set is utilized to provide the reference data, its consistency should be first checked. Secondly, as the method using the LCLM is built by the use of local pixels, parameters such as the block size and threshold for homogeneous pixels should be set carefully to ensure that there are enough samples for the local model building. Thirdly, efficient pre-processing, including cloud masking and geo-rectification, is also important to guarantee the accuracy and to avoid any possible errors.

References

Chen, J. M. 1999. “Spatial Scaling of a Remotely Sensed Surface Parameter by Contexture.” *Remote Sensing of Environment* 69 (1): 30–42. doi:10.1016/S0034-4257(99)00006-1.

Fox, H., and S. Weisberg. 2010. *An R Companion to Applied Regression*. Thousand Oaks, CA: Sage.

Franke, J., V. Heinzel, and G. Menz. 2006. “Assessment of NDVI-Differences Caused by Sensor Specific Relative Spectral Response Functions.” Paper presented at the IEEE International Conference on Geoscience and Remote Sensing Symposium, Denver, CO, July 31–August 4.

Gallo, K. P., and C. S. T. Daughtry. 1987. “Differences in Vegetation Indices for Simulated Landsat-5 MSS and TM, NOAA-9 AVHRR, and SPOT-1 Sensor Systems.” *Remote Sensing of Environment* 23 (3): 439–452. doi:10.1016/0034-4257(87)90100-3.

Gallo, K. P., and J. C. Eidenshink. 1988. “Differences in Visible and Near-IR Responses, and Derived Vegetation Indices, for the NOAA-9 and NOAA-10 AVHRRS: A Case Study.” *Photogrammetric Engineering and Remote Sensing* 54: 485–490.

Gan, W., S. Huanfeng, G. Wei, and Z. Liangpei. 2013. “Normalization of NDVI from Different Sensor System using MODIS Products as Reference.” Paper presented at the 35th International Symposium on Remote Sensing of Environment, Beijing, April 21–26.

Gao, F., J. G. Masek, R. E. Wolfe, and C. Huang. 2010. "Building a Consistent Medium Resolution Satellite Data set using Moderate Resolution Imaging Spectroradiometer Products as Reference." *Journal of Applied Remote Sensing* 4 (1): 043526. doi:10.1117/1.3430002.

Guyot, G., and X.-F. Gu. 1994. "Effect of Radiometric Corrections on NDVI-Determined from SPOT-HRV and Landsat-TM Data." *Remote Sensing of Environment* 49 (3): 169–180. doi:10.1016/0034-4257(94)90012-4.

Holben, B. N. 1986. "Characteristics of Maximum-Value Composite Images from Temporal AVHRR Data." *International Journal of Remote Sensing* 7 (11): 1417–1434. doi:10.1080/01431168608948945.

Huber, P. J. 1964. "Robust Estimation of a Location Parameter." *The Annals of Mathematical Statistics* 35 (1): 73–101. doi:10.1214/aoms/1177703732.

Huete, A., H.-J. Kim, and T. Miura. 2005. "Scaling Dependencies and Uncertainties in Vegetation Index-Biophysical Retrievals in Heterogeneous Environments." Paper presented at the IEEE International Geoscience and Remote Sensing Symposium, Seoul, July 25–29.

Jiang, Z., A. R. Huete, J. Chen, Y. Chen, J. Li, G. Yan, and X. Zhang. 2006. "Analysis of NDVI and Scaled Difference Vegetation Index Retrievals of Vegetation Fraction." *Remote Sensing of Environment* 101 (3): 366–378. doi:10.1016/j.rse.2006.01.003.

Jonas, F., and G. Menz. 2004. "Sensor Intercalibration-Adjustment of MODISNDVI- to NOAA-AVHRR-NDVI Data." Paper presented at the IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, September 20–24.

Ju, J., D. P. Roy, E. Vermote, J. Masek, and V. Kovalskyy. 2012. "Continental-Scale Validation of MODIS-Based and LEDAPS Landsat ETM+ Atmospheric Correction Methods." *Remote Sensing of Environment* 122: 175–184. doi:10.1016/j.rse.2011.12.025.

Martínez-Beltrán, C., M. A. Osann Jochum, A. Calera, and J. Meliá. 2009. "Multisensor Comparison of NDVI for a Semi-Arid Environment in Spain." *International Journal of Remote Sensing* 30 (5): 1355–1384. doi:10.1080/01431160802509025.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and T.-K. Lim. 2006. "A Landsat Surface Reflectance Dataset for North America, 1990–2000." *IEEE Geoscience and Remote Sensing Letters* 3 (1): 68–72. doi:10.1109/LGRS.2005.857030.

Miura, T., A. Huete, and H. Yoshioka. 2006. "An Empirical Investigation of Cross-Sensor Relationships of NDVI and Red/Near-Infrared Reflectance using EO-1 Hyperion Data." *Remote Sensing of Environment* 100 (2): 223–236. doi:10.1016/j.rse.2005.10.010.

Obata, K., T. Miura, and H. Yoshioka. 2012. "Scaling Effects in Area-Averaged Values of Two-Band Spectral Vegetation Indices Represented in a General Form." *Journal of Applied Remote Sensing* 6 (1): 063585. doi:10.1117/1.JRS.6.063585.

Obata, K., T. Miura, H. Yoshioka, and A. R. Huete. 2013. "Derivation of a MODIS-Compatible Enhanced Vegetation Index from Visible Infrared Imaging Radiometer Suite Spectral Reflectances using Vegetation Isoline Equations." *Journal of Applied Remote Sensing* 7 (1): 073467. doi:10.1117/1.JRS.7.073467.

Obata, K., T. Wada, T. Miura, and H. Yoshioka. 2012. "Scaling Effect of Area-Averaged NDVI: Monotonicity along the Spatial Resolution." *Remote Sensing* 4 (12): 160–179. doi:10.3390/rs4010160.

Olthof, I., D. Pouliot, R. Fernandes, and R. Latifovic. 2005. "Landsat-7 ETM+ Radiometric Normalization Comparison for Northern Mapping Applications." *Remote Sensing of Environment* 95 (3): 388–398. doi:10.1016/j.rse.2004.06.024.

Price, J. C. 1994. "How Unique are Spectral Signatures?" *Remote Sensing of Environment* 49 (3): 181–186. doi:10.1016/0034-4257(94)90013-2.

Quattrochi, D. A., and M. F. Goodchild. 1997. *Scale in Remote Sensing and GIS*. Boca Raton, FL: CRC Press.

Steven, M. D., T. J. Malthus, F. Baret, H. Xu, and M. J. Chopping. 2003. "Intercalibration of Vegetation Indices from Different Sensor Systems." *Remote Sensing of Environment* 88 (4): 412–422. doi:10.1016/j.rse.2003.08.010.

Thenkabail, P. S. 2004. "Inter-Sensor Relationships between IKONOS and Landsat-7 ETM+ NDVI Data in Three Ecoregions of Africa." *International Journal of Remote Sensing* 25 (2): 389–408. doi:10.1080/0143116031000114842.

Trishchenko, A. P., J. Cihlar, and Z. Li. 2002. "Effects of Spectral Response Function on Surface Reflectance and NDVI Measured with Moderate Resolution Satellite Sensors." *Remote Sensing of Environment* 81 (1): 1–18. doi:10.1016/S0034-4257(01)00328-5.

Tucker, C. J., and P. J. Sellers. 1986. "Satellite Remote Sensing of Primary Production." *International Journal of Remote Sensing* 7 (11): 1395–1416. doi:10.1080/01431168608948944.
Van Leeuwen, W. J. D., B. J. Orr, S. E. Marsh, and S. M. Herrmann. 2006. "Multi-Sensor NDVI Data Continuity: Uncertainties and Implications for Vegetation Monitoring Applications." *Remote Sensing of Environment* 100 (1): 67–81. doi:10.1016/j.rse.2005.10.002.
Vermote, E. F., N. Z. El Saleous, and C. O. Justice. 2002. "Atmospheric Correction of MODIS Data in the Visible to Middle Infrared: First Results." *Remote Sensing of Environment* 83 (1–2): 97–111. doi:10.1016/S0034-4257(02)00089-5.
Vermote, E. F., D. Tanre, J. L. Deuze, M. Herman, and M. J.-J. 1997. "Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: An Overview." *IEEE Transactions on Geoscience and Remote Sensing* 35 (3): 675–686. doi:10.1109/36.581987.
Wu, H., B.-H. Tang, and Z.-L. Li. 2013. "Impact of Nonlinearity and Discontinuity on the Spatial Scaling Effects of the Leaf Area Index Retrieved from Remotely Sensed Data." *International Journal of Remote Sensing* 34 (9–10): 3503–3519. doi:10.1080/01431161.2012.716537.
Yoshioka, H., T. Miura, and A. R. Huete. 2003. "An Isoline-Based Translation Technique of Spectral Vegetation Index using EO-1 Hyperion Data." *IEEE Transactions on Geoscience and Remote Sensing* 41 (6): 1363–1372. doi:10.1109/TGRS.2003.813212.
Yoshioka, H., T. Wada, K. Obata, and T. Miura. 2008. "Monotonicity of Area Averaged NDVI as a Function of Spatial Resolution Based on a Variable Endmember Linear Mixture Model." Paper presented at the Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International.
Zhang, L., H. Shen, W. Gong, and H. Zhang. 2012. "Adjustable Model-Based Fusion Method for Multispectral and Panchromatic Images." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42 (6): 1693–1704. doi:10.1109/tsmcb.2012.2198810.
Zhang, T. 2011. "Cross Comparison of ASTER and Landsat ETM+ Multispectral Measurements for NDVI and SAVI Vegetation Indices." *Spectroscopy and Spectral Analysis* 31 (7): 1902–1907.
Zhang, Z. 1997. "Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting." *Image and Vision Computing* 15 (1): 59–76. doi:10.1016/S0262-8856(96)01112-2.