A Moving Weighted Harmonic Analysis Method for Reconstructing High-Quality SPOT VEGETATION NDVI Time-Series Data

Gang Yang, Huanfeng Shen, *Senior Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, Zongyi He, and Xinghua Li, *Graduate Student Member, IEEE*

***Abstract*—Global or regional environmental change is of wide concern. Extensive studies have indicated that long-term vegetation cover change is one of the most important factors reflecting environmental change, and normalized difference vegetation index (NDVI) time-series data sets have been widely used in vegetation dynamic change monitoring. However, the significant residual effects and noise levels impede the application of NDVI time-series data in environmental change research. This study develops a novel and robust filter method, i.e., the moving weighted harmonic analysis (MWHA) method, which incorporates a moving support domain to assign the weights for all the points, making the determination of the frequency number much easier. Additionally, a four-step process flow is designed to make the data approach the upper NDVI envelope, so that the actual change in the vegetation can be detected. A total of 487 test pixels selected from SPOT VEGETATION 10-day MVC NDVI time-series data from January 1999 to December 2001 were used to illustrate the effectiveness of the new method by comparing the MWHA results with the results of another four existing methods. Finally, the long-term SPOT VEGETATION 10-day maximum-value compositing (MVC) NDVI time series for China from April 1998 to May 2014 was reconstructed by the use of the proposed method, and a test region in China was utilized to validate the effectiveness of the proposed MWHA method. All the results indicate that the reconstructed high-quality NDVI time series fits the actual growth profile of the vegetation and is suitable for use in further remote sensing applications.**

***Index Terms*—Global Land Cover 2000 Project (GLC 2000), harmonic analysis of time series (HANTS), harmonic, local model, normalized difference vegetation index (NDVI), SPOT VEGETATION, time series.**

Manuscript received October 8, 2014; revised February 27, 2015 and April 20, 2015; accepted May 2, 2015. This work was supported in part by the National High Technology Research and Development Program of China under Grant 2013AA12A301 and in part by the National Natural Science Foundation of China under Grants 41271376 and 41422108.

G. Yang, Z. He, and X. Li are with the School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China (e-mail: love64080@163.com; 837381190@qq.com; lixinghua5540@whu.edu.cn).

H. Shen is with the School of Resource and Environmental Sciences and the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: shenhf@whu.edu.cn).

L. Zhang is with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing and the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: zlp62@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2431315

I. Introduction

DYNAMIC monitoring of vegetation provides important information reflecting the trend of climate change because vegetation has the obvious characteristics of both seasonal and interannual variations. For this reason, the monitoring of vegetation dynamics, as well as the analysis of the relationship between vegetation variation and climate change, has become an important field in global climate change research. Normalized difference vegetation index (NDVI) time-series data sets have been widely accepted as a promising approach for vegetation dynamic change monitoring, the macroscopic classification of vegetation, and the extraction of plant biophysical parameters [1]–[3]. The problem is that the application of NDVI time-series data sets derived from NOAA Advanced Very High Resolution Radiometers (AVHRR), Terra/Aqua Moderate Resolution Imaging Spectroradiometer (MODIS), or SPOT VEGETATION products is greatly limited by the unwanted residual effects and noise levels in the data [4], although these data sets are usually preprocessed, for example, by using the maximum-value compositing (MVC) technique [5] to eliminate the effect of cloud contamination and the residual atmosphere. Therefore, before an NDVI time-series data set can be used in a further application, an appropriate smoothing technique must be applied to reconstruct a high-quality NDVI time-series data set.

In an attempt to address the challenges faced by NDVI time-series data sets, a number of researchers have been exploring denoising methods to reconstruct high-quality NDVI time-series data over the last three decades. The developed denoising methods can generally be grouped into three types, each type having both advantages and drawbacks in preprocessing NDVI time-series data.

The first type is the time-domain local filter methods. For example, Viovy *et al.* [6] proposed the best index slope extraction method to reduce noise in an AVHRR NDVI time series. This method is based on the predictability of vegetation change, and the filtered results are often close to the original NDVI time series [7]–[9]. However, when using this method, a sliding period is required to determine the regrowth percentage of the NDVI to remove the outliers caused by clouds and other elements; it is also necessary to determine a smooth area and to indicate whether the increase in this area is within an acceptable range.

Therefore, the method makes the extracted temporal information unreliable when the NDVI time series presents a long-term declining trend [10]. Julien and Sobrino [11] used an iterative interpolation for data reconstruction (IDR) method to reconstruct Pathfinder AVHRR NDVI time-series data, which is an iterative process that compares the original NDVI time series to the alternative NDVI time series obtained as the mean between the immediate and following observations, and their method performed well in identifying the upper envelope of the original time-series data. However, the IDR method always overestimates the upper envelope when the transition between vegetation activity and dormancy is abrupt [11]. The mean-value iteration filter (MVI) method [12] uses a similar idea to reduce noise, and has been used to reconstruct Pathfinder AVHRR NDVI time-series data for the northwest of China. This method is carried out by comparing the difference to an average of different years, which is problematic for a short-length time series and for areas suffering high interannual change [11]. The Savitzky–Golay filter method, which is a local polynomial fitting method [13], [14], was modified to portray the patterns of SPOT VEGETATION NDVI time-series change by Chen (2004). However, the excessive influence of clouds and atmospheric conditions on the NDVI invalidates the smoothing effect [10]. In addition, Zhu *et al.* [15] proposed a changing-weight filter (CWF) method to reconstruct high-quality NDVI time-series data by the use of a 250-m 16-day MODIS NDVI product. The CWF method was designed to preserve the shape and amplitude of the NDVI time series, where a mathematical morphology algorithm and a rule-based decision are used to identify the local maximum/minimum, and a three-point changing-weight convolution filter is used to generate the new NDVI time series. However, its reconstruction results are not stable or reliable, with irregular fluctuations due to successive atmospherically contaminated values.

The second type of method is the denoising methods in the frequency domain. For example, Sellers *et al.* [16] developed the Fourier smoothing method, which is applied to an AVHRR NDVI time series by fitting the first three harmonics with a least squares solver. This approach has been widely used in many studies, such as calculating terrestrial biophysical parameters, evaluating net primary productivity dynamics, extracting the number and the timing of growing seasons, and extracting vegetation growth rates from AVHRR NDVI time-series data [17]–[20]. The NDVI curves obtained by the use of this method are quite smooth, with obvious dynamic characteristics of the vegetation growth. However, it is not suitable for irregular or asymmetric NDVI data since this method is critically dependent on symmetric sine and cosine functions. Similar to the Fourier series, the "harmonic analysis of time series" (HANTS) method was proposed by Verhoef [21] and has been successfully used in reconstructing cloud-free NDVI composites [22]. This method selects only the most significant frequencies in the time profiles and uses a least squares curve fitting procedure based on the harmonic components. However, it tends to overestimate the maximum NDVI values in the plateau of a time series and underestimates NDVI values when meeting several successive atmospherically contaminated values [11]. Lu *et al.* [23] proposed a wavelet transform method to generate high-quality terrestrial MODIS products. This is an effective method for reducing the noise in time-series data; however, it also reduces some reasonable high values, which limits its practical usage [23].

The third type of method is the function-based methods. These methods are based on least squares of the piecewise functions to the NDVI time series. The Gaussian function fitting approach performs well in obtaining a high-quality AVHRR NDVI time series [24], and the double logistic function fitting method [25] has been previously applied to Global Inventory Modeling and Mapping Studies (GIMSS) data to derive global land-surface phenology trends [26]. However, both of these methods have difficulty in fitting an NDVI time series with an irregular seasonal temporal profile, and their application may be limited when the region is large with diverse biomes [15].

A number of methods that consider certain particular factors have also been proposed. For example, Moody *et al.* [27] proposed an ecosystem-dependent temporal interpolation technique to fill missing or seasonally snow-covered surface albedos derived from MOD43B3 products. This technique is mainly based on the concept that pixels of the same ecosystem classification should exhibit roughly the same phenological or temporal behavior. In addition, Fang *et al.* [28] proposed the temporal spatial filter, which integrates both the spatial and temporal characteristics for different plant functional types to reconstruct leaf area index (LAI) products. Moreover, some other integrated methods can be found in [29]–[32].

The goals of this study are as follows: 1) to propose an innovative but effective denoising method (the "moving weighted harmonic analysis" [MWHA] method) to remove the potential anomalous effects of clouds and atmospheric contamination from an NDVI time series, while preserving the integrity of the vegetation phenology; and 2) to reconstruct a long-term high-quality SPOT VEGETATION 10-day MVC NDVI time series for China from April 1998 to May 2014.

II. Methodology

NDVI time-series data sets have been the most commonly used means for the temporal analysis of land-surface and climate change; however, their use is based on certain facts. First, vegetation change is the main characteristic of the NDVI; hence, an NDVI time series will show an annual cycle of increase, decrease, or stable tendency with the vegetation changes based on phenology. Plants growing under different climates show different growth tendencies and are seriously influenced by temperature and precipitation. For example, the NDVI time series of a cropland area will show different growth cycles in one year with different temperature and precipitation conditions. In an extremely cold area, an NDVI time series will show only one growth cycle in one year because crops grow slowly due to the low temperature. In a temperate zone, an NDVI time series of a cropland area will usually show three cycles over two years due to the evident seasonal variation in the temperature and precipitation. In subtropical and tropical zones, the high temperature and adequate moisture will result in an NDVI time series showing two or more growth cycles in one year. Different types of forests also show different growth profiles. For

example, deciduous forests show seasonal cycles as the season changes, which differs from evergreen forests, which have an active growth state all year round. Furthermore, arid areas suffer from low vegetation coverage, which results in an NDVI time series showing low values with no obvious increase or decrease. Second, certain kinds of contamination, such as cloud contamination, atmospheric variability, bidirectional effects, etc., can result in a decrease in NDVI values, making the NDVI time series inconsistent with the vegetation changes. It should be noted that these decreases differ from those caused by catastrophic natural events (e.g., drought, insect infestation, and fire). Because it takes a long period of time for vegetation to recover after a catastrophic natural event, the NDVI time series will show a gradual change profile from low values to normal high values over this time range. However, a decrease caused by contamination usually happens over a short time range, without a gradual change profile. In light of these two facts, the proposed method based on the harmonic analysis algorithm is designed to locally fit the growth or decline profile of the NDVI time series. Being clearly different from the HANTS algorithm, this strategy avoids the periodic influence that must be detected to reasonably confirm the frequency number. With the help of the support domain, NDVI time-series data are similarly grouped into many local parts, in which the data are fitted within a period of a seasonal cycle.

A. MWHA Method

The HANTS algorithm was developed based on the Fourier transform, where the series is the combination of a constant and superimposed sequence of sines and cosines based on the time interval, to deal with the irregularly spaced time series [33]. This algorithm is carried out iteratively and is a global model for the denoising of NDVI time-series data. However, the method tends to overestimate the maximum NDVI values in the plateau of a time series and underestimates NDVI values when meeting several successive atmospherically contaminated values.

In this paper, the proposed method is based on a modification of the HANTS algorithm from a global model to a local model. The local model [34] is to fit or interpolate for each t_i of $y(t)$ on a support domain $t \in I$. Let $y(t)$ be a time series, and $y(t) = \widetilde{y}(t) + \varepsilon(t)$ is its noisy counterpart. $\{y(t) = \widetilde{y}(t) + \varepsilon(t)\}$ describes the set of samples of the noisy time series. The parameterized model of y can then be described by the form

$$\widetilde{y}(t) = a_0 + \sum_{j=1}^{n} [a_j \cos(2\pi f_j t) + b_j \sin(2\pi f_j t)], \quad t = 1, \ldots, N \tag{1}$$

where $a = \{a_j\}_{j=0}^{n}$ and $b = \{b_j\}_{j=1}^{n}$ are the set of the parameters of the trigonometric components with the frequency f_j. N is the length of the time series, and n is the number of harmonics, of which the maxima must satisfy $2n + 1 \leq N$. The local model is to estimate $a = \{a_j\}_{j=0}^{n}$ and $b = \{b_j\}_{j=1}^{n}$ from the noisy time series y or its samples $\{y\}$ in the support domain of a point t.

Suppose that we want to denoise a noisy time series y or $\{y\}$. Then, around $t = t_0$, we can estimate the parameters of (1) by solving

$$\arg\min \left\| y(t_0) - \left[a_0 + \sum_{j=1}^{n} [a_j \cos(2\pi f_j t_0) + b_j \sin(2\pi f_j t_0)] \right] \right\| \tag{2}$$

in some support domain of t_0 and a distance norm $\|\cdot\|$. This minimization gives us the best local estimate of $\widetilde{y}(t_0)$ as follows:

$$\widetilde{y}(t_0) = a_0 + \sum_{j=1}^{n} [a_j \cos(2\pi f_j t_0) + b_j \sin(2\pi f_j t_0)]. \tag{3}$$

Repeat this process along the time series step by step for every location of t, and obtain the "moving" best estimate of y.

For each step, it is very important to set the weight for the neighborhood points according to the distance. Commonly, the weighted least squares distance [35] is chosen as a generalization for the local model as follows:

$$\left\| y(t) - \left[a_0 + \sum_{j=1}^{n} [a_j \cos(2\pi f_j t) + b_j \sin(2\pi f_j t)] \right] \right\|_w = \int \left[y(t_i) - \left[a_0 + \sum_{j=1}^{n} [a_j \cos(2\pi f_j t) + b_j \sin(2\pi f_j t)] \right] \right]^2 w(t_i - t) dy \tag{4}$$

where $w(\cdot)$ is a weight function. It localizes the estimation of the parameters a and b and should be nonnegative and monotonically decreasing, following the increasing of $\|t - t_i\|_2$. The weight value is not equal to 0 in the support domain; otherwise, it is 0. Spline functions are commonly used as the weight function, and in this work, a cubic spline is used to calculate the weight values. We denote $s' = |t - t_i|$ and $s = s'/r$, where r is the support domain radius. The weight value is described as follows:

$$w(s) = \begin{cases} \frac{2}{3} - 4s^2 + 4s^3, & s \leq \frac{1}{2} \\ \frac{4}{3} - 4s + 4s^2 - \frac{4}{3}s^3, & \frac{1}{2} < s \leq 1 \\ 0, & s > 1. \end{cases} \tag{5}$$

In this algorithm, as in the HANTS algorithm, input data points that are identified as outliers are removed by assigning a weight of 0 to them. Therefore, in the local model, these outliers are automatically detected and are not used to calculate the parameters. The outlier weight is

$$p(t) = \begin{cases} 0, & y(t) < \min \text{ or } y(t) > \max \\ 1, & \text{others} \end{cases} \tag{6}$$

where min and max are the minimum threshold and the maximum threshold, respectively, which can identify the outliers of the time series. Therefore, the final weight of (4) is described as follows:

$$w = w(s) \cdot p. \tag{7}$$

As mentioned in Section II, certain kinds of contamination can cause a sudden decrease in NDVI values, making the NDVI time series inconsistent with the vegetation changes. These

frequent sudden decreases make the reconstruction results deviate from the normal growth tendency of the vegetation. These sudden decreases should therefore be regarded as noise points, and these points are removed from the reconstruction process by setting their weights to 0. The problem is that, when the sudden decrease points are removed by assigning a weight of 0 to them, the number of available points within the support domain may not satisfy the condition $2n+1 \leq N$, where N is the number of available points within the support domain. For this reason, we execute a strategy by assigning a criterion to adaptively change the support domain radius. The criterion is described as $N_{w=0} \leq N-2n-d$, where $N_{w=0}$ is the number of points with a weight of 0, d is the degree of overdeterminedness, and the value is greater than or equal to 1. When the criterion is not satisfied, the support domain radius adds a constant value until the criterion is satisfied. This is a safety measure to satisfy the criterion for curve fitting. However, this criterion is always satisfied under normal circumstances and is not time consuming.

With the continuous and discrete cases mentioned earlier, we can compute the local parameters a and b to estimate the time series $\widetilde{y}$, not only at t_0, but at all t. Like the moving least squares estimation, we call this process the MWHA method.

There are two parameters influencing the smoothness of the reconstructed time series when this method is used for NDVI time-series smoothing or interpolation. The first parameter is the support domain radius r, which is used to calculate the weights and to identify the valid values in the support domain. Usually, a larger value of r makes the result smoother at the expense of flattening sharp peaks. The second parameter is the frequency number n, of which the maximum must meet the requirement that $2n+1$ is equal to or less than the number of valid points in the support domain. A smaller value of n makes a smoother result.

B. Four-Step Process Flow of the MWHA Method With an NDVI Time Series

When the MWHA method is implemented on an NDVI time series, we propose a number of strategies to identify the contaminated NDVI points and to ensure that the data approach the upper NDVI envelope. The four processing steps are presented as follows (see Fig. 1).

Step 1—Preprocessing of the Contaminated NDVI Time Series: Frequent contaminated points can make the reconstruction result deviate from the normal growth profile of the vegetation. Therefore, it is very important to identify these noise points. Most remote sensing data include a quality assurance (QA) band providing a valuable indication of the quality of each data point. Although these data do not indicate all the contaminated data points, they can be used to estimate low-quality data points for further analysis and application [10], [15], [23], [24]. In the proposed approach, the QA bands are used as ancillary data to indicate contaminated points, and these points are replaced by linearly interpolated values using their adjacent points. A linear interpolation strategy is chosen here since vegetation shows a nearly linear profile over a short period, such as one month for a three-point time span in a 10-day MVC NDVI time series [15]. Concretely, assuming that

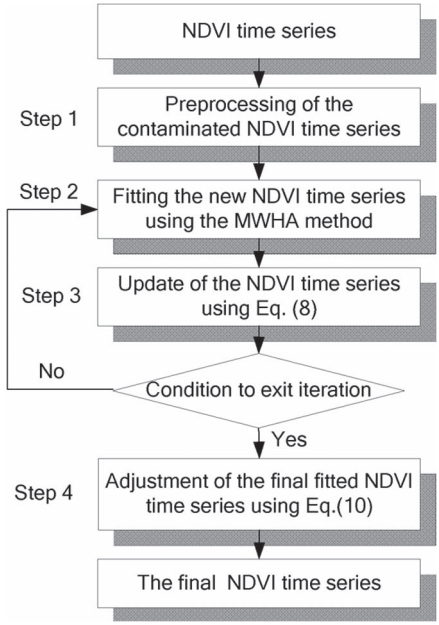


Fig. 1. Flowchart of the reconstruction process of the NDVI time series by the use of the proposed method.

there is an NDVI time series of data points $(t_i, N_i, Q_i), i = 1, 2, 3, \ldots, n$, where t_i is the time node, N_i is the NDVI value, and Q_i is the quality status flag, then the ith point N_i will be replaced by the linearly interpolated value using the high-quality adjacent points when Q_i is identified as a cloud or shadow point. In addition, some noisy points with a random NDVI increase of greater than 0.4 over the 20 days are also rejected and replaced by the linearly interpolated values using their adjacent points, since a normal vegetation change cannot show such a large increase [10]. After the contaminated NDVI points are interpolated, a new NDVI time series representing the NDVI seasonal change profile is obtained as $(t_i, N_i^0), i = 1, 2, 3, \ldots, n$, where t_i is the time node, and N_i^0 is the NDVI value after the linear interpolation. Fig. 3(a) shows the linearly interpolated NDVI time series (t_i, N_i^0) according to the quality status flags. Here, it is clear that three points showing a sudden decrease in the time series are identified by the quality status flags, while some other sudden decrease points are not identified.

It should be noted that, in the experiments, we executed both the proposed method and the other alternative methods based on the preprocessed NDVI time series.

Step 2—Fitting the New NDVI Time Series by the Use of the MWHA Method: After linearly interpolating the contaminated NDVI points, there may still be some sudden drops in the NDVI time series, meaning that the NDVI time series does not follow the gradual process of the annual vegetation cycle. In this case, it is very important to replace them with new values to fit the seasonal change profile. The new values should be above the corresponding original values, and the seasonal change profile curve should follow the variation profile of the annual vegetation cycle, without losing the temporal detail in the NDVI time series. For this reason, based on the new time series (t_i, N_i^0), the MWHA method is used to fit the variations between all the

NDVI points to obtain the new seasonal change profile. As mentioned earlier, it is very important to correctly select the combination of frequency number and support domain radius. The characteristics of the MWHA method mean that if the value of r (support domain radius) is too small, it can "overfit" the data points, and, in contrast, if the value of r is too large, it can result in a smooth curve and neglect some detail information in the NDVI time series. A sensitivity analysis of the MWHA method to the combinations of frequency number and support domain radius, which is based on a synthetically generated NDVI time series with low, moderate, and high levels of noise, was therefore undertaken, which determined the optimum combination as (1,5) (details in Section III-B and C). The new fitted NDVI time series using the MWHA method [as shown in Fig. 3(b)] is described as (t_i, N_i^{new}). It is clear that the new NDVI time-series curve preserves the temporal detail information, and the sudden decrease points can be seen to be below the new larger values.

Step 3—Update of the NDVI Time Series: In order to make the results approach the upper envelope of the original NDVI time series, a comparison should be made between time series (t_i, N_i^0) and the new time series (t_i, N_i^{new}). Then, we need to identify the points in the time series (t_i, N_i^k) that should be replaced by the new values. As such, a new time series (t_i, N_i^k) is generated using

$$N_i^k = \begin{cases} N_i^{k-1}, & N_i^{\text{new}} < N_i^{k-1} \\ N_i^{\text{new}}, & N_i^{\text{new}} > N_i^{k-1} \end{cases} \tag{8}$$

where $i = 1, 2, 3, \ldots, n$; $k = 1, 2, 3, \ldots, m$. Fig. 3(c) shows the first fitted curve (t_i, N_i^1) and the linearly interpolated time series (t_i, N_i^0). Here, it is clear that the first fitted curve is closer to the upper NDVI envelope.

It is apparent that it is a gradual process to approach the upper NDVI envelope, and an iterative process should be designed to make the fitted results approach the upper NDVI envelope as soon as possible. In each of the iterations, a time series difference (t_i, N_i^k) between the new time series (t_i, N_i^{new}) and the previous time series (t_i, N_i^{k-1}) should be calculated, as described in

$$D_i^k = |N_i^{\text{new}} - N_i^{k-1}|. \tag{9}$$

Normally, the time-series difference decreases after each iteration until the iterations stop when the maximum difference is less than a given threshold of 0.02 (details in Section III-B). Fig. 3(d) shows the final fitted curve, which is close to the upper NDVI envelope.

Step 4—Adjustment of the Final Fitted NDVI Time Series: After the iterative process, all the values of the original preprocessed NDVI time series (t_i, N_i^0) will be less than or equal to those of the final fitted NDVI time series $(t_i, N_i^{\text{final}})$. All the sudden decrease points, in particular, will be seen to fall below the final fitted time-series curve. In fact, it is not necessary for many of the original points along the correct vegetation change profile to be changed; on the contrary, it is better to keep the final fitted values close to the original values. In addition, the upper envelope may be overestimated when meeting an abrupt transition between vegetation activity and dormancy. In order

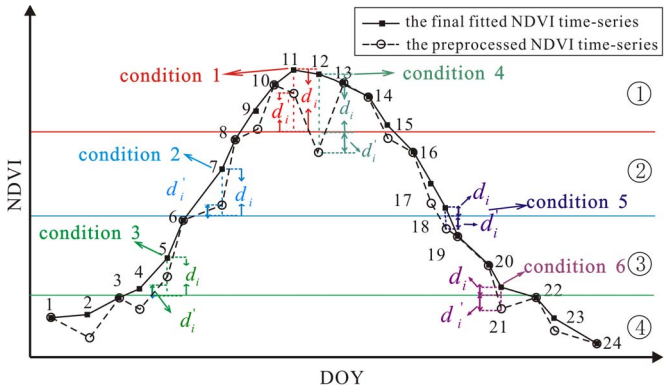


Fig. 2. Schematic of the six conditions in step 4.

to decrease the overestimation and increase the fidelity of the reconstructed time series, this step is designed to adjust the final fitted time series as follows:

$$N_i^{\text{adj}} = \begin{cases} \frac{d_i - d_i'}{d_i} \cdot N_i^{\text{final}} + \frac{d_i'}{d_i} \cdot N_i^0, & \text{conditions } 1, 2, \text{ and } 3 \\ \frac{\max(d_i, d_i')}{d_i + d_i'} \cdot N_i^{\text{final}} + \frac{\min(d_i, d_i')}{d_i + d_i'} \cdot N_i^1, & \text{conditions } 4, 5, \text{ and } 6 \\ N_i^{\text{final}}, & \text{others} \end{cases} \tag{10}$$

where N_i^{adj} is the adjusted value, N_i^{final} is the final fitted value of step 3, N_i^0 is the preprocessed value, N_i^1 is the first fitted value, d_i is the vertical distance of the final fitted point to the correct line according to the conditions, and d_i' is the vertical distance of the preprocessed point to the correct line according to the conditions. Specifically, as Fig. 2 shows, the blue line is the mean line of the preprocessed NDVI time series, the red line is the mean line of the points that are above the blue line in the preprocessed NDVI time series, and the green line is the mean line of the points that are below the blue line in the preprocessed NDVI time series. These three lines divide the plane into four parts, where condition 1 (see points 9–11, 13, and 14, with the red line as the correct line), condition 2 (see points 7, 8, 16, and 17, with the blue line as the correct line), and condition 3 (see points 5, 6, 19, and 20, with the green line as the correct line) refer to the situations where the final fitted NDVI points and the corresponding preprocessed points are all simultaneously located in part ①, part ②, and part ③, respectively. Condition 4 (see points 12 and 15, with the red line as the correct line) is the situation where the final fitted NDVI points and the corresponding preprocessed points are located in part ① and part ②, respectively. Similarly, condition 5 (see point 18, with the blue line as the correct line) means that the final fitted NDVI points and the corresponding preprocessed points are located in part ② and part ③. Condition 6 (see points 4 and 21, with the green line as the correct line) means that the final fitted NDVI points and the corresponding preprocessed points are located in part ③ and part ④, respectively. Other conditions refer to the situations where the final fitted NDVI points and the corresponding preprocessed points (see points 1–3 and 22–24) are simultaneously located in part ④. Fig. 3(e) shows the final adjusted NDVI time series (thick solid line), with the adjustments made by comparing the results with the final fitted NDVI

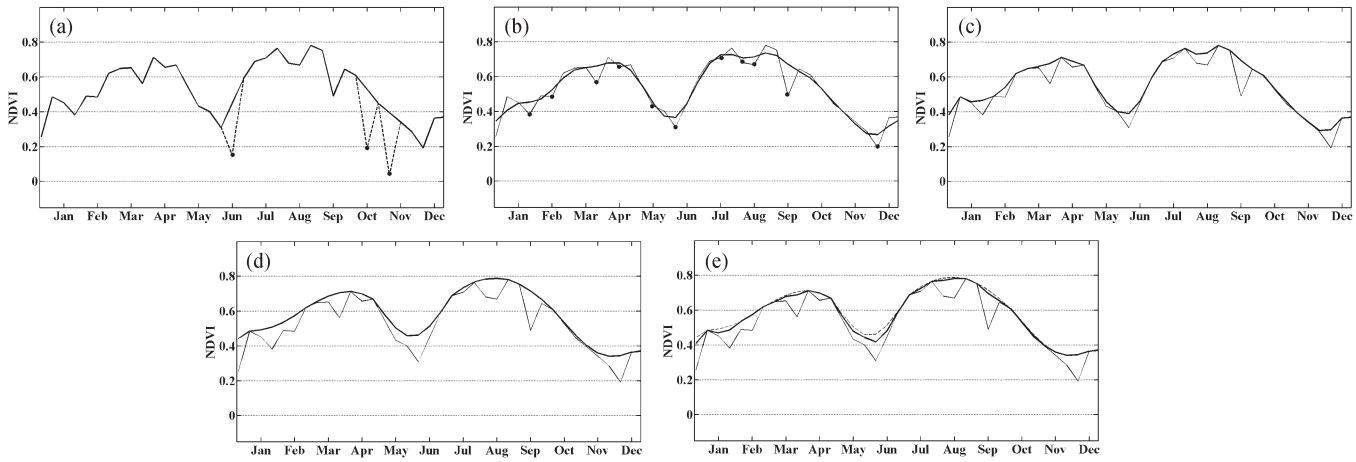


Fig. 3. Example of an NDVI time series in the different steps of the proposed method (NDVI data from the fourth test pixel in Fig. 5). (a) Original NDVI time series (short dashed line) with contaminated NDVI points (solid black points in the figure) and the linearly interpolated NDVI time series (t_i, N_i^0). (thick solid line). (b) Seasonal change profile curve (thick solid line) fitted by the use of the MWHA method. The solid black points are new identified noise points. (c) First fitted NDVI time series (t_i, N_i^1) by the use of (8), plotted as a thick solid line. (d) Final fitted NDVI time series (thick solid line). (e) Final adjusted NDVI time series (thick solid line) after all the processing steps and the final fitted NDVI time series in (d) (short dashed line).

time series (short dashed line). It is clear that the final adjusted NDVI time series approaches the upper envelope and preserves the temporal detail information of the seasonal change profile.

III. Results

A. Data

The SPOT VEGETATION data set used in this study consists of NDVI images acquired by the SPOT VEGETATION instruments (VGT1 and VGT2), which are equipped with wide-angle chip cameras, thereby avoiding the spatial distortion at off-nadir viewing angles. The data set ranges from April 1998 to May 2014, and the 16 years of data were acquired by the SPOT-4 and SPOT-5 satellites. The SPOT-4 satellite, with the VGT1 instrument on board, functioned from March 1998 to July 2013. The SPOT-5 satellite was launched on May 4, 2002, with the VGT2 instrument on board, and has been functioning correctly up to the end of May 2014. Every day, a new image of the global vegetation status is processed, archived, and distributed at the Image Processing Center of VITO in Belgium (available from the VEGETATION website at http://www.vgt.vito.be/index.html).

The SPOT VEGETATION NDVI subsets for Southeast Asia (68–147° E, 5–55° N) obtained from the VEGETATION website (http://www.vgt.vito.be/pages/policy.html) are 10-day MVC composites at a 1-km spatial resolution, which are generated from the atmospherically corrected P products. The P products are processed for band-to-band radiometric and geometric corrections [36]. Although the most commonly used NDVI data sets, such as the Pathfinder Land data set, the GIMMS NDVI data set, and the SPOT VEGETATION products [37]–[39], are usually composited by means of the so-called MVC algorithm, in order to suppress atmospheric effects, they still contain noise limiting the further applications. For example, the directional effects caused by the complex light reflection function are well known in the MVC composites and are still visible in the VEGETATION synthesis. Coastlines are also often heterogeneous due to the surrounding sea or cloud margin, for the reason that the MVC algorithm selects cloudy pixels in preference to sea pixels. Due to a number of defective detectors in the instruments, some oblique lines with high values are often visible in the shortwave infrared (SWIR) band, as well as in the VEGETATION synthesis. These processing results are due to the atmospheric influence and the viewing geometry of the sensor system at the time of imaging.

We obtained the SPOT VEGETATION NDVI subsets for China (73–135° E, 5–53° N) by cropping the subsets for Southeast Asia by the use of vector data. There are two processes required for these data. First, the adverse edge effect of the fitting quality must be pointed out, in that the first r points and the last r points do not have a complete support domain when the proposed method is carried out. Additionally, given the fact that the NDVI time series is cyclic, and in order to be consistent with the Global Land Cover 2000 Project (GLC 2000), we chose the SPOT VEGETATION NDVI data sets for the period of January 1999 to December 2001 to test the proposed method. In this experiment, 487 test pixels were selected from the SPOT VEGETATION NDVI time series by the use of the Create Random Points tool in ArcGIS, in which the GLC 2000 data were used for the stratification. Fig. 4 shows the GLC 2000 data and the distribution of the 487 test pixels. These test pixels include nearly all the vegetation types shown in the GLC 2000 data. The other process was that the NDVI values must be recalculated from the digital number (DN) value range to the original value range of -1.0–1.0, according to the formula as follows: $\text{NDVI} = 0.004 \times \text{DN} - 0.1$ [40]. In addition, the SPOT VEGETATION composite product includes a QA band, which is named the status map (SM) and restored as 8-bit data. It provides not only the radiometric quality for the four spectral bands but also the cloud, shadow, water, and ice/snow flags for each point of the product. Because the combination of Bit NR 0 and Bit NR 1 corresponds to the cloud flag, we selected Bit NR 0 and Bit NR 1 of the SM band to identify the contaminated points.

In this paper, the global land-cover database for the year 2000 (GLC 2000) was used for the stratification to select the test pixels from the SPOT VEGETATION NDVI time series. This database is coordinated and implemented by the Global Vegetation Monitoring Unit, in collaboration with a network of

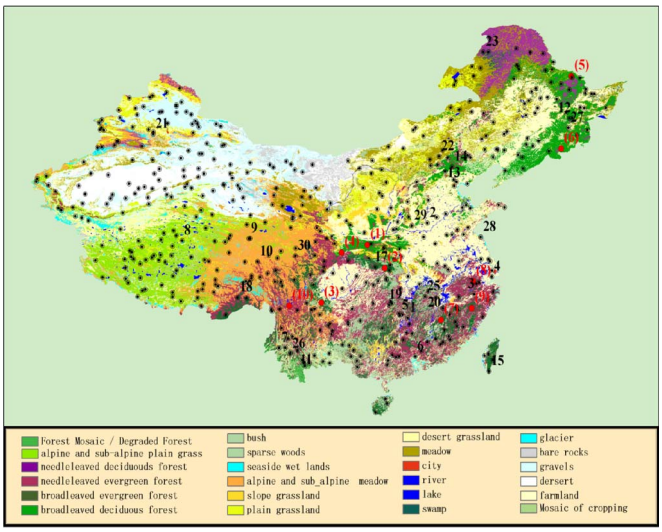


Fig. 4. GLC 2000 (mostly China) and the distribution of the 487 pixels.

partners around the world. The aim of the project is to update the existing global land-cover data for the year 2000 and to provide a basic database of global land cover for environmental assessment in relation to various activities. The land-cover maps are all based on daily data from the VEGETATION sensor aboard SPOT 4, while the mapping of some regions with specific issues is completed using data from other Earth-observing sensors [41]. The data of the Chinese region were provided by the Database of Global Change Parameters, Chinese Academy of Sciences (http://globalchange.nsdc.cn).

B. Application of Existing Methods and Parameter Determination

In this paper, four other denoising methods were selected to compare with the proposed method by plotting each of the 487 selected test pixels in the NDVI time series. These methods, namely, the Savitzky–Golay filter method, the HANTS algorithm, the IDR method, and the CWF method, are all efficient methods for denoising an NDVI time series.

For the analysis of the SPOT VEGETATION 10-day MVC NDVI time series, the parameters of all the filter methods must be correctly determined. For the Savitzky–Golay filter method, two parameters m (half-width of the smoothing window) and d (degree of polynomial) need to be determined. We adopted the best combination of (m, d) as being (4,6), which was determined by the method proposed by Chen *et al.* [10], using a SPOT VEGETATION NDVI time series for Southeast Asia. After the successful application of the IDR method in the reconstruction of a 15-day MVC AVHRR NDVI time series [11] and a daily NDVI time series [42], with the threshold parameters being set as 0.02 and 0.005, respectively, we selected the value of 0.015 as the threshold for the 10-day MVC SPOT VEGETATION NDVI time series after several experiments. For the CWF method, the window width of the structuring element is the main parameter, which should cover a full vegetation growth cycle [15]. In this paper, we set it as 11, i.e., equal to the support domain width of the MWHA method, which could cover a full vegetation growth cycle for the 487 test points. The other two thresholds (maximum/minimum point difference tolerance and fit error tolerance) were set as 0.1 and 0.01, as described in [15]. Although Roerink *et al.* [22] pointed out that one of the disadvantages of the HANTS method is that there are no objective rules for determining the parameter settings, we determined the parameter settings for HANTS by running several different combinations of the parameters. Finally, we set the frequency number as 15 for the three years and the fit error tolerance as 0.02 because they bring satisfactory results both qualitatively and quantitatively.

In the MWHA method, the support domain radius r and the frequency number n are the two main parameters to be determined, according to the facts mentioned at the beginning in Section II. The first fact is that vegetation has the obvious characteristics of both seasonal and interannual variations; hence, the NDVI time series should show an annual cycle of increase, decrease, or stable tendency with the vegetation changes. However, the MWHA method is carried out with a local model restricted by the support domain, and the points included in the support domain show no more than one period of a full vegetation growth profile. For the shortest seasonal period, one frequency is enough to describe the vegetation change profile. After determining the frequency number as 1, the support domain radius is then the main parameter influencing the reconstructed NDVI time series. According to a sensitivity analysis of the MWHA method to combinations of frequency number and support domain radius, the radius was determined as 5. Finally, when the maximal difference in the time-series difference in step 3 is less than a given threshold, the iteration is terminated. The given threshold was set as a lower threshold of 0.02, in order to denoise the contaminated NDVI time series according to [11].

C. Comparison Results

1) Qualitative Assessment: In order to assess the filtering effectiveness of the proposed method, certain criteria have to be established. Commonly, a qualitative assessment is the main method used in the comparison of alternative denoising methods, since we cannot obtain the ground truth for an NDVI time series. For the qualitative assessment, we undertook a visual inspection of the original, preprocessed, and filtered NDVI time series obtained through the five methods. Specifically, we plotted the NDVI time series of the selected 487 test points for the period from January 1999 to December 2001. The qualitative assessment for the different filtered NDVI time series could then be given, whether the noisy points were identified by inspection or not. These inspection results were used to calculate the ideal modeled NDVI data for the quantitative assessment. In order to display more detailed comparison results within a limited space, only the qualitative assessment for the data set of the year 2000 is given here.

Fig. 5 shows the fitting results of the NDVI time series carried out by the five methods. Due to the space limitations, we just show the results of 30 test pixels, which contain nearly all the vegetation types mentioned in Fig. 4, with the corresponding vegetation types shown in Table I. All the information obtained from the results can be explained from three aspects. First, the SPOT VEGETATION NDVI product still includes a lot of noise

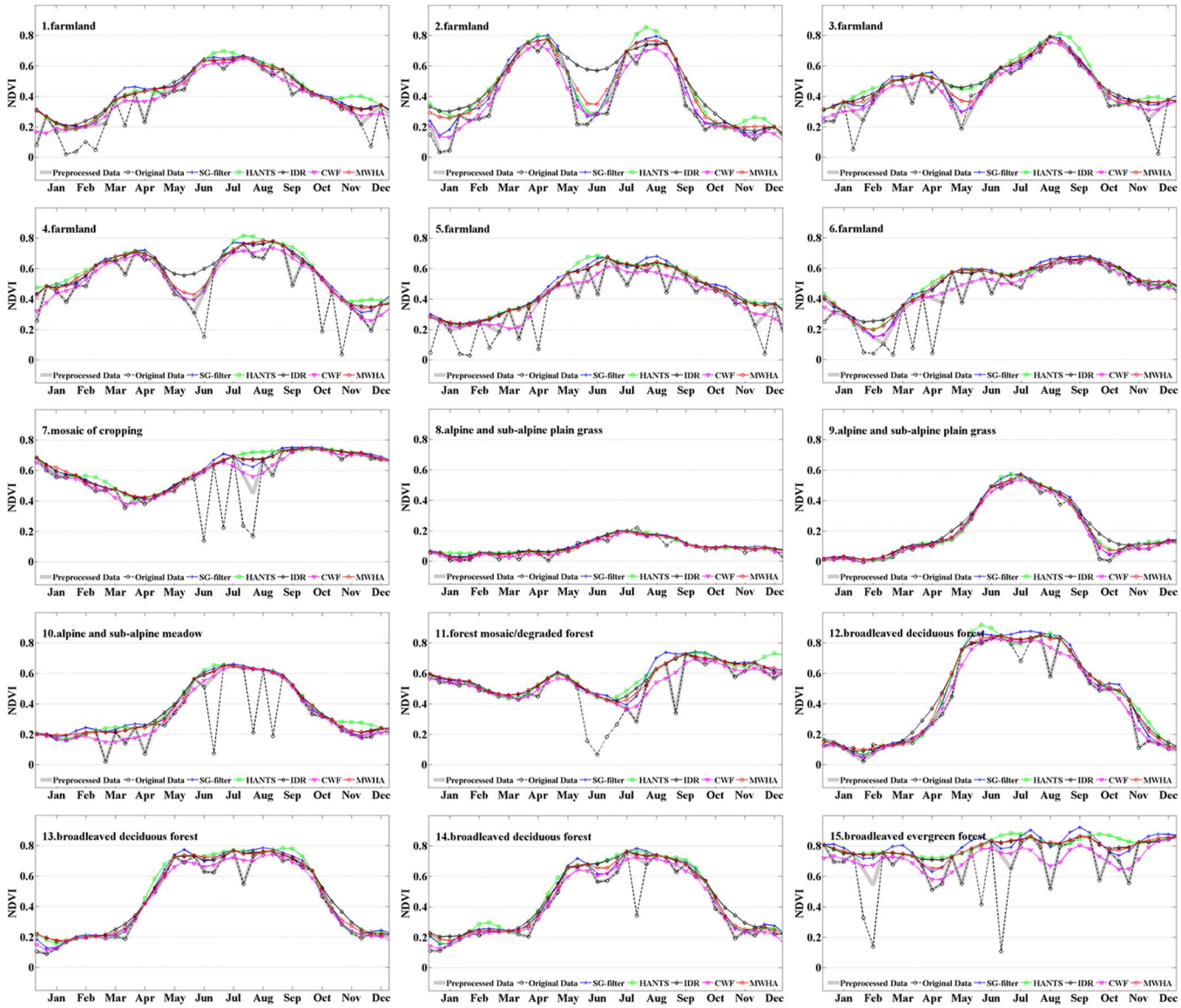


Fig. 5. Results of the reconstructed NDVI time series generated by the MWHA method, the IDR method, the HANTS algorithm, the Savitzky–Golay filter method, and the CWF method.

after the 10-day maximum compositing, as the original values of test pixels 6, 7, 10, 15–20, and 24–26 show. Therefore, the product cannot be directly used for analysis without preprocessing. Second, using the SM bands can identify some but not all the contaminated points in the NDVI time series. Finally, the general variation profile of the NDVI time series shows that these five methods can denoise correctly and can effectively reconstruct high-quality NDVI time-series data sets in most cases.

Of the five methods, the Savitzky–Golay filter method is designed to approach the upper envelope of the NDVI time series. This method could identify the continuous contaminated points effectively, including all the low values shown in each figure of Fig. 5. However, its limitation is that it could not fit a correct seasonal variation profile following actual conditions with many fluctuations. Furthermore, the reconstructed time series tended to overestimate NDVI values for the frequent fluctuations (see, for example, test pixels 15, 18, and 24). The HANTS algorithm tended to approach the upper envelope of the NDVI time series from below, and the reconstructed time series had roughly the same variation profile as the original time series. However, many of the values in the time series were changed and showed a large deviation away from the original time-series values, which meant that the fitting curve was the smoothest in the frequent fluctuation areas (see for example test pixels 15, 18, 24, and 29). The HANTS method made the time series underestimate the NDVI values for the dates surrounding the plateau areas (see, for example, test pixels 2, 4, 12, 13, and 23). The IDR approach obtained a better result than the aforementioned two methods. It not only identified the contaminated values correctly and fitted them to the upper envelope of the NDVI time series, which was not always the case with the aforementioned two methods (see, for example, test pixels 12, 15, 17, and 18), but also conserved the values of the pixels, which were not identified as contaminated points during the iterative reconstruction. However, it tended to overestimate the upper envelope with successive low-value points, such as test pixel 24. In addition, overestimation also happened in the dormancy period of the vegetation when the transition between the vegetation

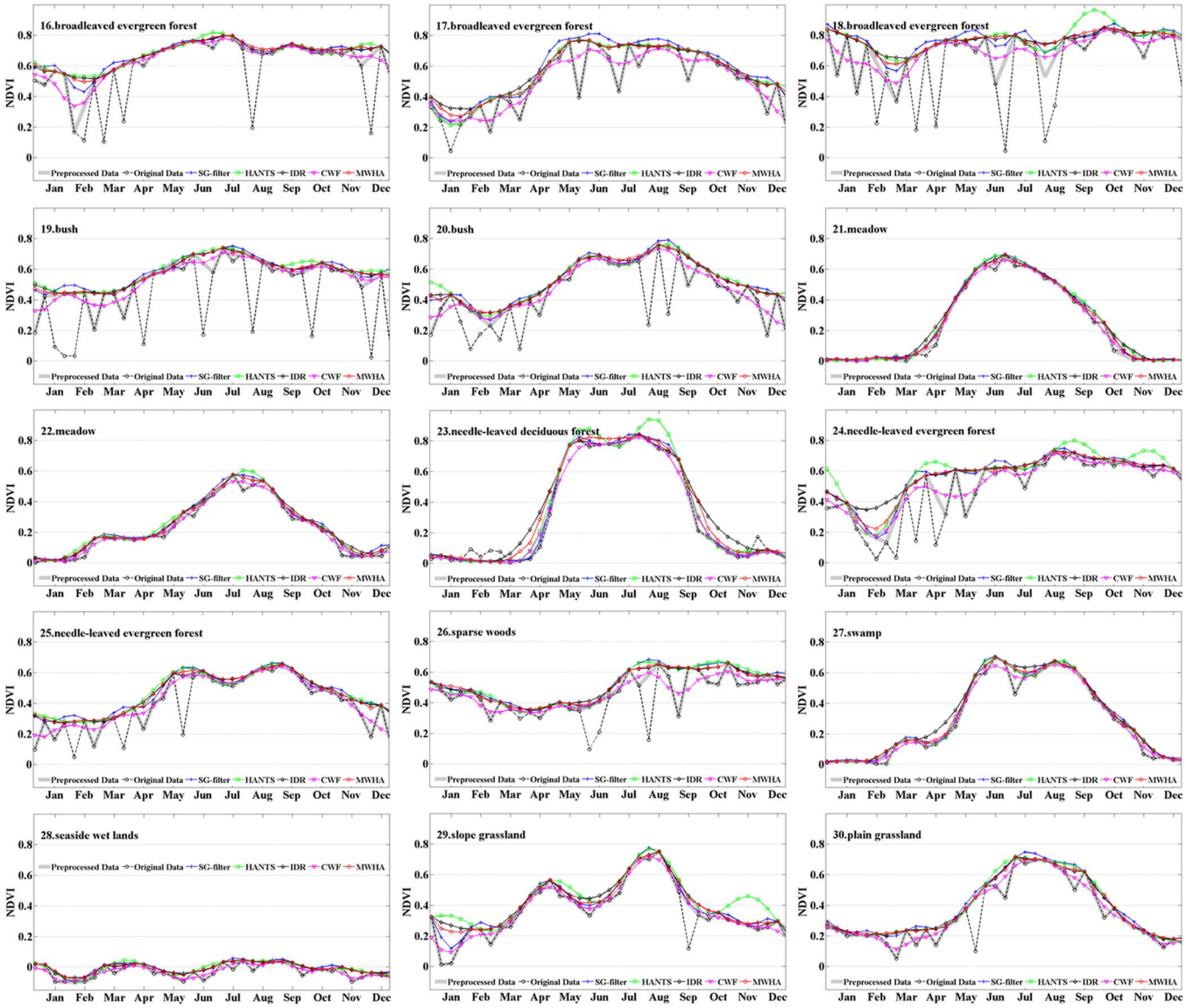


Fig. 5. *(Continued.)* Results of the reconstructed NDVI time series generated by the MWHA method, the IDR method, the HANTS algorithm, the Savitzky–Golay filter method, and the CWF method.

TABLE I
CORRESPONDENCE BETWEEN THE TEST PIXELS AND THE VEGETATION TYPE IN FIG. 4

Test pixels	Vegetation type	Test pixels	Vegetation type	Test pixels	Vegetation type
1	farmland	11	forest mosaic/degraded forest	21	meadow
2	farmland	12	broadleaved deciduous forest	22	meadow
3	farmland	13	broadleaved deciduous forest	23	needle-leaved deciduous forest
4	farmland	14	broadleaved deciduous forest	24	needle-leaved evergreen forest
5	farmland	15	broadleaved evergreen forest	25	needle-leaved evergreen forest
6	farmland	16	broadleaved evergreen forest	26	sparse woods
7	mosaic of cropping	17	broadleaved evergreen forest	27	swamp
8	alpine and sub-alpine plain grass	18	broadleaved evergreen forest	28	seaside wet lands
9	alpine and sub-alpine plain grass	19	bush	29	slope grassland
10	alpine and sub-alpine meadow	20	bush	30	plain grassland

growing season and the dormancy was abrupt (as in test pixels 12, 23, and 27). Differing from the aforementioned three methods, the CWF method aims to preserve the shape and amplitude of the NDVI time series rather than the upper envelope of the NDVI values. For most of the test points, the CWF method showed an excellent performance in denoising the contaminated NDVI time series, and the general temporal profile shape was preserved. However, when the NDVI time series was continually contaminated or showed an irregular curve with several fluctuations, this method could not fit a correct

seasonal variation profile following the actual conditions (see, for example, test pixels 15, 17, and 24).

Compared with the other four methods, the MWHA method showed a superior performance in denoising the irregular NDVI time series. The results in Fig. 5 show that the general profile of the reconstructed time series is approaching the actual time series. In addition, the sudden decrease values were identified, and the result tended to approach the upper envelope of the time series correctly with the help of the four-step process, which demonstrates that the MWHA method was clearly better than the CWF method because the CWF method could not make the results follow the upper envelope. Differing from the IDR method, the MWHA method could correctly estimate the pixel values in the period of vegetation dormancy, which can be seen in the experimental results of the broadleaved evergreen forest, such as test pixels 15, 16, 17, and 18. Although the results of the same vegetation type showed different performances, they were all reasonable. For example, test pixel 17 showed a seasonal change profile with high NDVI values in summer because it was selected from the midland of China. Furthermore, test pixels 16 and 18 were selected in the southeast and southwest of China, respectively, almost at the same latitude, and their vegetation showed the same change profile, which was likely influenced by the highland climate. Finally, for test pixel 15, which was selected in Taiwan province, the NDVI had high values all year round.

Overall, for pixels of a regular NDVI time series (see test pixels 1, 8, 9, 21, and 22), all the results of the five methods showed a similar performance, but the results with the HANTS method were the smoothest. For pixels with two growth cycles in one year (see test pixels 2, 3, 4, and 29), the IDR method tended to combine the two growth cycles into one growth cycle for all cases, while the HANTS method did the same thing for just one case. It should be noted that the selected parameter values of these two methods may not be the optimal for eliminating the overestimation of the reconstructed data for rapid transitions. The Savitzky–Golay filter method, the CWF method, and the MWHA method all preserved the curve shape with two growth cycles very well. For pixels with successive sudden drops (see test pixels 15, 17, 18, and 24), the Savitzky–Golay filter method tended to overestimate the values, and the CWF method is not suitable for an NDVI time series with successive sudden drops, which perhaps lead to the results containing noise that was not identified by the cloud flags. On the other hand, some of the noise might be wrongly identified as the minimum points by the mathematical morphology algorithm, thus making the filtered NDVI time series follow an irregular curve with several fluctuations. The aforementioned analysis of the experimental results indicates that the MWHA method is a better choice, because it successfully eliminated the disadvantages of the other four methods and preserved the integrity of the phenology in many cases.

2) Quantitative Assessment: In this paper, a quantitative assessment was used to assess the proposed method and the other four existing methods. For the quantitative assessment, we followed the method carried out in [15], which adopted the quantitative analysis framework described in [43]. This method can empirically evaluate the denoising methods under a variety of conditions. First, the ideal modeled NDVI time series was synthesized using the averages of the filtered NDVI time series obtained through the five methods for each of the 487 test points. We then introduced varying (low, moderate, and high) levels of noise into the ideal modeled NDVI time series by replacing random selections of 10%, 40%, and 70% of the points, as described in [15]. These pixel values were randomly reduced by 5%–50% with an interval of 5%. Finally, we used the root-mean-square error (RMSE) to evaluate the fidelity of the filtered results obtained through the five methods. Lower RMSE values indicate that a method is better able to keep the high fidelity of the filtered NDVI time series. Considering the adverse edge effect mentioned in Section III-A, we calculated the RMSE after removing the first and last five points.

TABLE II
STATISTICS OF THE RMSE VALUES FOR THE FIVE DENOISING METHODS

Level of noise	RMSE					
	MWHA	SG	IDR	CWF	HANTS	None
Low	0.0132	0.0153	0.0199	0.0247	0.0284	0.0365
Moderate	0.0229	0.0313	0.0254	0.0547	0.0312	0.0745
High	0.0372	0.0486	0.0388	0.0812	0.0480	0.0987

Note: None = no noise reduction.

Table II shows the RMSE values for the five denoising methods after filtering the modeled NDVI time series with three levels of introduced noise. These results are in accordance with the visual inspection results, in that all five methods could effectively reduce the noise. However, the MWHA method obtained the lowest RMSE values among all five methods for all the levels of noise, indicating that the proposed method was more effective than the other alternative methods. For the low level of noise, the Savitzky–Golay filter method showed the second-best results of all five methods, but it was no better than the IDR method for the moderate and high levels of noise. The CWF method was the most unstable method and showed the worst RMSE values for the NDVI time series with moderate and high levels of noise. The HANTS method showed average results for the moderate and high levels of noise and obtained the worst results for the low level of noise.

3) Sensitivity Analysis: Although the qualitative and quantitative assessments validated the better performance of the MWHA method in reconstructing a high-quality NDVI time series, the influence of the frequency number and the support domain radius should not be neglected. Therefore, a sensitivity analysis of the MWHA method to the combinations of frequency number and support domain radius was undertaken based on the ideal modeled NDVI time series of 487 test points with low, moderate, and high levels of noise. Fig. 6 shows the change in the RMSE values with different combinations of frequency number and support domain radius. For the low level of noise, it is clear that the lowest RMSE value was obtained for the combination of (1, 5). For the moderate level of noise, the lowest RMSE values appeared when the support domain radiuses were 5, 8, and 9 for 1, 2, and 3 frequency numbers, respectively. The three RMSE values were nearly at the same level, with the lowest RMSE value being obtained for the combination of (1, 5). For the high level of noise, the three lowest RMSE values appeared at the combinations of (1, 5), (2, 8), and (3, 12), respectively. Similarly, the three RMSE values were nearly at the same level, with the combination of (1, 5) being the lowest. The frequency number and the support domain radius were therefore set as 1 and 5 for the MWHA method to reconstruct a high-quality SPOT VEGETATION time series. In

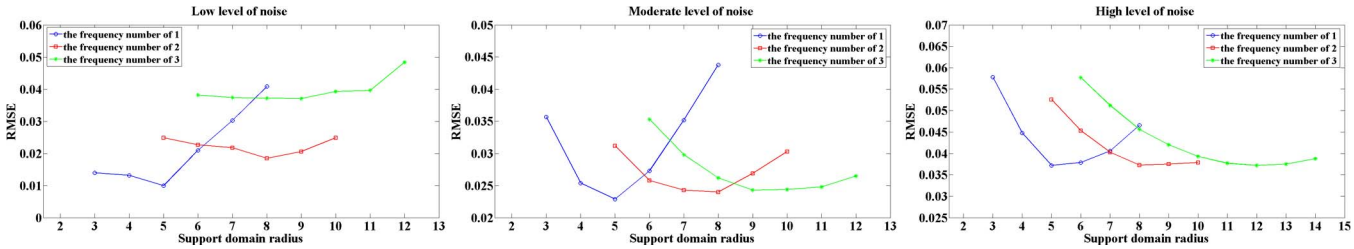


Fig. 6. Sensitivity of the MWHA method to the combinations of frequency number and support domain, based on the moderate and high levels of noise.

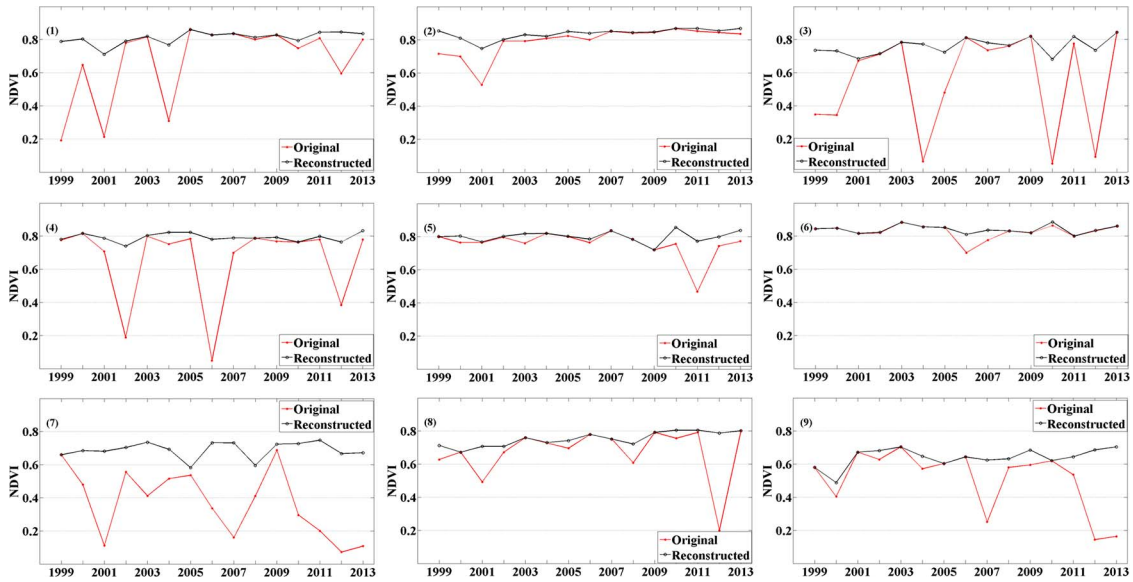


Fig. 7. Comparison of the interannual NDVI before and after reconstruction for each of the 10 test points.

addition, with the increase of the frequency number and the support domain radius, the process time also increased. Therefore, considering the process time, the most time-saving combination (1, 5) was chosen since the three lowest RMSE values were nearly at the same level. It was reasonable that the whole sliding window included 11 points of the NDVI time series (100 days) when the support domain was 5, which could just cover a full vegetation growth cycle. It should be noted that only three cases with the frequency number of 1, 2, and 3 were analyzed. This was because that, although there was a suitable support domain radius to match a larger frequency number, it would have been very time consuming for all the processes. For the 10-day SPOT VEGETATION NDVI product, there are three NDVI points in a month and 36 NDVI points in a year. According to the change in the seasons, three months as a local window width is reasonable. In other words, at least every 10 NDVI points (as a local part) are processed by the proposed method, and the support domain radius is reasonably set as 5.

D. Reconstruction of the High-Quality SPOT VEGETATION NDVI Time Series

After the validation, we were able to conclude that the new method is an effective filtering method. Therefore, we used it to reconstruct the SPOT VEGETATION 10-day MVC NDVI time series for China covering the period from April 1998 to May 2014. We also selected 9 points from the 487 test points to show the interannual change curves, to illustrate the effectiveness of the new method. Compared to other vegetation types, forest is

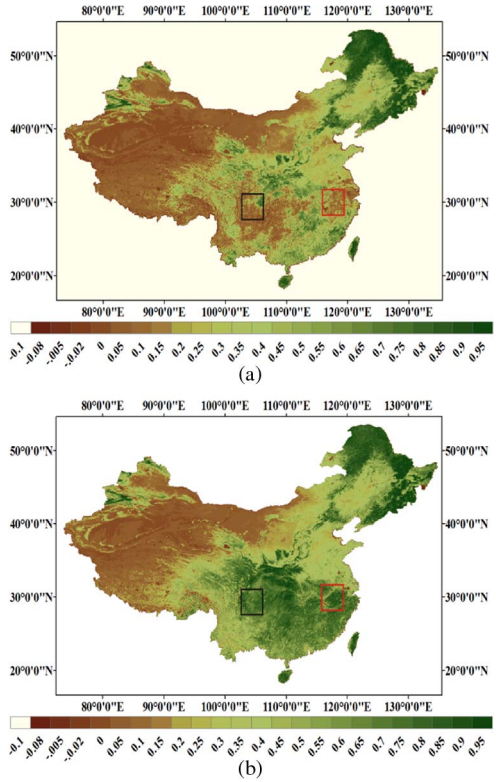


Fig. 8. NDVI values for China before and after the implementation of the MWHA method for June 21–30, 1999. The MWHA NDVI shows a striking improvement for southwestern China and southeastern China. (a) SPOT VEGETATION NDVI for June 21–30, 1999. (b) MWHA NDVI for June 21–30, 1999.

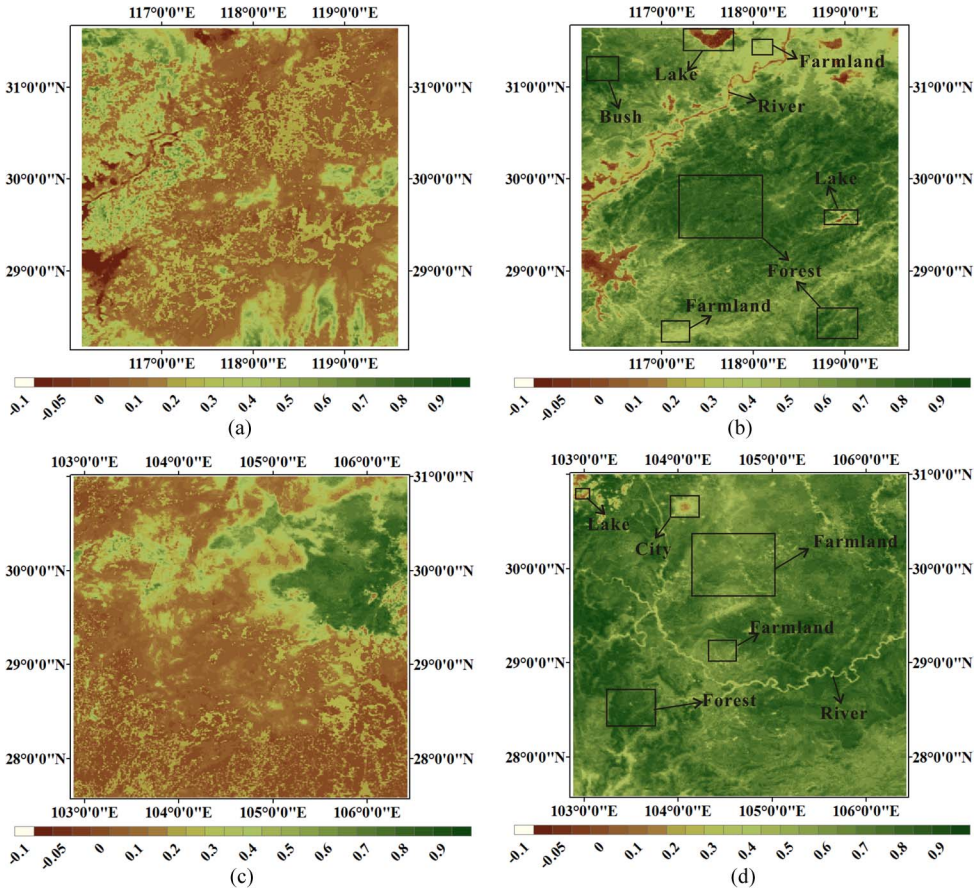


Fig. 9. Zoomed-in images of the black and red rectangles in Fig. 8 respectively show the details of the changes, in which some specific geographic features can be clearly distinguished in both areas. (a) Zoom in on the red rectangles in Fig. 8(a). (b) Zoom in on the red rectangles in Fig. 8(b). (c) Zoom in on the black rectangles in Fig. 8(a). (d) Zoom in on the black rectangles in Fig. 8(b).

more stable due to the diversity of biotic community. In addition, June, which is a month of the northern hemisphere summer, is the exuberant growth period for vegetation in China; hence the NDVI values will stay at a certain high level. As such, all the 10 points were selected with the forest vegetation type and are marked with solid red circles in Fig. 4. Fig. 7 shows a comparison of the interannual NDVI before and after reconstruction for each of the 10 test points generated using the NDVI of 15 years (1999–2013, because the data for 1998 and 2014 are not enough for one year) at the time of June 1–10, from which we can clearly see that the noise is effectively removed and the NDVI values stay almost at the same level.

To further test the effectiveness of the MWHA method at a regional scale, we show an example of the original SPOT VEGETATION NDVI values [see Fig. 8(a)] and the MWHA NDVI values for mostly China [see Fig. 8(b)] and two examples for local regions in southeastern China [see Fig. 9(a) and (b)] and southwestern China [see Fig. 9(c) and (d)] derived from the third 10-day composite period of June 1999. Fig. 8(a) shows that the image was seriously degraded by atmospheric perturbations due to the rainy season of southern China, and the values of the contaminated parts were very low. A direct comparison between Fig. 8(a) and (b) shows that these low values were effectively increased by the new method. Fig. 9 shows more detailed information of two local regions, as indicated by the red and black rectangles in Fig. 8(a) and (b), where some specific geographic features can be clearly distinguished after the MWHA reconstruction, such as the rivers and lakes shown in Fig. 9(b) and (d) and the city area shown in Fig. 9(d). The area covered by the black rectangle is located in the Sichuan Basin, China, with typical broadleaved evergreen forest, needle-leaved evergreen forest, meadow, and farmland, according to the GLC 2000 database. The city area is of Chengdu, for which the low pixel values were effectively retained. The area covered by the red rectangle is located in the Tai Lake Basin, with typical broadleaved evergreen forest, needle-leaved evergreen forest, bush, and farmland. Here, the specific geographic features were also effectively reconstructed. In summary, the reconstructed NDVI shows clearer geographic features than the original NDVI.

IV. Conclusion

Long-term NDVI time-series data can be used to monitor vegetation changes and support phenological interpretations, but further analysis and applications are limited because of the contamination caused by cloud presence and other atmospheric contamination. Therefore, a high-quality NDVI time series must be constructed before the NDVI time-series data can be used in further applications. To this end, several methods have been developed to reduce noise or to identify and interpolate contaminated values in NDVI time-series data. However, some of these methods are not sufficiently flexible or effective in filtering or interpolation. To deal with this issue, this paper has proposed a new method based on a modification of the HANTS

algorithm. A support domain is designed to assign the weights of all the points, and a four-step process flow for the NDVI time series is designed to make the data approach the upper NDVI envelope. We used the SPOT VEGETATION 10-day MVC NDVI product from January 1999 to December 2001 to illustrate the filtering effectiveness, by comparing the results with those from the Savitzky–Golay filter method, the HANTS algorithm, the IDR method, and the CWF method. From the comparison, we found that the proposed method has the following advantages over the existing methods: 1) it can identify the contaminated points and can make the data correctly approach the upper envelope of the time series; 2) it can correctly estimate the pixel values in a period of vegetation dormancy and is a robust strategy for most cases; and 3) it can be implemented with NDVI time-series data sets with different time intervals.

We have reconstructed high-quality SPOT VEGETATION 10-day MVC NDVI time-series data for China ranging from April 1998 to May 2014 by the use of the proposed method. We have also tested the performance of the MWHA method at a regional scale. From the results obtained, we can confirm that this high-quality SPOT VEGETATION NDVI time-series data could be used in global environmental change research and the monitoring of vegetation dynamics and for analyzing the relationship between environmental change and climate. We will explore how to better use such data sets combined with other data sets in our future research, such as the temporal analysis and change detection of land cover for multiple years, as Julien *et al.* [44], [45] have previously undertaken.

Although the four-step process flow was designed based on the MWHA method to reconstruct the high-quality SPOT VEGETATION 10-day MVC NDVI time series, it could be used not only for NDVI with different scales and time intervals but also for remote sensing time-series data sets from different sensors, such as all kinds of reflectance products, the LAI, and albedo, after a small adjustment. However, it should be noted that there are some limitations to the proposed method. In general, the information in different images of the same location does not change much as time goes by. For example, the surface temperature time-series product is not suitable for use with this new method because temperature changes irregularly, and the temporal relationship observed by the proposed method will not reflect the actual change profile. Additionally, a daily time-series data set is more vulnerable to successive cloud contamination than a composite time-series data set, and there may not be enough high-quality pixels in a single-pixel time series to satisfy this method. Therefore, when the proposed method is implemented in a daily time-series data set, other ancillary data should be used.

REFERENCES

[1] P. Sellers *et al.*, "Modeling the exchanges of energy, water, and carbon between continents and the atmosphere," *Science*, vol. 275, no. 5299, pp. 502–509, Jan. 1997.

[2] R. R. Nemani *et al.*, "Climate-driven increases in global terrestrial net primary production from 1982 to 1999," *Science*, vol. 300, no. 5625, pp. 1560–1563, Jun. 2003.

[3] M. Zhao and S. W. Running, "Drought-induced reduction in global terrestrial net primary production from 2000 through 2009," *Science*, vol. 329, no. 5994, pp. 940–943, Aug. 2010.

[4] C. Simolo, M. Brunetti, M. Maugeri, and T. Nanni, "Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach," *Int. J. Climatol.*, vol. 30, no. 10, pp. 1564–1576, Aug. 2010.

[5] B. N. Holben, "Characteristics of maximum-value composite images from temporal AVHRR data," *Int. J. Remote Sens.*, vol. 7, no. 11, pp. 1417–1434, Nov. 1986.

[6] N. Viovy, O. Arino, and A. Belward, "The best index slope extraction (BISE): A method for reducing noise in NDVI time-series," *Int. J. Remote Sens.*, vol. 13, no. 8, pp. 1585–1590, May 1992.

[7] J. Lovell and R. Graetz, "Filtering Pathfinder AVHRR land NDVI data for Australia," *Int. J. Remote Sens.*, vol. 22, no. 13, pp. 2649–2654, Jan. 2001.

[8] X. Xiao, S. Boles, J. Liu, D. Zhuang, and M. Liu, "Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data," *Remote Sens. Environ.*, vol. 82, no. 2/3, pp. 335–348, Oct. 2002.

[9] A. Ruimy, G. Dedieu, and B. Saugier, "TURC: A diagnostic model of continental gross primary productivity and net primary productivity," *Global Biogeochem. Cycles*, vol. 10, no. 2, pp. 269–285, Jun. 1996.

[10] J. Chen *et al.*, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter," *Remote Sens. Environ.*, vol. 91, no. 3/4, pp. 332–344, Jun. 2004.

[11] Y. Julien and J. A. Sobrino, "Comparison of cloud-reconstruction methods for time series of composite NDVI data," *Remote Sens. Environ.*, vol. 114, no. 3, pp. 618–625, Mar. 2010.

[12] M. Ma and F. Veroustraete, "Reconstructing Pathfinder AVHRR land NDVI time-series data for the Northwest of China," *Adv. Space Res.*, vol. 37, no. 4, pp. 835–840, 2006.

[13] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[14] J. Steinier, Y. Termonia, and J. Deltour, "Comments on smoothing and differentiation of data by simplified least square procedure," *Anal. Chem.*, vol. 44, no. 11, pp. 1906–1909, Sep. 1972.

[15] W. Zhu *et al.*, "A changing-weight filter method for reconstructing a high-quality NDVI time series to preserve the integrity of vegetation phenology," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1085–1094, Apr. 2012.

[16] P. Sellers *et al.*, "A global 1 by 1 NDVI data set for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI," *Int. J. Remote Sens.*, vol. 15, no. 17, pp. 3519–3545, Nov. 1994.

[17] L. Andres, W. A. Salas, and D. Skole, "Fourier analysis of multi-temporal AVHRR data applied to a land cover classification," *Int. J. Remote Sens.*, vol. 15, no. 5, pp. 1115–1121, Mar. 1994.

[18] C. M. Malmström *et al.*, "Interannual variation in global-scale net primary production: Testing model estimates," *Global Biogeochem. Cycles*, vol. 11, no. 3, pp. 367–392, Sep. 1997.

[19] M. Menenti, S. Azzali, W. Verhoef, and R. Van Swol, "Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images," *Adv. Space Res.*, vol. 13, no. 5, pp. 233–237, May 1993.

[20] A. Moody and D. M. Johnson, "Land-surface phenologies from AVHRR using the discrete Fourier transform," *Remote Sens. Environ.*, vol. 75, no. 3, pp. 305–323, Mar. 2001.

[21] W. Verhoef, "Application of harmonic analysis of NDVI time series (HANTS)," in *Fourier Analysis of Temporal NDVI in the Southern African and American Continents*, S. Azzali and M. Menenti, Eds. Wageningen, The Netherlands: DLO Winand Staring Centre, Report 108, 1996, pp. 19–24.

[22] G. J. Roerink, M. Menenti, and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *Int. J. Remote Sens.*, vol. 21, no. 9, pp. 1911–1917, Jan. 2000.

[23] X. Lu, R. Liu, J. Liu, and S. Liang, "Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 10, pp. 1129, Oct. 2007.

[24] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1824–1832, Aug. 2002.

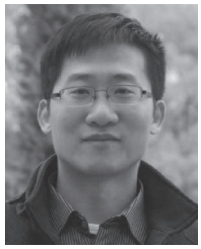
[25] P. S. Beck, C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI," *Remote Sens. Environ.*, vol. 100, no. 3, pp. 321–334, Feb. 2006.

[26] Y. Julien and J. A. Sobrino, "Global land surface phenology trends from GIMMS database," *Int. J. Remote Sens.*, vol. 30, no. 13, pp. 3495–3513, Jan. 2009.

[27] E. G. Moody, M. D. King, S. Platnick, C. B. Schaaf, and F. Gao, "Spatially complete global spectral surface albedos: Value-added datasets derived from Terra MODIS land products," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 144–158, Jan. 2005.
[28] H. Fang, S. Liang, J. R. Townshend, and R. E. Dickinson, "Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 75–93, Jan. 2008.
[29] A. Verger, F. Baret, M. Weiss, S. Kandasamy, and E. F. Vermote, "The CACAO method for smoothing, gap filling, and characterizing seasonal anomalies in satellite time series," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 1963–1972, Apr. 2013.
[30] Z. Jin and B. Xu, "A novel compound smoother—RMMEH to reconstruct MODIS NDVI time series," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 942–946, Jul. 2013.
[31] Z. Xiao, S. Liang, J. Wang, J. Song, and X. Wu, "A temporally integrated inversion method for estimating leaf area index from MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2536–2545, Aug. 2009.
[32] Z. Xiao, S. Liang, J. Wang, B. Jiang, and X. Li, "Real-time retrieval of leaf area index from MODIS time series data," *Remote Sens. Environ.*, vol. 115, no. 1, pp. 97–106, Jan. 2011.
[33] E. B. Brooks, V. A. Thomas, R. H. Wynne, and J. W. Coulston, "Fitting the multitemporal curve: A Fourier series approach to the missing data problem in remote sensing analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3340–3353, Sep. 2012.
[34] S. Shem-Tov, G. Rosman, G. Adiv, R. Kimmel, and A. M. Bruckstein, *On Globally Optimal Local Modeling: From Moving Least Squares to Over-parametrization*. Berlin, Germany: Springer-Verlag, 2013, ch. 17, sec. 2, pp. 379–405.
[35] G. Farnebäck, "Spatial domain methods for orientation and velocity estimation," Lic. thesis, Dept. Elect. Eng., Linköping Univ., Linköping, Sweden, 1999.
[36] E. Tarnavsky, S. Garrigues, and M. E. Brown, "Multiscale geostatistical analysis of AVHRR, SPOT-VGT, and MODIS global NDVI products," *Remote Sens. Environ.*, vol. 112, no. 2, pp. 535–549, Feb. 2008.
[37] J. Pinzon, "Using HHT to successfully uncouple seasonal and interannual components in remotely sensed data," in *Proc. SCI*, Orlando, FL, USA, vol. XIV, Jul. 14–18, 2002, pp. 287–292.
[38] J. Pinzon, M. E. Brown, and C. J. Tucker, "Satellite time series correction of orbital drift artifacts using empirical mode decomposition," in *Hilbert-Huang Transform: Introduction and Applications*. Singapore: World Scientific, 2005, pp. 167–186.
[39] C. Tucker *et al.*, "An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data," *Int. J. Remote Sens.*, vol. 26, no. 20, pp. 4485–4498, Oct. 2005.
[40] *SPOT Vegetation User Guide*, VEGETATION Programme, Mol, Belgium, 1998. [Online]. Available: http://www.vgt.vito.be/userguide/book_1/2/e2a.htm
[41] E. Bartholomé and A. Belward, "GLC2000: A new approach to global land cover mapping from Earth observation data," *Int. J. Remote Sens.*, vol. 26, no. 9, pp. 1959–1977, May 2005.
[42] J. S. Sobrino, Y. Julien, and G. Soria, "Phenology estimation from Meteosat second generation data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1653–1659, Jun. 2013.
[43] J. N. Hird and G. J. McDermid, "Noise reduction of NDVI time series: An empirical comparison of selected techniques," *Remote Sens. Environ.*, vol. 113, no. 1, pp. 248–258, Jan. 2009.
[44] Y. Julien and J. A. Sobrino, "The yearly land cover dynamics (YLCD) method: An analysis of global vegetation from NDVI and LST parameters," *Remote Sens. Environ.*, vol. 113, no. 2, pp. 329–334, Feb. 2008.
[45] Y. Julien *et al.*, "Temporal analysis of normalized difference vegetation index (NDVI) and land surface temperature (LST) parameters to detect changes in the Iberian land cover between 1981 and 2001," *Int. J. Remote Sens.*, vol. 32, no. 7, pp. 2057–2068, Mar. 2011.

Gang Yang received the M.S. degree in geographical information system from Hunan University of Science and Technology, Xiangtan, China, in 2012. He is currently working toward the Ph.D. degree with the School of Resource and Environmental Sciences, Wuhan University, Wuhan, China.

His current research interests focus on image inpainting and temporal reconstruction of remote sensing time-series products.



Huanfeng Shen (M'10–SM'13) received the B.S. degree in surveying and mapping engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

In July 2007, he joined the School of Resource and Environmental Sciences, Wuhan University, where he is currently a Full Professor. His research interests include image quality improvement, remote sensing mapping and application, data fusion and assimilation, and regional and global environmental change. He has published more than 100 research papers.

Dr. Shen serves as a Director of the China Association of Remote Sensing Application. He is currently a member of the Editorial Board of the *Journal of Applied Remote Sensing*. He has been supported by several talent programs, such as the China National Science Fund for Excellent Young Scholars (2014), the New Century Excellent Talents by the Ministry of Education of China (2011), and the Hubei Science Fund for Distinguished Young Scholars (2011).

Liangpei Zhang (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He is also a Chang-Jiang Scholar Chair Professor appointed by the Ministry of Education of China. He is currently a Principal Scientist for the China State Key Basic Research Project (20112016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has more than 410 research papers. He is the holder of 15 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a Fellow of the Institution of Engineering and Technology, an Executive Member (Board of Governor) of the China National Committee of the International Geosphere-Biosphere Program, and an Executive Member of the China Society of Image and Graphics. He was a recipient of the 2010 the Best Paper Boeing Award and the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing (ASPRS). He regularly serves as a Cochair of the series SPIE Conferences on Multispectral Image Processing and Pattern Recognition, the Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and geoinformatics symposiums. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the *International Journal of Digital Multimedia Broadcasting*, the *Journal of Geospatial Information Science*, the *Journal of Remote Sensing*, and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Zongyi He received the B.S. degree and the M.S. degree in cartography from Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1982 and 1984, respectively.

He is currently a Full Professor with the School of Resource and Environmental Sciences, Wuhan University. His research interests include digital mapping theory and technology, multiscale spatial data processing, the principle and method of spatial analysis model, visualization of spatial information, and cartography theory.

Xinghua Li (S'14) received the B.S. degree in geographical information system from Wuhan University, Wuhan, China, in 2011. He is currently working toward the Ph.D. degree with the School of Resource and Environmental Sciences, Wuhan University.

His current research interests focus on missing information reconstruction of remote sensing image, cloud removal of remote sensing image, compressed-sensing-based image processing, and sparse representation.