One-Step High-Quality NDVI Time-Series Reconstruction by Joint Modeling of Gradual Vegetation Change and Negatively Biased Atmospheric Contamination

Xinxin Liu, *Member, IEEE*, Huanfeng Shen, *Senior Member, IEEE*, Qiangqiang Yuan, *Member, IEEE*, Xiliang Lu, and Shutao Li, *Fellow, IEEE*

***Abstract*— The normalized difference vegetation index (NDVI) can reflect the plant life cycle of growth and senescence and has become a widely used tool for many applications related to phenology, ecology, and environment. However, unwanted disturbance from cloud, snow, and other atmospheric effects greatly lowers the NDVI quality and hinders its further application. In this article, differing from the previous research attempting to approach the upper NDVI envelope by local adjustment or threshold-related iteration, a novel one-step global variational reconstruction (OGVR) method for NDVI time series is proposed via joint modeling of the gradual vegetation change and negatively biased atmospheric contamination. Two versions of the proposed method are designed for processing NDVI data with or without auxiliary flag information. Long-term and global-scale Advanced Very High Resolution Radiometer (AVHRR) global inventory monitoring and modeling system (GIMMS) data were applied in simulated and real-data experiments to verify the proposed method. The results show that the proposed method can successfully estimate the natural vegetation change from seriously contaminated NDVI time series and can conquer the problem of continuous low-value gaps. The qualitative and quantitative comparisons with five other widely used methods indicate that the proposed method has significant advantages in terms of both effectiveness and stability.**

Manuscript received July 5, 2021; revised September 14, 2021 and October 25, 2021; accepted October 25, 2021. Date of publication November 2, 2021; date of current version February 14, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61901166, Grant 11871385, and Grant 41922008; in part by the Fundamental Research Funds for the Central Universities under Grant 531118010209; in part by the Hunan Provincial Natural Science Foundation of China under Grant 2020GK2038; and in part by the National Key Research and Development Program of China under Grant 2020YFA0714200. *(Corresponding author: Xinxin Liu.)*

Xinxin Liu and Shutao Li are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China, and also with the Key Laboratory of Visual Perception and Artificial Intelligence of Hunan Province, Changsha 410082, China (e-mail: liuxinxin@hnu.edu.cn; shutao_li@hnu.edu.cn).

Huanfeng 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).

Qiangqiang Yuan is with the School of Geodesy and Geomatics and the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: qqyuan@sgg.whu.edu.cn).

Xiliang Lu is with the School of Mathematics and Statistics and the Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan 430072, China (e-mail: xllv.math@whu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2021.3124798

***Index Terms*— Advanced Very High Resolution Radiometer (AVHRR) global inventory monitoring and modeling system (GIMMS), atmospheric contamination, normalized difference vegetation index (NDVI) time series, one-step reconstruction, variational optimization.**

I. Introduction

REMOTELY sensed normalized difference vegetation index (NDVI) characterizes vegetation coverage conditions and vegetation growth status. The large-scale and long-term use of the NDVI plays an essential role in many applications, including monitoring of vegetation phenological dynamics [1]–[3], detection of land-cover change [4]–[6], analysis of regional drought characteristics [7], [8], and investigation of climate response to vegetation activities [9]–[11]. Accordingly, access to high-quality NDVI time series is of great importance. However, the accuracy of NDVI time series is easily affected by the unwanted disturbance stemming from cloud, snow, dust, ozone, and some other factors. This disturbance shows up as negatively biased noise, which can cause sudden drops in NDVI values. To offset the bias, most of the remote sensing vegetation products are preprocessed via the maximum-value composite (MVC) technique [12]. Examples of such products are the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) global vegetation index data, Aqua/Terra Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation index products, and SPOT VEGETATION products. Although a considerable amount of noise has been eliminated in these products, the residual noise related to continuous cloudy days, snow coverage, and bidirectional effects still need to be handled. Hence, it is of critical importance to develop an effective reconstruction method to obtain high-quality NDVI time series for reliable analysis in further applications.

Over the past few decades, a number of noise removal methods have been proposed for NDVI data. Under the similar assumption that the noise in NDVI data is mainly negatively biased, the existing methods are generally characterized by the use of an iterative operation to approach the upper NDVI

envelope and can be broadly divided into three groups. The first group is the function-based fitting methods, where mathematical functions are utilized to shape the seasonal cycle of vegetation growth. The double logistic (DL) function [13], asymmetric Gaussian (AG) function [14], cubic spline function [15], and Fourier series [16]–[18] are the commonly used fitting techniques for NDVI time series. These methods perform well when the NDVI time series conforms to the law of natural change with regular seasonal cycles. However, low-quality results are sometimes unavoidable for those irregular NDVI profiles caused by abnormal climate, natural disasters, or human activity. For instance, the Fourier-based methods provide periodic sine and cosine functions to fit the cycles of vegetation growth, but difficulties remain in capturing the nonperiodic vegetation response to the abrupt phenomena in certain years. Adaptations combining local adjustments [19] alleviate the problem by using more parameters. Unfortunately, the complexity of the parameter settings increases at the same time, especially when processing heterogeneous land surface vegetation cover.

The statistical-based filtering/smoothing approaches constitute, to some extent, the most popular group. The main idea behind these approaches is to fill the missing or seriously contaminated pixels through valid observations in a local moving window. Statistical calculation, usually combined with a threshold, assists in different tasks to control the pixel quality judgment, the filtering degree, or the step approaching the upper envelope. Typical examples include the mean value iteration (MVI) filter [20], the Savitzky–Golay (SG) filter [21], [22], the Whittaker smoother (WS) [1], [23], [24], iterative interpolation for data reconstruction (IDR) [25], the running median, mean value, maximum operation, endpoint processing, and hanning (RMMEH) smoothing method [26], and the best index slope extraction (BISE) method [27]. However, due to the use of a local moving window, it is easy to ignore the global property, such as seasonal cycles, of NDVI time series in the filtering process, which can lead to some underestimation or overestimation in local parts. Taking low NDVI values as an example, when using only local judgment and processing, it is impossible to distinguish whether a low value is from a nongrowing or growing season, and the correct low values in nongrowing seasons are very likely to be overestimated after the same degree of smoothing. As the thresholds in these methods are always important (or even sensitive) but not self-adaptive, it needs lots of experience to set different thresholds properly for different types of vegetation cover.

Researchers have also tried to integrate some other techniques into NDVI reconstruction [28]–[32]. For example, Gu *et al.* [29] made use of a data assimilation method to calculate NDVI time series by weighting the background field (capturing the annual features of vegetation change) and the observed NDVI data, Qiu *et al.* [30] drew support from continuous wavelet transform to smooth MODIS enhanced vegetation index (EVI) time series, without any local adjustments, and Cao *et al.* [31] designed a spatiotemporal SG filter to fill the missing data under the assumption that neighboring pixels within the same land-cover type exhibit similar phenological behavior curves. However, drawbacks, such as the high computational load, complex selection from redundant wavelet coefficients, or dependency on the historical data, leave room for further improvement.

Despite the considerable effort that has been made, the universality of the existing methods is still limited, mainly for the following reasons.

A. Strict Requirements for Input Data

Ancillary data or historical data are important in NDVI reconstruction and are strictly embedded into the processing flows of many of the existing methods. However, for the case of insufficient or unavailable ancillary/historical data, alternative strategies are rarely given.

B. Unstable Performance or Low Efficiency

The land surface vegetation index changes spatially by the land-cover type and temporally by the season. Different abrupt environmental factors diversify the shapes of NDVI time series, which increases the difficulty of a stable filtering performance. Strategies, such as local adjustment or spatial information utilization, are commonly used improvements in most of the previous methods, but they always sacrifice the processing efficiency.

C. Complex Setting of Numerous Parameters

Parameter setting is generally closely related to the noise level of each NDVI time series. Since pixels from different land-cover types suffer varying degrees of temporal noise levels, it is difficult to choose universal and suitable values for the existing methods with sensitive parameters, and it is far harder in the case of numerous sensitive parameters.

In this article, we propose a novel one-step global variational NDVI reconstruction model by taking the properties of both natural vegetation change and negatively biased atmospheric contamination into consideration. Since the different information is well utilized in the proposed optimization-based model, NDVI reconstruction can be directly completed with no more preprocessing or postprocessing steps. This, on the one hand, accelerates the processing speed and, on the other hand, avoids the adverse effects such as accumulated errors. To increase the universality of the proposed model, a version that uses ancillary data and a version that does not use ancillary data are provided to adapt to different processing cases involving various NDVI products and self-calculated NDVI/EVI time series. There are as few parameters as possible in the model, which can be set as fixed values for ease of use in real applications. Experiments conducted on both simulated and real NDVI time series at a global scale demonstrate the effectiveness and robustness of the proposed model. Even without ancillary data, the performance of the proposed model is competitive when compared with the existing well-known NDVI reconstruction approaches.

The rest of this article is organized as follows. Section II introduces the proposed one-step global variational NDVI reconstruction model and its numerical solution. The experimental results and the comparisons with other methods are

provided in Section III. A further discussion and our conclusions are, respectively, provided in Sections IV and V.

II. Methodology

NDVI time-series data reflect the vegetation change related to the phenological characteristics. As such, an NDVI time series usually shows the plant life cycle of growth and senescence or a more stable growth state for the evergreen species. However, except for the normal vegetation changes, the negative bias (caused by cloud, snow, and other atmospheric effects) and the minor technical errors coexisting in the NDVI time-series products can greatly lower the NDVI data quality and should thus be removed. In this article, differing from the previous research utilizing comparisons followed by local adjustment or threshold-related iteration to approach the upper NDVI envelope, a one-step global optimization-based method is proposed to effectively and efficiently estimate high-quality NDVI time series by taking the properties of both latent NDVI data and atmospheric contamination into consideration. The method can be applied to different input conditions, both with or without an ancillary data quality flag. Furthermore, very few parameters are required in the calculation and, most importantly, they can be assigned as fixed values at the expense of very little accuracy loss. Moreover, there are no restrictions with regard to the sensor type, i.e., the proposed method is valid for use with various NDVI products and even self-calculated NDVI or EVI time series. In the following, we first describe how the observation model is constructed for NDVI data and then introduce the proposed method in detail.

A. Problem Formulation

Supposing that the atmospheric contamination in the observed NDVI time series can be regarded as additive noise, then the relationship between the noisy series $\mathbf{y}$ and its latent ideal series $\mathbf{x}$ can be described as

$$\mathbf{y} = \mathbf{x} + \mathbf{z} \tag{1}$$

where $\mathbf{y} = (y_1, \ldots, y_n)$, $\mathbf{x} = (x_1, \ldots, x_n)$, and $\mathbf{z} = (z_1, \ldots, z_n)$ denote the remotely sensed NDVI series, the ideal smooth time series, and the noise part, respectively.

B. Universal Optimization Framework for Curve Smoothing

According to the requirements and characteristics of curve smoothing, the task can be summarized as an optimization problem with two competing aspects. One is to minimize the residual between $\mathbf{x}$ and $\mathbf{y}$, and the other is to alleviate the unnatural disturbance to ensure a smooth $\mathbf{x}$. Hence, a universal optimization framework to estimate the smooth curve $\mathbf{x}$ or, equivalently, the noise component $\mathbf{z}$ ($\mathbf{z} = \mathbf{y} - \mathbf{x}$) can be built as

$$\min \frac{1}{p}||\mathbf{x} - \mathbf{y}||_p^p + \frac{\lambda}{q}||D_k\mathbf{x}||_q^q \tag{2}$$

where $p \in (0, 2]$ and $q \in (0, 2]$, and the first term in the above objective function measures the fitting error, while the second term penalizes the roughness of the estimated curve. $\lambda \geq 0$ is the regularization parameter used to balance the functioning degree of these two terms. D_k denotes the k-order difference matrix, which is constructed as a Toeplitz matrix (in which each descending diagonal from left to right is constant) with $n - k$ rows and n columns. Each row of D_k contains the pattern of the k-order difference operator with the first row equaling to, for example, $[1 \ -1 \ 0 \ \cdots \ 0]$ when $k = 1$ and $[1 \ -2 \ 1 \ 0 \ \cdots \ 0]$ when $k = 2$. In general, the choice of a specific combination of the ℓ^p-norm and ℓ^q-norm mainly depends on the property of the noise component $\mathbf{z}$ and the ideal series $\mathbf{x}$. For example, $p = 1$ fits better than $p = 2$ for capturing $\mathbf{z}$ when the noise component is in accordance with an impulse noise distribution but fits worse than $p = 2$ when it conforms to a Gaussian distribution.

At the time, when $p = 2$, $q = 2$, and $k = 2$, the ℓ^p-norm and ℓ^q-norm in (2) then become an ℓ^2-norm or Euclidean norm, while D_2 is the second-order difference matrix. Namely, the universal curve smoothing framework concretes into the well-known WS (which is also known as the Hodrick–Prescott filter [33], [34]), which has been widely used in biology and economics analysis [35] and has also been modified for use in remote sensing-based vegetation index reconstruction [1], [24]. In order to recover the negative bias in the NDVI time series, an iterative process is employed to gradually approach the upper envelope of the data by choosing and integrating those higher values from the curves before and after Whittaker smoothing. Unfortunately, this modification inherits the sensitivity to the regularization parameter λ and lacks the constraint of the noise property in the optimization modeling, so can sometimes cause oversmoothing. Comparatively speaking, the NDVI property tailored modeling shows more potential for high-quality time-series reconstruction.

C. Proposed Variational NDVI Reconstruction Method

1) Negatively Biased Atmospheric Contamination Constraint: In line with the assumption that the noise component $\mathbf{z}$ in an NDVI time series is almost negative (due to the poor atmospheric conditions, such as clouds), the relationship between $\mathbf{x}$ and $\mathbf{y}$ then satisfies the inequality

$$\mathbf{z} = \mathbf{y} - \mathbf{x} \leq 0. \tag{3}$$

It is easy to understand that $\mathbf{y} - \mathbf{x} < 0$ only occurs when the observed series is contaminated by negative noise, while $\mathbf{y} - \mathbf{x} = 0$ is always kept in a normal case. Inspired by the popular rectified linear unit (ReLU) activation function [36], [37], we rewrite (3) as

$$(\mathbf{y} - \mathbf{x})_+ = 0 \tag{4}$$

where $(\cdot)_+$ represents the operation to take the higher value between 0 and $\mathbf{y} - \mathbf{x}$. Clearly, the output of $(\mathbf{y} - \mathbf{x})_+$ is always nonnegative, so the minimization of $(\mathbf{y} - \mathbf{x})_+$ aims to penalize the case of $\mathbf{y} > \mathbf{x}$. In other words, it encourages $\mathbf{x}$ to approach the upper envelope of $\mathbf{y}$, which exactly corresponds with the requirement in NDVI filter design.

2) Joint Modeling of Gradual Vegetation Change and Negatively Biased Atmospheric Contamination: Note that the measure $(\mathbf{y} - \mathbf{x})_+$ alone is not functional. It is thus combined in the proposed method with the universal curve smoothing

optimization framework, and the objective function is given as

$$\min \frac{1}{p}||\mathbf{x}-\mathbf{y}||_p^p + \frac{\lambda}{q}||D_k\mathbf{x}||_q^q \quad \text{subject to} \quad (\mathbf{y}-\mathbf{x})_+ = 0. \tag{5}$$

Since the natural vegetation change possesses a gradual process profile, we choose the ℓ^2-norm to minimize the second-order difference of $\mathbf{x}$ ($q = 2$ and $k = 2$) to penalize the roughness of $\mathbf{x}$. As for the noise component modeling, we then use the ℓ^1-norm ($p = 1$) to penalize the differences between $\mathbf{x}$ and $\mathbf{y}$. As the noise in NDVI time series appears as sudden drops with large bias, $p = 1$ rather than $p = 2$ (as used in the WS) is finally chosen. This is because the ℓ^1-norm can not only accept huge jumps to deal with the abrupt changes similar to impulse noise, but it can also lead to sparsity (i.e., with many zero elements) for better preservation of the phenology-related key points in NDVI reconstruction [34], [38].

From the optimization point of view, the problem (5) is a constrained optimization problem, and a penalized method can be applied, i.e., we solve the following unconstraint optimization problem:

$$\min ||\mathbf{x}-\mathbf{y}||_1 + \frac{\lambda}{2}||D_2\mathbf{x}||_2^2 + \frac{\mu}{2}(\mathbf{y}-\mathbf{x})_+^2. \tag{6}$$

If the penalty parameter μ is chosen to be large enough, (6) is a good approximation of (5). However, from the practical point of view, the constraint condition of $(\mathbf{y}-\mathbf{x})_+ = 0$ holds true only in ideal cases, which means that problem (5) imposes quite strict constraints in real NDVI reconstruction. To obtain a more effective and flexible model, we convert $\mu \geq 0$ into another regularization parameter in (6). Thus, problem (6) can also be considered as a relaxation of (5).

In model (6), the larger the value of parameter μ, the stronger the influence of $(\mathbf{y}-\mathbf{x})_+^2$, and the closer to the upper envelope that $\mathbf{x}$ can be estimated. Above all, the proposed method takes full consideration of the fidelity of the data (especially those phenology-related key points), the smoothness of natural vegetation change, and the negatively biased property of atmospheric contamination. As a result, the proposed method is well suited to NDVI time-series reconstruction and similar tasks such as EVI time-series reconstruction.

3) More General Version With Quality Assurance Information: Quality assurance/control information [1], [19], [22], [39]–[41] is used to assess the quality of each pixel, and cloud, water, shadow, and ice/snow flags are provided along with the remote sensing product. These ancillary data can help to control the reconstruction process to make the results more reliable. Hence, when the ancillary quality information is available, a more precise model is as follows:

$$\min ||C(\mathbf{x}-\mathbf{y})||_1 + \frac{\lambda}{2}||D_2\mathbf{x}||_2^2 + \frac{\mu}{2}(C(\mathbf{y}-\mathbf{x}))_+^2 \tag{7}$$

where C denotes the diagonal matrix with diagonal elements $c_{i,i} \in [0, 1]$ related to the corresponding quality of the observed pixel y_i. According to the quality assurance information, all the observed pixels can be divided into five groups: good, uncertain, marginal, missing, and seriously contaminated. In practice, for simplicity, good pixels are set with the maximum weight of 1, and uncertain or marginal pixels are given a moderate weight of 0.5. The minimum weight of 0 is used for the missing or seriously contaminated pixels. If part of the quality assurance information is missing, those pixels can be regarded as good, with a weight of 1. When no ancillary quality data are provided, Q becomes an identity matrix, and in such a case, (7) is equal to (6). Since the model in (6) is just a special case of (7), a more general description of the proposed model, i.e., (7), is used throughout.

D. Numerical Solution

Before analyzing the optimization of the proposed NDVI reconstruction method, we first discuss the upper envelope measure and expand it as

$$(\mathbf{y}-\mathbf{x})_+^2 = \begin{cases} (y_i - x_i)^2, & i \in \mathcal{A} \\ 0, & i \in \mathcal{A}^c \end{cases} \tag{8}$$

where $\mathcal{A} \triangleq \{i : x_i < y_i\}$ and $\mathcal{A}^c$ denotes the complement set of $\mathcal{A}$. By introducing a diagonal matrix U with diagonal values $u_{i,i} = 1$ for $i \in \mathcal{A}$ and $u_{i,i} = 0$ for $i \in \mathcal{A}^c$, (8) can be rewritten in a matrix form as

$$(\mathbf{y}-\mathbf{x})_+^2 = U(\mathbf{y}-\mathbf{x})^2. \tag{9}$$

Then, to solve the whole problem, the iteratively reweighted least-squares (IRLS) method [42]–[44] is employed due to its efficiency in the ℓ^p-norm ($p \in (0,\ 2]$) minimization. At each iteration, the solution is equal to weighted least-squares minimization of

$$\hat{\mathbf{x}}^{(k+1)} = \arg\min\left\{\frac{1}{2}||\ W^{(k)\,1/2}C(\mathbf{x}-\mathbf{y})\ ||_2^2 + \frac{\lambda}{2}||\ D_2\mathbf{x}\ ||_2^2 + \frac{\mu}{2}U^{(k)}(C(\mathbf{y}-\mathbf{x}))^2\right\} \tag{10}$$

with weighting matrix

$$W^{(k)} = \operatorname{diag}\left(\mathcal{T}\left(C\left(\mathbf{x}^{(k)}-\mathbf{y}\right)\right)\right) \tag{11}$$

where

$$\mathcal{T}(s) = \begin{cases} |\ s\ |^{-1}, & \textbf{if}\ |\ s\ | > \varepsilon \\ \varepsilon^{-1}, & \textbf{if}\ |\ s\ | \leq \varepsilon \end{cases} \tag{12}$$

and ε is a small positive number.

Thus, the closed-form solution of $\hat{\mathbf{x}}^{(k+1)}$ is

$$\hat{\mathbf{x}}^{(k+1)} = \left(C^T W^{(k)} C + \lambda D_2^T D_2 + \mu U^{(k)} C^T C\right)^{-1} \times \left(C^T W^{(k)} C + \mu U^{(k)} C^T C\right)\mathbf{y}. \tag{13}$$

As the boundary processing technique has an influence on the reconstruction results of the endpoints, we can expand the original data based on the Neumann boundary condition (setting the data outside the domain as a reflection of the data inside) [45] to alleviate the possible adverse edge effect. More specifically, by denoting the expanded data as $(y_{-m+1}, \ldots, y_0, y_1, \ldots, y_n, y_{n+1}, \ldots, y_{n+m})$, the data outside $\mathbf{y}$ then satisfy

$$\begin{cases} y_0 = y_1 \\ y_{-m+1} = y_m \end{cases} \quad \text{and} \quad \begin{cases} y_{n+1} = y_n \\ y_{n+m} = y_{n-m+1}. \end{cases} \tag{14}$$

III. Results

A. Data

To evaluate the proposed variational NDVI reconstruction method, the global inventory monitoring and modeling system (GIMMS) dataset (version number 3g.v1) generated from data obtained by the AVHRR sensor on the NOAA satellites was used in this study. The GIMMS NDVI3g.v1 product has a temporal resolution of twice a month (composited by the MVC technique) and a spatial resolution of 1/12 of a degree [25], [46]. The temporal coverage of the dataset ranges from July 1981 to December 2015. The flags provided along with the GIMMS data marked by values of 0, 1, or 2 separately indicate that the NDVI values are calculated from the raw data, processed by spline interpolation, or may correspond to snow/cloud cover. In the experiments, the NDVI values calculated from the raw data were regarded as good (with a weight of 1), and the spline interpolated values were regarded as uncertain (with a moderate weight of 0.5), while the possible snow/cloud covered values were considered as bad (with a weight of 0).

The International Geosphere–Biosphere Program (IGBP) classification maps [5] in the MODIS Land Cover Climate Modeling Grid Product (MCD12C1) were exploited to test and compare the NDVI reconstruction performance over different land-cover types. The MCD12C1 data at a spatial resolution of 0.05° have been released year by year since 2001. For spatial and temporal consistency, the IGBP maps were resampled to the GIMMS spatial resolution and selected from 2001 to 2015.

B. Implementation of the Comparison Methods

Five widely used NDVI reconstruction methods were selected for comparison: 1) the RMMEH method [26]; 2) the IDR method [25]; 3) the SG method [22]; 4) the WS method [1]; and 5) the Fourier-based moving weighted harmonic analysis (MWHA) method [19]. The different versions of the proposed one-step global variational reconstruction (OGVR) method (with and without ancillary data) named OGVR and OGVR_nw (OGVR with no weight) were both tested and compared. Among the different methods, only RMMEH, IDR, and OGVR_nw are methods that are independent of ancillary flag information.

The parameters in the five comparison methods were set based on the advice given by the respective authors. In detail, RMMEH can be implemented easily as neither ancillary data nor parameters are required in the procedure. For the IDR method, only one threshold parameter needs to be set. Here, the value of 0.02 was used, as in [25]. Due to the embedded least-squares fitting [22] and the cross-validated optimization [1] techniques introduced in the SG and WS methods, the two parameters (the half-width of the smoothing window and the degree of the polynomial) in the SG processing flow and the one smoothing parameter in the WS method were calculated automatically in this study, at the cost of increased time. There are two main parameters, i.e., the support domain radius r_d and the frequency number n_f, in the MWHA method. After multiple tests, r_d and n_f were, respectively, set as 5 and 1, which were the same as the values used in [19]. The parameters of the proposed model were fixed as $\lambda = 100$ and $\mu = 100$ in all the experiments. More details of the parameter setting of the proposed method can be found in Section IV.

C. Simulated Data Experiments

In the simulated experiments, the reference NDVI curve was synthesized to represent a basic pattern of the vegetation coverage at each pixel, which was then used as the ground truth to evaluate the reconstruction methods quantitatively [40], [47], [48]. Specifically, for the AVHRR GIMMS data, the NDVI values calculated by spline interpolation or corresponding to snow/cloud cover were first regarded as the poor-quality points to be discarded. The other valid NDVI values in the same day of year (DOY) from July 1981 to December 2015 were then averaged to generate the reference NDVI. There were a total of 24 composites in the reference NDVI. If the good points from the same DOY numbered less than nine (approximately 1/4 of the total points), the corresponding mean value had much less chance of reflecting the essential state of the vegetation. Thus, cubic interpolation was performed on the reference NDVI to compute a more reliable value from the neighboring representative composites. As each composite in the reference NDVI was calculated from different years over a long period of time, the disturbances related to annual variations were further suppressed by the IDR method. Through the procedures introduced above, the reference NDVI curves were obtained in which most of the noise and interannual disturbances were alleviated.

The next step was to add gaps and noise into the reference NDVI. In this study, three types of gaps/noise were added simultaneously.

1) *Random Gaps*, where 20% of the composites were randomly chosen and reduced by a random value between the range of [0.2, 0.4] to simulate the case of random cloudy days.
2) *Continuous Gap*, where no more than five consecutive composites, i.e., no more than 75 days, were randomly chosen to simulate the case of long-term cloud cover. These chosen composites were reduced by a random value within the same range used for the *Random Gaps*.
3) *Gaussian Noise*, where 15% of the composites were randomly chosen by adding a normally distributed random value with mean 0 and standard deviation 0.05. This kind of noise was added as the simulation of uncertain data points in the real NDVI curve.

Meanwhile, the corresponding flag information was also generated. All the untouched composites in the reference NDVI were marked as good with a weight of 1. The gap-related composites and Gaussian noise composites were separately marked as bad and uncertain with weights of 0 and 0.5, respectively, in the proposed method, and 0 and 1 in the comparison methods [1], [19], [22].

The quantitative assessment was based on three full-reference indices: the correlation coefficient (CC), the mean absolute error (MeanAE, also known as the mean bias), and the maximum absolute error (MaxAE). The CC was used to

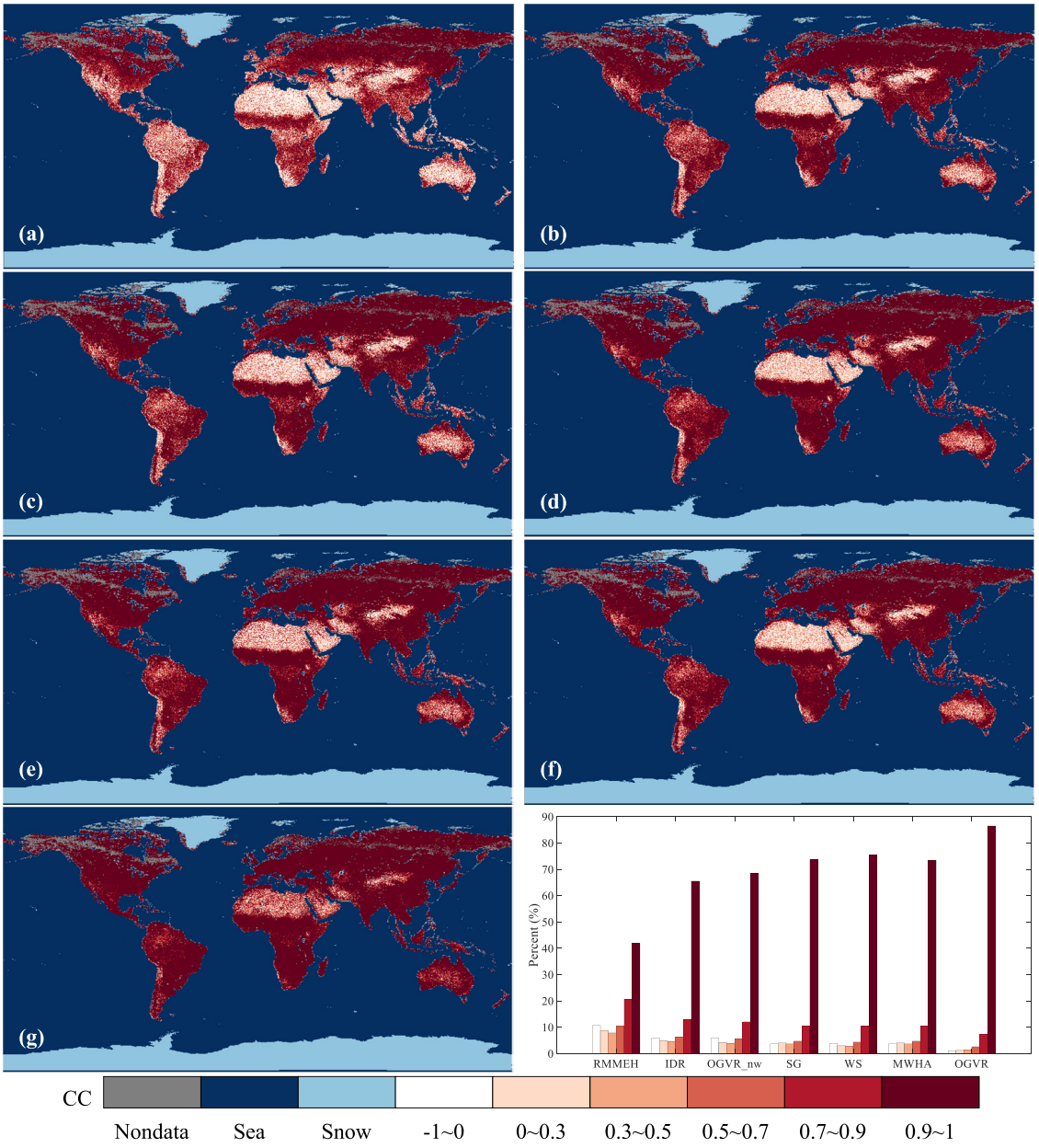


Fig. 1. Global distribution of the NDVI reconstruction under CC evaluation. The corresponding percentage distribution is shown in the bar graph. (a) RMMEH. (b) IDR. (c) OGVR_nw. (d) SG. (e) WS. (f) MWHA. (g) OGVR.

evaluate the similarity between the reference and reconstructed NDVI curves. The MeanAE was taken as the indicator to measure the model error, whereas the MaxAE was employed to reflect the model stability. Their numerical definitions are given as follows:

$$\mathbf{CC} = \frac{\sum_{i=1}^{n}(x_i - \mu_{\mathbf{x}})(\hat{x}_i - \mu_{\hat{\mathbf{x}}})}{\sqrt{\sum_{i=1}^{n}(x_i - \mu_{\mathbf{x}})^2}\sqrt{\sum_{i=1}^{n}(\hat{x}_i - \mu_{\hat{\mathbf{x}}})^2}} \tag{15}$$

$$\mathbf{MeanAE} = \frac{\sum_{i=1}^{n}|\hat{x}_i - x_i|}{n} \tag{16}$$

$$\mathbf{MaxAE} = \max_{1\le i\le n}|\hat{x}_i - x_i| \tag{17}$$

where $\mu_{\mathbf{x}}$ and $\mu_{\hat{\mathbf{x}}}$, respectively, denote the mean values of the whole reference NDVI curve $\mathbf{x} = (x_1, \ldots, x_n)$ and the reconstructed NDVI curve $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_n)$. A better performance can be reflected by higher CC and lower MeanAE and MaxAE.

Figs. 1–3 show the visualized quantitative evaluation of the different NDVI reconstruction methods at a global scale. Since two groups of methods (using ancillary data or not) were both tested, the comparison within the group requiring no auxiliary information is first introduced for the sake of fairness. Among the methods of RMMEH, IDR, and OGVR_nw, RMMEH achieves the lowest curve similarity [Fig. 1(a)] and the highest data error [Figs. 2(a) and 3(a)] over the entire region, which indicates that its sensitivity to noise is relatively high. IDR and OGVR_nw exhibit obvious advantages over the RMMEH method in the visualized quantitative global evaluation, but

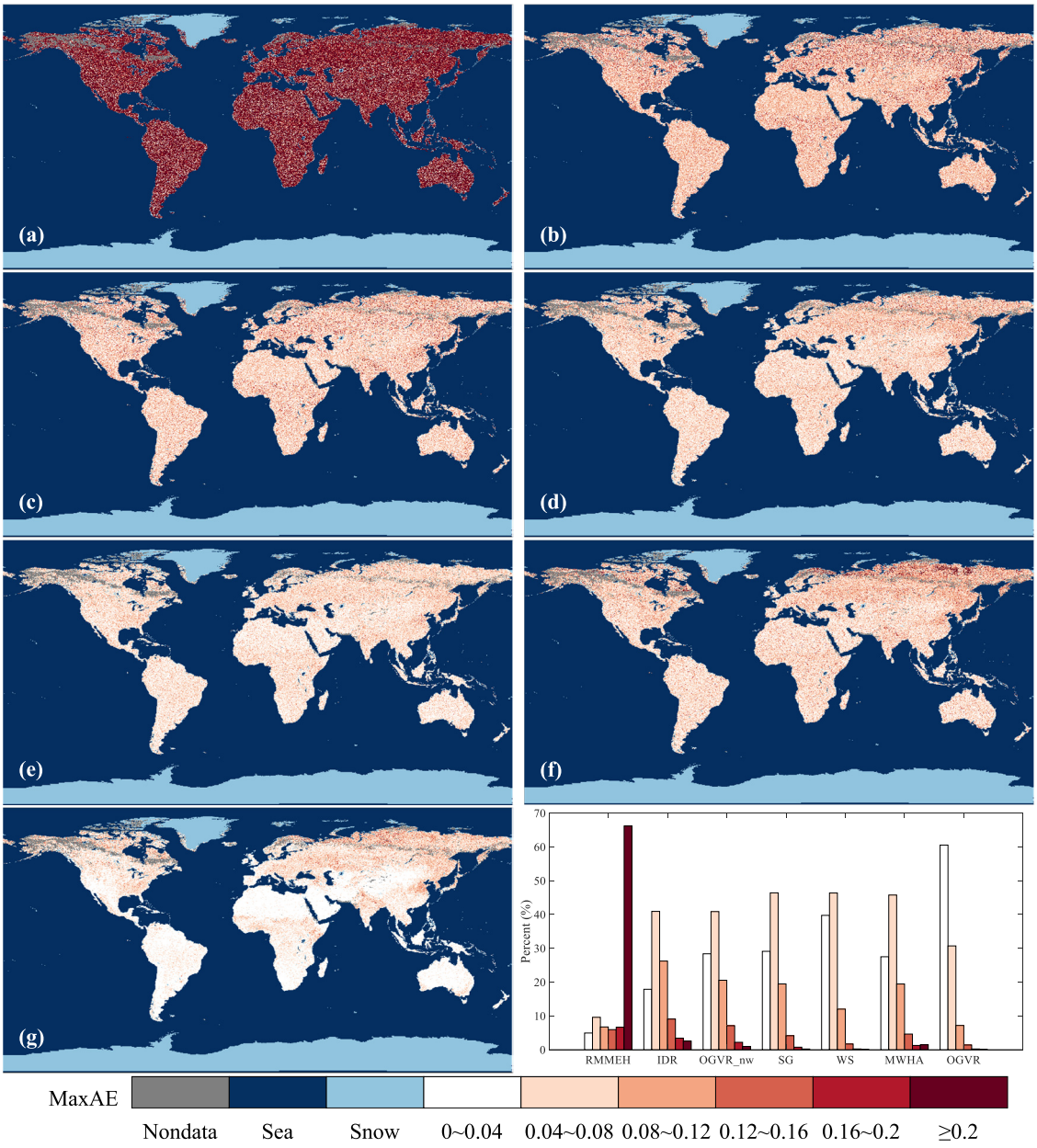


Fig. 2. Global distribution of the NDVI reconstruction under MaxAE evaluation. The corresponding percentage distribution is shown in the bar graph. (a) RMMEH. (b) IDR. (c) OGVR_nw. (d) SG. (e) WS. (f) MWHA. (g) OGVR.

the differences between each method are much less clear. Based on the statistical results shown in the bar graphs, the proposed OGVR_nw method outperforms the IDR method as it obtains a superior performance in two of the three index evaluations. Given that the reference NDVI in this work was calculated from the IDR method, more significant advantages of OGVR_nw over IDR can be expected in other tasks.

By adding the flag information into the NDVI reconstruction, the SG, WS, MWHA, and OGVR methods perform better, overall, than the RMMEH, IDR, and OGVR_nw methods. For the SG, WS, and MWHA methods, although a slightly higher point percentage with CC $\geq$ 0.9 can be found in the bar graph (from Fig. 1) compared to all the methods requiring no auxiliary information, their advantages decrease in different degrees under the MaxAE and MeanAE measurements. For example, the point percentage of MaxAE $\leq$ 0.04 in the SG and MWHA methods is nearly equal to the value in the OGVR_nw method, as shown in Fig. 2. A similar example can be seen in Fig. 3, in which the percentage distribution of the WS method under MeanAE evaluation is almost the same as that of the IDR method. Unlike the methods described above, the proposed OGVR model recovers the noisy NDVI curves well, with the most widely distributed points of the highest CC and the lowest MaxAE and MeanAE values. Namely, the proposed method can not only recover the correct curve trend (the vegetation growth trend) under noise interference but also preserve most of the useful information in the time series.

It is worth noting that points in rain forest and desert show much smaller greenness changes within a year, and thus, larger errors are more easily generated after removing the corresponding added noise from the reference NDVI time

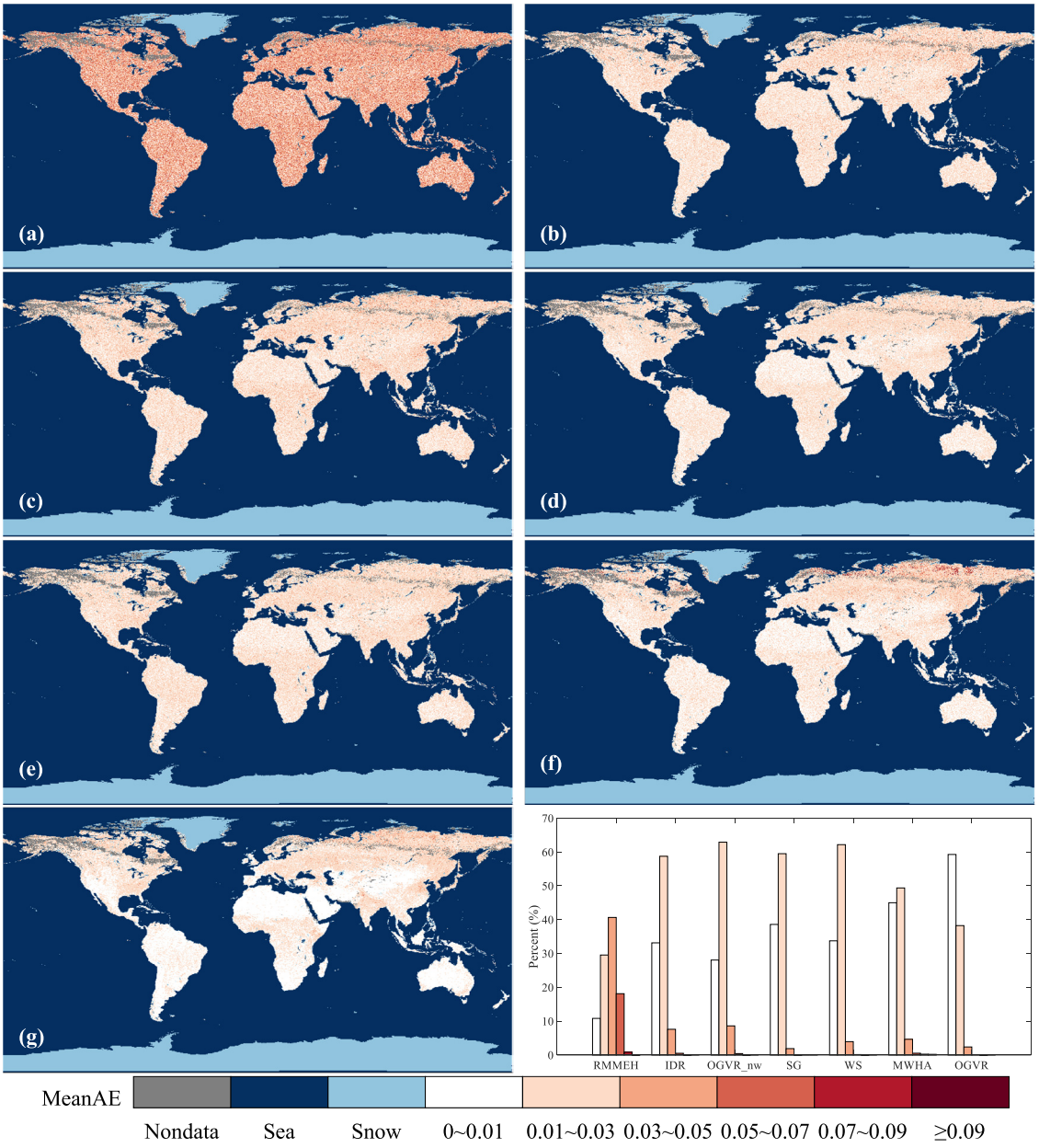


Fig. 3. Global distribution of the NDVI reconstruction under MeanAE evaluation. The corresponding percentage distribution is shown in the bar graph. (a) RMMEH. (b) IDR. (c) OGVR_nw. (d) SG. (e) WS. (f) MWHA. (g) OGVR.

series. These points can be used to test the stability of the NDVI filtering methods. In the comparison methods, the points with lower similarity or larger error are concentrated in the regions of rain forest and desert. In contrast, the proposed OGVR method shows a good performance by recovering the noisy NDVI curves well for different land-cover types, particularly in the challenging areas of rain forest and desert. According to the overall results, the proposed model exhibits a satisfactory denoising capability, whether the flag information is employed or not. Even in the case of making no use of the flag data, the proposed OGVR_nw method is still competitive when compared to the existing ancillary data-dependent NDVI reconstruction methods.

D. Real-Data Experiments

Without the help of reference data, qualitative assessment through visual comparison is the main way to judge the performance of different methods in real-data experiments. To make the experiments more convincing, the long-term NDVI time series for different land-cover types based on the IGBP classification were used. More specifically, the GIMMS data and the IGBP maps within the same time span from 2001 to 2015 were first collected. An IGBP class was then assigned to each GIMMS pixel when two conditions were simultaneously satisfied: 1) the resampled IGBP pixel was generated by pixels from the same land-cover class (spatial homogeneity) and 2) the class type of the resampled IGBP pixel remained unchanged over the entire 15 years (temporal homogeneity). The candidate points with IGBP class information are shown in Fig. 4. The areas colored in gray correspond to those pixels with a lack of spatial homogeneity or temporal homogeneity. Due to the space limitations, ten control points were randomly chosen from the different land-cover types for the comparison. Their geographical locations along with

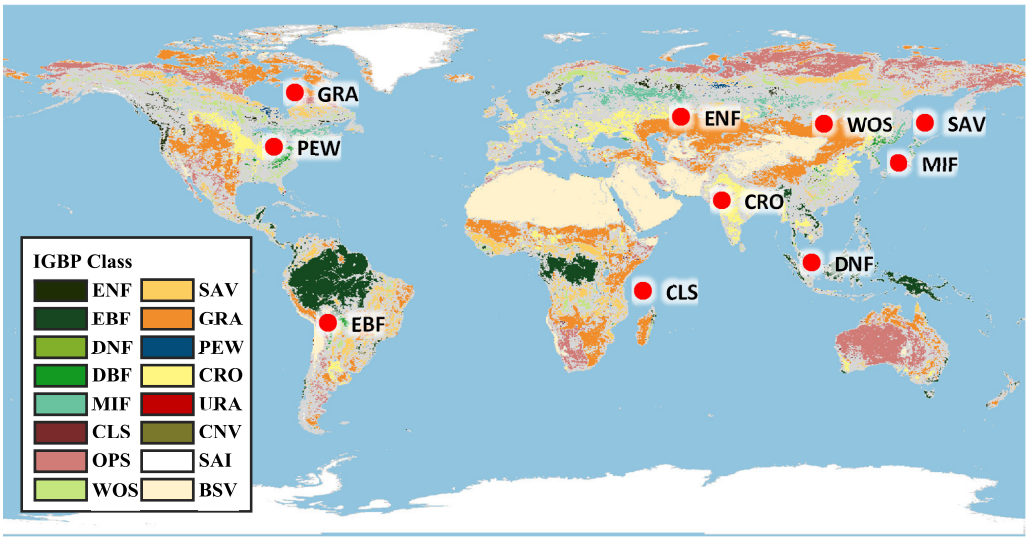


Fig. 4. Spatial distribution of ten randomly chosen control points for comparison. Water body and noncandidate points are masked by light blue and gray colors, respectively. ENF: evergreen needleleaf forests. EBF: evergreen broadleaf forests. DNF: deciduous needleleaf forests. DBF: deciduous broadleaf forests. MF: mixed forests. CSH: closed shrublands. OSH: open shrublands. WSA: woody savannas. SAV: savannas. GRA: grasslands. WET: permanent wetlands. CRO: croplands. URB: urban and built-up lands. CVM: cropland/natural vegetation mosaics. SNO: permanent snow and ice. BSV: barren or sparsely vegetated.

the classification information are given in Fig. 4. In order to maintain consistency with the original settings [1], [19], [22], the raw NDVI values obtained from the data, the NDVI values obtained by spline interpolation, and those values corresponding to possible snow/cloud cover were set to 1, 0.5, and 0 for the proposed method, and 1, 1, and 0 for the comparison methods, respectively.

The reconstructed NDVI time series calculated by the six approaches at the ten control points are shown in Fig. 5. It can be seen clearly that although the GIMMS data are composited by the MVC technique, a lot of noise still remains in the NDVI time series. In addition, the flag information only indicates some of the contaminated points, and to complicate matters further, these contaminated points can appear even beyond the local low points. Faced with different cases of complicated GIMMS NDVI time series, the proposed OGVR method can successfully eliminate the unnatural disturbances, maintaining a good agreement with the upper envelope of the good points, while capturing the gradual change of vegetation and also conquering the problem of continuous cloud/snow. The satisfactory and stable performance in all the control points demonstrates the effectiveness and robustness of the OGVR method. As a method that is independent of the ancillary flag information, the RMMEH method reconstructs the time series based mainly on the local mean and median values and is thus sensitive to local oscillations [e.g., Fig. 5(a) and (c)]. Compared with the RMMEH method, the other methods that are independent of the ancillary flag information, i.e., IDR and OGVR_nw, output better reconstruction results. Although the reconstruction performances of the IDR and OGVR_nw methods look similar, their close-up details reveal the differences. The OGVR_nw method generates smoother time series and is better able to capture the gradual changes of vegetation by using a global constraint instead of local calculation. According to the results shown in Fig. 5, the estimated NDVI curves of the RMMEH, IDR, and OGVR_nw methods are consistent with the proposed OGVR method in most cases but are significantly worse than OGVR in dealing with contaminated high-value points [e.g., Fig. 5(c) and (d)]. This is understandable as OGVR utilizes more information (flag information) than RMMEH, IDR, and OGVR_nw.

Under the same input conditions, the proposed OGVR method was then further evaluated and compared with the SG, WS, and MWHA methods. The SG method can conquer parts of the fluctuations in the NDVI time series but is vulnerable to frequent fluctuations. For instance, for the evergreen broadleaf forest (EBF) and woody savannas (WSA) land-cover types, the SG method is unable to capture the gradual changes of vegetation growth [Fig. 5(b) and (f)]. Compared to SG, the WS method always obtains very smooth time series, whereas the results shown in Fig. 5(c), (h), and (i) reveal its heavy dependency on good flag information, in which regionally large bias is caused by the method being misled by some incorrect good points or outliers. The MWHA method inherits the advantage of the Fourier-based fitting method, in which it can reconstruct smooth curves, but also the disadvantage, in which it easily generates large bias from the original NDVI values. Although the local adjustment is designed in MWHA to alleviate the unwanted bias, the residual bias is nonnegligible. Examples are shown in Fig. 5(d), (e), and (i). It should also be mentioned that the MWHA method fails to reconstruct the long-term continuous cloud/snow time series in the *c* and *h* control points. This is mainly because the parameters of MWHA fail to converge for the seriously contaminated cases. In summary, the proposed OGVR method can capture the gradual change of vegetation well and can properly handle the contaminated points, no matter if they are locally high or low. Even in seriously contaminated cases, such as the continuous cloud/snow problem, the OGVR method still performs stably and is unaffected by the frequent oscillations. Considering that, in the OGVR reconstruction, all the parameters are fixed and the ancillary flag information is not strictly required (OGVR_nw), the proposed method can also be regarded as having advantages in execution.

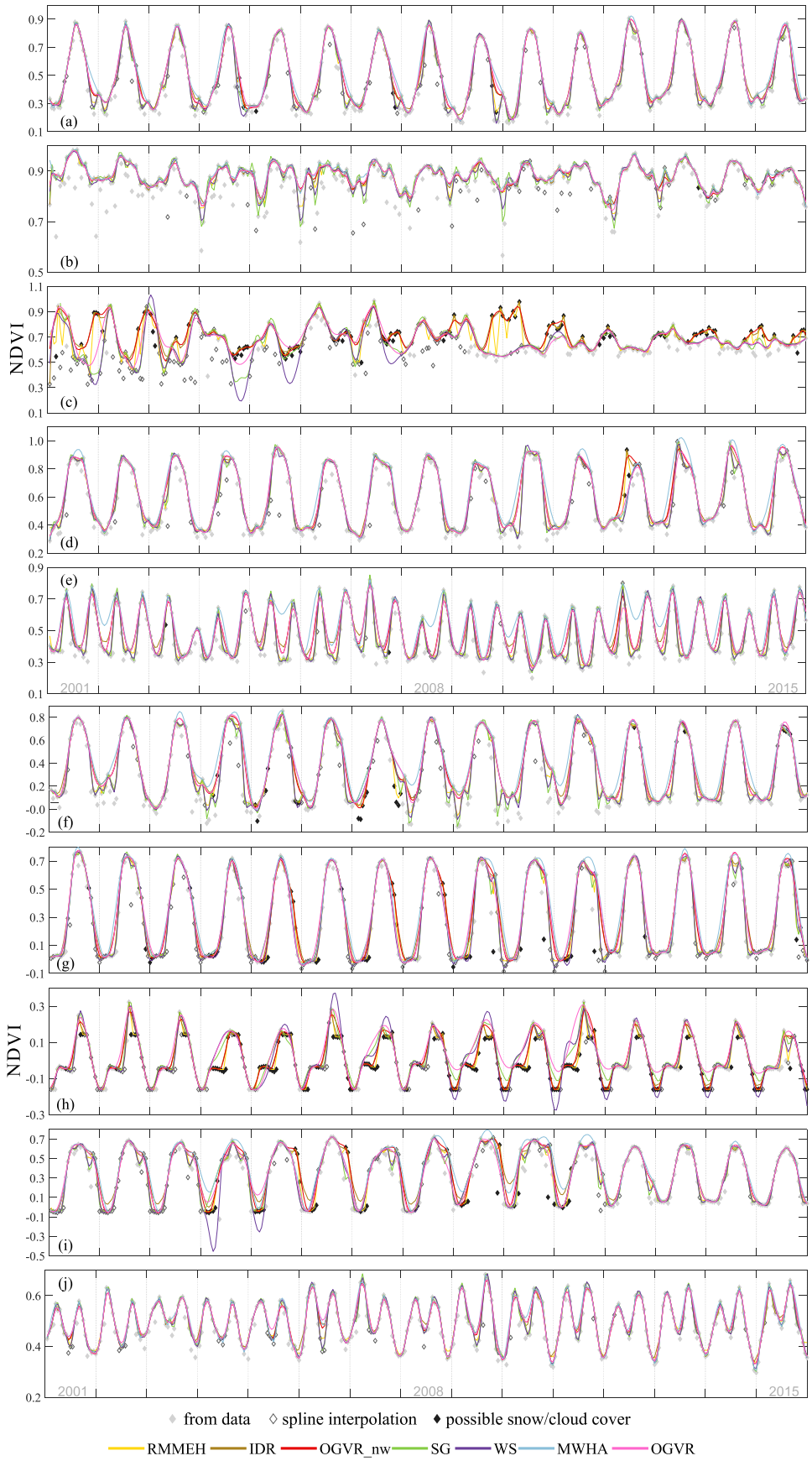


Fig. 5. Performance of the various methods for reconstructing NDVI time series during 2001–2015 at 10 control points (Fig. 4). (a) ENF. (b) EBF. (c) DNF. (d) MF. (e) CSH. (f) WSA. (g) SAV. (h) GRA. (i) WET. (j) CRO.

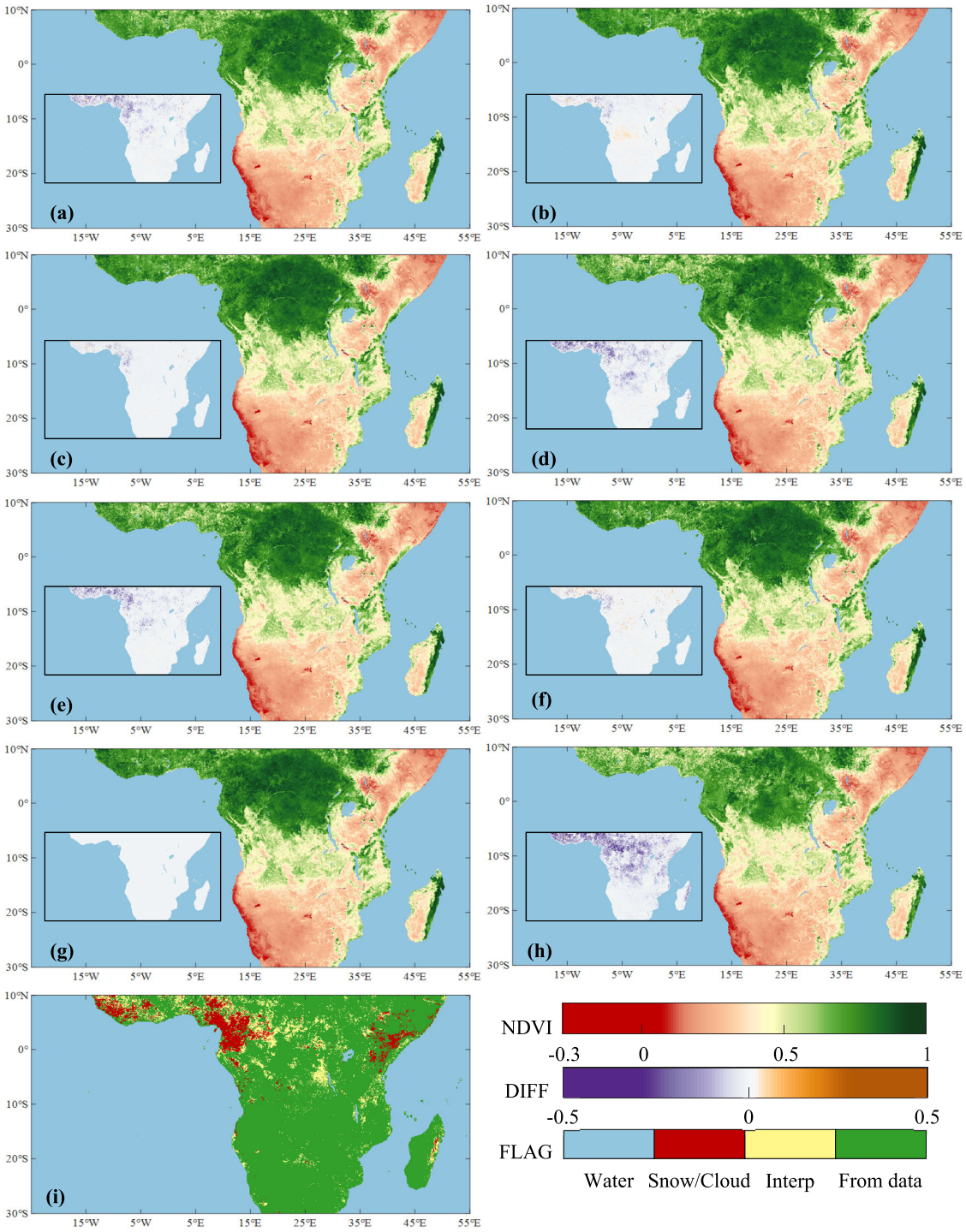


Fig. 6. Regional performance of the various methods for NDVI reconstruction in the first half of August in 2005 in south-central Africa. DIFF: difference image obtained via subtracting the results of the other methods from those of OGVR. Interp: spline interpolation. Snow/Cloud: possible snow/cloud cover. (a) RMMEH. (b) IDR. (c) OGVR_nw. (d) SG. (e) WS. (f) MWHA. (g) OGVR. (h) Raw image. (i) Flag.

Examples of performance comparisons at regional scales with extensive cloud cover or spline interpolation are given in Figs. 6 and 7. The first example shown in Fig. 6 is selected from south-central Africa, where the Congo Basin is located, which contains the world's second-largest contiguous rainforest. Due to the nature of the terrain, the western part of the Congo rainforest is always covered by cloud. In the first half of August, the savanna areas north of the Congo rainforest are in the rainy season, so the NDVI values both near and north of the equator should remain at a high level. However, many abnormal low values can be found in the raw NDVI in 2005 (Fig. 6), and the distribution of these low values is consistent with the location of the cloud/interpolation flags [see Fig. 6(h) and (i)]. The other example is chosen from East Asia (Fig. 7), where the majority of the areas belong to the monsoon climate zone, which is

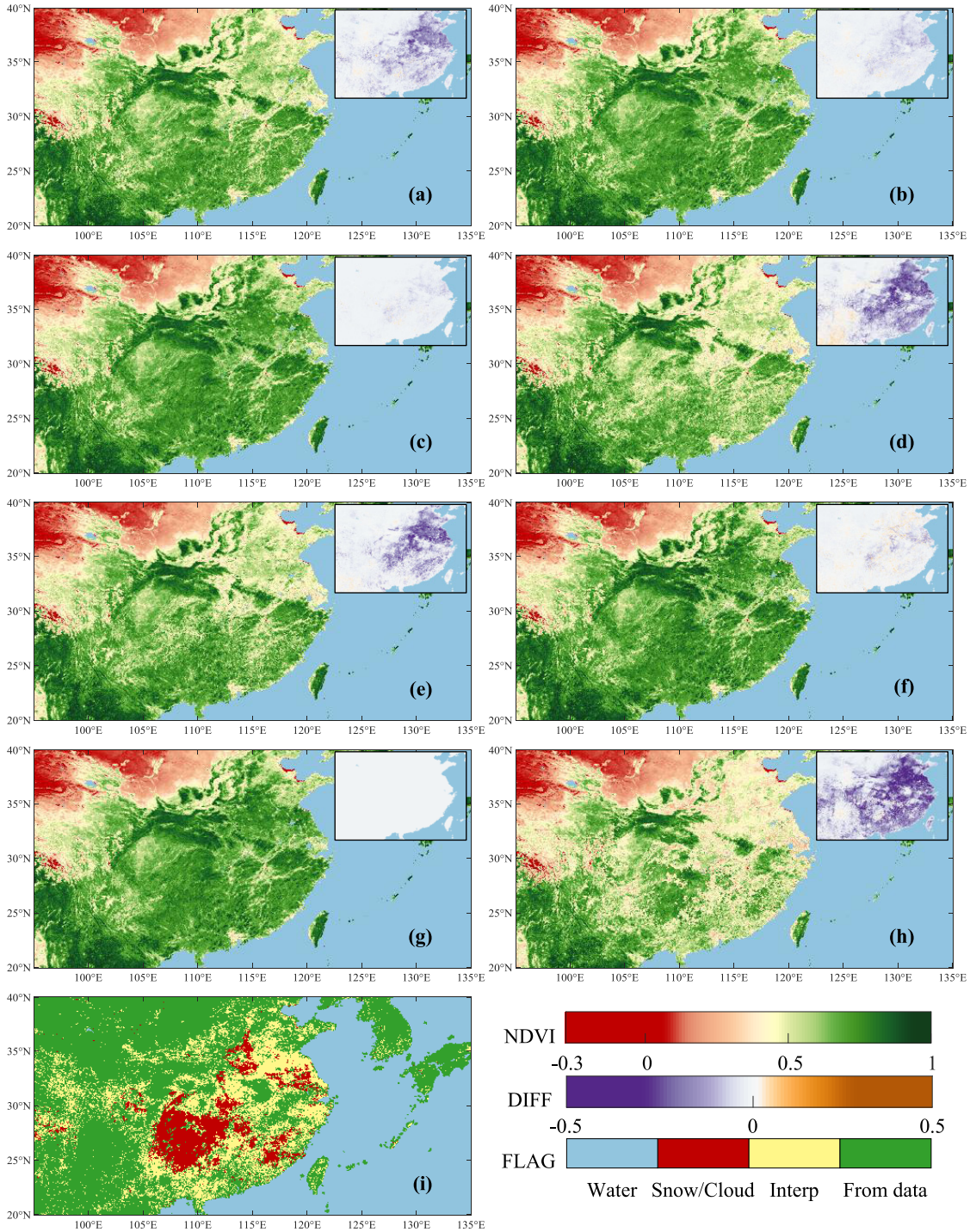


Fig. 7. Regional performance of the various methods for NDVI reconstruction in the first half of June in 2007 in East Asia. DIFF: difference image obtained via subtracting the results of the other methods from those of OGVR. Interp: spline interpolation. Snow/Cloud: possible snow/cloud cover. (a) RMMEH. (b) IDR. (c) OGVR_nw. (d) SG. (e) WS. (f) MWHA. (g) OGVR. (h) Raw image. (i) Flag.

characterized by humid summers and dry winters. In the first half of June, areas affected by the monsoon have entered into the rainy season, while the vegetation coverage also approaches the highest level. However, the corresponding raw NDVI in 2007 is seriously contaminated by either cloud or interpolation, appearing as a wide range of low values [see Fig. 7(h) and (i)].

For a clearer observation, difference maps obtained via subtracting the results of the other methods from those of OGVR are also given. Hence, a negative value in the difference map indicates that the NDVI reconstruction result of this pixel is lower than that of OGVR and vice versa. According to the reconstruction results shown in Figs. 6 and 7, the RMMEH, SG, and WS methods all fail to increase some

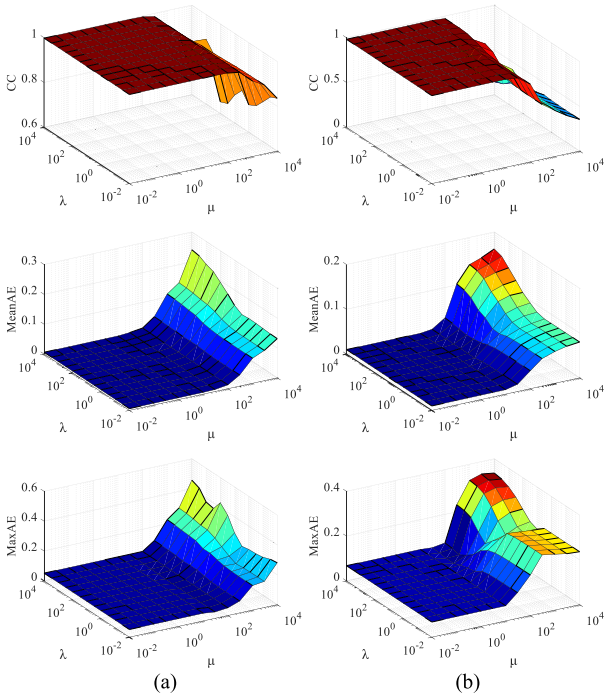


Fig. 8. Sensitivity analysis between parameters λ and μ in the OGVR model at (a) control point a (ENF) and (b) control point e (CSH).

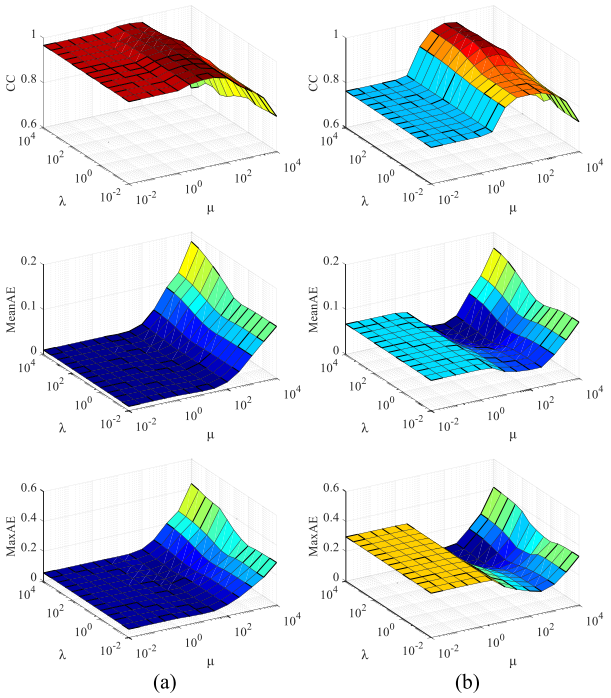


Fig. 9. Sensitivity analysis between parameters λ and μ in (a) OGVR model and (b) OGVR_nw model, based on the average reconstruction performance at the ten control points.

parts with low NDVI values in the test regions. Although the spatial distributions of their recovered NDVI become much smoother, the remaining discontinuity is still nonnegligible. Consistent results can be found in their difference maps, where many negative values are distributed over a large area. For the RMMEH method, it recovers the depressed NDVI by employing the statistics of the median and mean in the time series and can thus be affected by short-term oscillations. The reason why the SG and WS methods obtain relatively poor performances is that these two methods are heavily reliant on good-point flag (in GIMMS). Namely, when the NDVI values with good-point flag are, in fact, not good, the results can be misdirected. The MWHA method is better able to increase the contaminated NDVI values in Figs. 6(f) and 7(f), but some "speckle" artifacts reveal its instability. Taking Fig. 6(f) as an example, the artifacts with obvious low values near the Congo River are distinct in both the NDVI and difference maps. The IDR, OGVR_nw, and OGVR methods obtain satisfactory and similar visual results in Figs. 6 and 7. The main differences between the IDR, OGVR_nw, and OGVR methods lie in the cloud-flag and interpolation-flag areas, which can be seen more clearly in their difference maps. With the help of the flag information, the recovered NDVI obtained by the OGVR method is more homogeneous, with more continuous coloring throughout the whole space [see Figs. 6(g) and 7(g)]. These results confirm the effectiveness and stability of the proposed model, whether flag information is employed or not.

IV. Discussion

A. Parameter Selection

In the OGVR and OGVR_nw methods, the properties of both the latent NDVI time series and the atmospheric

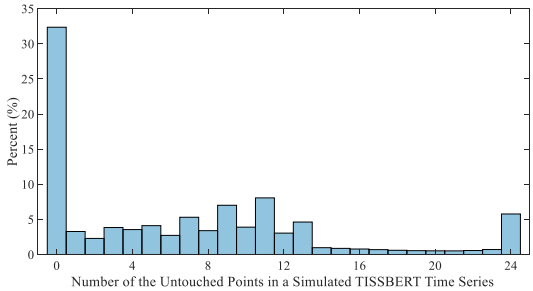


Fig. 10. Global statistics of the untouched points in TISSBERT simulated time series.

contamination are fully considered and constructed as constraint terms to estimate the high-quality NDVI time series. Therefore, the regularization parameters λ and μ in (7) are very important. λ works to control the smoothness of the estimated NDVI curve. A larger λ induces a smoother curve, whereas a smaller λ captures more frequent changes. For μ, it tunes the degree to approach the upper envelope of the original NDVI time series. The estimated curve goes through the middle of the original time series as $\mu \to 0$ and goes by the top as $\mu \to \infty$. To better understand the influence of λ and μ on the reconstruction performance, parameter sensitivity analyses were undertaken based on the synthesized reference NDVI curves at the locations of the ten control points (see Fig. 4). Since the OGVR model can be regarded as a more general version of the OGVR_nw model, the sensitivity analysis for OGVR was done first. By taking control point *a* from the evergreen needleleaf forests (ENF) category and control point *e* from the closed shrublands (CSH) category as examples, the changes of the CC, MeanAE, and MaxAE values varying with the different combinations of parameters λ and

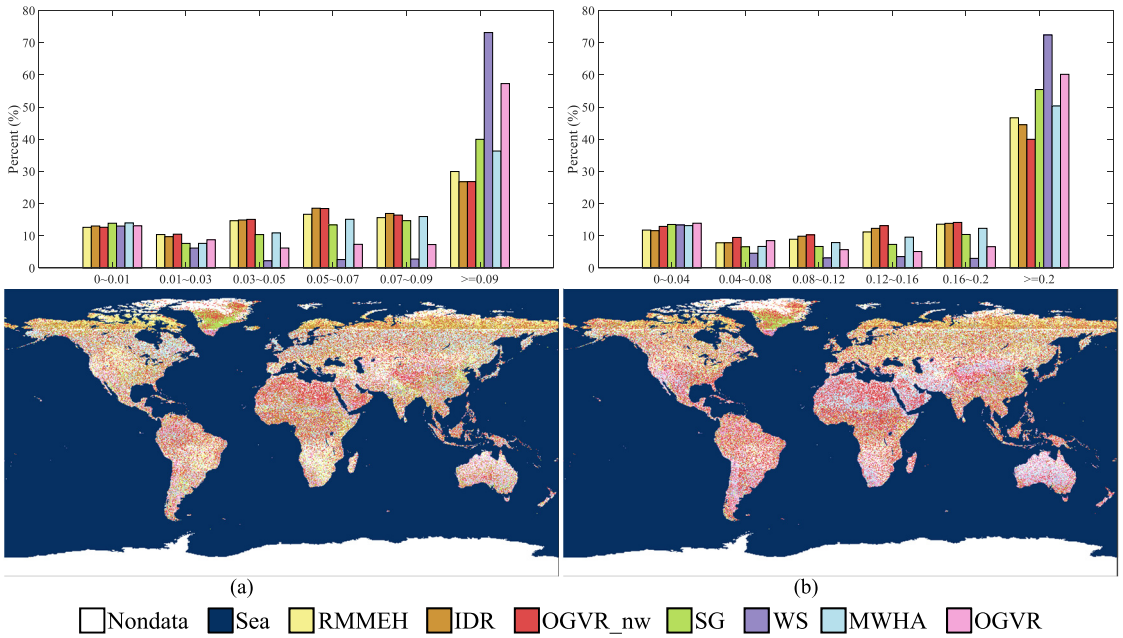


Fig. 11. Method comparison in TISSBERT data under MeanAE and MaxAE evaluations. From top to bottom: the percentage distributions of NDVI reconstruction errors and the corresponding global distributions of the best performing method. (a) MeanAE. (b) MaxAE.

TABLE I

GLOBAL RATIOS (%) OF THE BEST PERFORMING METHOD IN TISSBERT SIMULATED DATA

Evaluation Index	MeanAE	MaxAE
RMMEH	16.7	16.5
IDR	19.8	10.7
OGVR_nw	**24.7**	**33.2**
SG	6.0	5.9
WS	2.0	3.3
MWHA	21.0	16.6
OGVR	9.8	13.8

TABLE II

RUNNING TIME(S) OF THE COMPARISON METHODS WITH THREE TIME-SERIES LENGTHS

Data Size	24 ×10000	120×10000	360×10000
RMMEH	0.148	0.170	0.219
IDR	0.181	0.711	1.953
OGVR_nw	2.211	3.743	5.631
SG	1.524	1.374	1.531
WS	55.957	552.22	2524.175
MWHA	13.423	176.812	6279.856
OGVR	2.573	3.982	5.806

μ are shown in Fig. 8. Ideally, the CC, MeanAE, and MaxAE should be, respectively, equal to 1, 0, and 0. It is clear that the proposed OGVR method performs satisfactorily and stably with the changes of λ and μ under the different evaluation indices at the control points of the different land-cover types. Specifically, within the range of $\lambda \in [10^{-2}, 10^{4}]$ and $\mu \in [10^{-2}, 10^{2}]$, the impact on reconstruction performance caused by the changes of parameters λ and μ is almost negligible. In order to further verify the stability of the proposed method, a sensitivity comparison between OGVR and OGVR_nw was conducted by using the average reconstruction performance at all ten control points. According to the results shown in Fig. 9, OGVR doubtlessly exhibits satisfactory stability in average performance. Although the sensitivity of OGVR_nw is higher than that of OGVR, due to the lack of constraints from flag information, a relatively stable and pleasing performance can also be found around the point when both λ and μ are equal to 100. For convenience, λ and μ were fixed as 100 in all our practical usage.

B. Suggestions on Version Selection of the Proposed Method

There remains a question about when to use OGVR and when to use OGVR_nw. Although OGVR is a more precise model and should be more suitable than OGVR_nw when the ancillary quality information is available, its performance may be worse than OGVR_nw in the case when untouched points in an NDVI time series are few. This is understandable as OGVR sets the weight of 0 for contaminated pixels, i.e., abandons the contaminated pixels in NDVI reconstruction. To illustrate the issue better, time-series simulation for benchmarking of reconstruction techniques (TISSBERTs) data [47], [48] was employed to conduct a supplementary experiment. TISSBERT is a global daily NDVI statistics data synthesized from 15 years AVHRR data, which provides a global test and reference dataset representative of real-world NDVI time series. TISSBERT data consist of one year of standardized parameters, including DOY cloudy average, DOY cloudy standard deviation, DOY clear average, DOY clear standard deviation, and DOY probability of clouds. In the experiment, daily synthetic cloud-free and cloudy time series were first generated. Concretely, cloud-free reference time series were determined as the DOY clear average time series, while cloudy test time series were synthesized from a mixture of cloudy time series and clear average time series according to the DOY probability of clouds. Then, both reference and

test time series were composited bimonthly using the MVC technique, with a total of 24 composites.

Before method comparison, the global statistics of the untouched composites in the TISSBERT simulated time series were made. The result is shown in Fig. 10 (sea composites were excluded). It can be found that more than 30% of the test time series have zero untouched composites, and nearly 60% of the test time series have untouched composites less than or equal to 7 (30% of all 24 composites). Two groups of methods using flag information or not were both tested. Since the methods using flag information abandon the contaminated points in NDVI reconstruction, it will be very difficult when untouched composites are few. The percentage distributions of NDVI reconstruction errors in TISSBERT data under MeanAE and MaxAE evaluations are shown in Fig. 11. The corresponding global distributions of the best performing method are also given in Fig. 11. As a supplement, the global ratios of the best performing method are listed in Table I. It is obvious that the class using flag information shows a significant overall decline in performance. With very limited information of untouched composites, SG, WS, MWHA, and OGVR are much easier to produce higher errors than RMMEH, IDR, and OGVR_nw methods (see Fig. 11). Although the performance of OGVR can be highly limited by the insufficient untouched points in TISSBERT simulated data, the proposed OGVR_nw that is independent of the flag information can always ensure a stable and high-quality NDVI reconstruction performance and achieves the best in this test. Therefore, combined with the previous experimental results, OGVR_nw is a more robust method that can always be relied on, while OGVR is recommended when flag information is available and the noise ratio is not too high. Alternatively, we can always choose OGVR but manually change its weight (from 0 to 1) for noisy pixels when processing seriously contaminated time series.

C. Running Time

For the proposed NDVI reconstruction method, the ℓ^p-norm ($p \in (0,\ 2]$) minimization in model (7) can be converted into a much simpler weighted least-squares minimization problem through the IRLS technique. As a result, both OGVR and OGVR_nw can efficiently estimate the latent ideal NDVI time series under the variational framework. In addition, since the regularization parameters in the proposed method are relatively insensitive, computation time is further saved by using the fixed parameters in real implementations. To speed up the calculation, all the experiments were conducted through the same parallel computation strategy in MATLAB on a desktop personal computer (CPU: eight processors with a frequency of 3.0 GHz and RAM: 48 GB). Based on 10 000 test points, the running times of all the test methods were recorded for cases of processing NDVI time series with different lengths. The average times calculated from the five rounds of repeated tests are given and compared in Table II. It should be noted that the running times of the SG and WS methods include the time taken to search for the best parameters, in accordance with their original works. According to the results listed in Table II, although the proposed OGVR and OGVR_nw, as global optimization methods, are not the fastest, their computation efficiency is still comparable to that of the highly efficient methods such as IDR and SG. Moreover, the running times of both versions of the proposed method grow slowly even when the data length increases significantly, indicating that the proposed method has good feasibility and applicability.

V. Conclusion

In this article, we have proposed a novel OGVR method to effectively and efficiently estimate high-quality NDVI time series via taking the properties of both the gradual vegetation change and negatively biased atmospheric contamination into consideration. To adapt to different processing cases, there are two versions of the proposed method, for data with and without ancillary flag information. The experiments undertaken on simulated and real NDVI time series at a global scale demonstrated that the proposed method has obvious advantages over the comparison methods in the following three aspects. First, the OGVR method is able to obtain effective and stable results for seriously contaminated NDVI time series, even if continuous missing data problems exist. Second, OGVR is convenient to use, without worrying about any tedious preprocessing/postprocessing steps or complex parameter settings. Third, OGVR possesses good universality and computing performance for different input conditions. The proposed method is not only suitable for use with GIMMS data, but it can also be easily applied to various NDVI products and self-calculated NDVI or EVI time series. These advantages of the proposed method will be conducive to research into vegetation monitoring, drought analysis, land-cover change detection, and other related topics.

Acknowledgment

The authors would like to thank Julien and Sobrino [47], [48] for making their data publicly available and also the anonymous reviewers for providing their constructive comments.

References

[1] C. Atzberger and P. H. C. Eilers, "A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America," *Int. J. Digit. Earth*, vol. 4, no. 5, pp. 365–386, Sep. 2011.

[2] 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. 2009.

[3] L. Zeng, B. D. Wardlow, D. Xiang, S. Hu, and D. Li, "A review of vegetation phenological metrics extraction using time-series multispectral satellite data," *Remote Sens. Environ.*, vol. 237, Feb. 2020, Art. no. 111511.

[4] R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy, "Land-cover change detection using multi-temporal MODIS NDVI data," *Remote Sens. Environ.*, vol. 105, no. 2, pp. 142–154, 2006.

[5] M. A. Friedl *et al.*, "Global land cover mapping from MODIS: Algorithms and early results," *Remote Sens. Environ.*, vol. 83, nos. 1–2, pp. 287–302, 2002.

[6] J. Zhou *et al.*, "Sensitivity of six typical spatiotemporal fusion methods to different influential factors: A comparative study for a normalized difference vegetation index time series reconstruction," *Remote Sens. Environ.*, vol. 252, Jan. 2021, Art. no. 112130.

[7] J. V. Keyserlingk, M. de Hoop, A. G. Mayor, S. C. Dekker, M. Rietkerk, and S. Foerster, "Resilience of vegetation to drought: Studying the effect of grazing in a Mediterranean rangeland using satellite time series," *Remote Sens. Environ.*, vol. 255, Mar. 2021, Art. no. 112270.

[8] 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, 2010.

[9] D. Kong, Q. Zhang, V. P. Singh, and P. Shi, "Seasonal vegetation response to climate change in the Northern Hemisphere (1982–2013)," *Global Planet. Change*, vol. 148, pp. 1–8, Jan. 2016.

[10] 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, 2003.

[11] I. H. Myers-Smith *et al.*, "Complexity revealed in the greening of the Arctic," *Nature Climate Change*, vol. 10, no. 2, pp. 106–117, 2020.

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

[13] P. S. A. 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, 2006.

[14] 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.

[15] J. M. Chen, F. Deng, and M. Chen, "Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2230–2238, Aug. 2006.

[16] P. J. 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, 1994.

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

[18] J. Zhou, L. Jia, and M. Menenti, "Reconstruction of global MODIS NDVI time series: Performance of harmonic ANalysis of time series (HANTS)," *Remote Sens. Environ.*, vol. 163, pp. 217–228, Jun. 2015.

[19] G. Yang, H. Shen, L. Zhang, Z. He, and X. Li, "A moving weighted harmonic analysis method for reconstructing high-quality SPOT VEGETATION NDVI time-series data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6008–6021, Nov. 2015.

[20] 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.

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

[22] 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, nos. 3–4, pp. 332–344, Jun. 2004.

[23] E. T. Whittaker, "On a new method of graduation," *Proc. Edinburgh Math. Soc.*, vol. 41, pp. 63–75, Feb. 1923.

[24] D. Kong, Y. Zhang, X. Gu, and D. Wang, "A robust method for reconstructing global MODIS EVI time series on the Google Earth engine," *ISPRS J. Photogramm. Remote Sens.*, vol. 155, pp. 13–24, Sep. 2019.

[25] 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.

[26] 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.

[27] 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, 1992.

[28] Y. Qiu, J. Zhou, J. Chen, and X. Chen, "Spatiotemporal fusion method to simultaneously generate full-length normalized difference vegetation index time series (SSFIT)," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 100, Aug. 2021, Art. no. 102333.

[29] J. Gu, X. Li, C. Huang, and G. S. Okin, "A simplified data assimilation method for reconstructing time-series MODIS NDVI data," *Adv. Space Res.*, vol. 44, no. 4, pp. 501–509, Aug. 2009.

[30] B. Qiu, M. Feng, and Z. Tang, "A simple smoother based on continuous wavelet transform: Comparative evaluation based on the fidelity, smoothness and efficiency in phenological estimation," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 47, pp. 91–101, May 2016.

[31] C. Ruyin *et al.*, "A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky–Golay filter," *Remote Sens. Environ.*, vol. 217, pp. 244–257, Nov. 2018.

[32] A. F. Militino, M. D. Ugarte, U. Perez-Goya, and M. G. Genton, "Interpolation of the mean anomalies for cloud filling in land surface temperature and normalized difference vegetation index," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6068–6078, Aug. 2019.

[33] R. J. Hodrick and E. C. Prescott, "Postwar U.S. Business cycles: An empirical investigation," *J. Money Credit Banking*, vol. 29, no. 1, pp. 1–16, Feb. 1997.

[34] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "ℓ_1 trend filtering," *SIAM Rev.*, vol. 51, no. 2, pp. 339–360, 2009.

[35] M. O. Ravn and H. Uhlig, "On adjusting the hodrick-prescott filter for the frequency of observations," *Rev. Econ. Statist.*, vol. 84, no. 2, pp. 371–376, 2002.

[36] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. Conf. Artif. Intell. Statist.*, 2011, pp. 315–323.

[37] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 807–814.

[38] H. L. Taylor, S. C. Bank, and J. F. McCoy, "Deconvolution with the ℓ_1 norm," *Geophys.*, vol. 44, no. 1, pp. 39–52, 1979.

[39] Y. Shao, R. S. Lunetta, B. Wheeler, J. S. Iiames, and J. B. Campbell, "An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data," *Remote Sens. Environ.*, vol. 174, pp. 258–265, Mar. 2016.

[40] R. Liu, R. Shang, Y. Liu, and X. Lu, "Global evaluation of gap filling approaches for seasonal NDVI with considering vegetation growth trajectory, protection of key point, noise resistance and curve stability," *Remote Sens. Environ.*, vol. 189, pp. 164–179, Feb. 2017.

[41] P. Jönsson and L. Eklundh, "TIMESAT—A program for analyzing time-series of satellite sensor data," *Comput. Geosci.*, vol. 30, no. 8, pp. 833–845, 2004.

[42] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least squares," *Commun. Stat.-Theory Methods*, vol. 6, no. 9, pp. 813–827, 1977.

[43] P. Rodríguez and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 322–332, Feb. 2009.

[44] X. Liu, H. Shen, Q. Yuan, X. Lu, and C. Zhou, "A universal destriping framework combining 1-D and 2-D variational optimization methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 808–822, Feb. 2018.

[45] M. K. P. Ng, R. H. Chan, and W.-C. Tang, "A fast algorithm for deblurring models with Neumann boundary conditions," *SIAM J. Sci. Comput.*, vol. 21, no. 3, pp. 851–866, 1999.

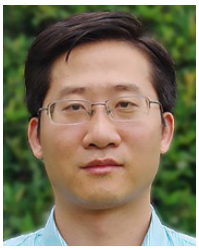
[46] W. Ye, A. I. J. M. van Dijk, A. Huete, and M. Yebra, "Global trends in vegetation seasonality in the GIMMS NDVI3g and their robustness," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 94, Feb. 2021, Art. no. 102238.

[47] Y. Julien and J. A. Sobrino, "TISSBERT: A benchmark for the validation and comparison of NDVI time series reconstruction methods," *Revista Teledetección*, vol. 51, pp. 19–31, Jun. 2018.

[48] Y. Julien and J. A. Sobrino, "Optimizing and comparing gap-filling techniques using simulated NDVI time series from remotely sensed global data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 76, pp. 93–111, Apr. 2019.

Xinxin Liu (Member, IEEE) received the B.S. degree in geographic information system and the Ph.D. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2013 and 2018, respectively.

In July 2018, she joined the College of Electrical and Information Engineering, Hunan University, Changsha, China, where she is currently an Assistant Professor. Her research interests include image quality improvement, remote sensing image processing, and remote sensing mapping and application.



Huanfeng Shen (Senior Member, IEEE) 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 2007, he joined the School of Resource and Environmental Sciences (SRES), Wuhan University, where he is currently a Luojia Distinguished Professor and the Associate Dean of SRES. He was or is the Principal Investigator of two projects supported by the National Key Research and Development Program of China and six projects supported by the National Natural Science Foundation of China. He has authored or coauthored more than 100 research articles in peer-reviewed international journals. His research interests include remote sensing image processing, multisource data fusion, and intelligent environmental sensing.

Dr. Shen is a Council Member of the China Association of Remote Sensing Application, an Education Committee Member of the Chinese Society for Geodesy Photogrammetry and Cartography, and a Theory Committee Member of the Chinese Society for Geospatial Information Society. He is currently a member of the Editorial Board of *Journal of Applied Remote Sensing* and *Geography and Geo-Information Science*.

Qiangqiang Yuan (Member, IEEE) 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 2006 and 2012, respectively.

In 2012, he joined the School of Geodesy and Geomatics, Wuhan University, where he is currently a Professor. He has published more than 90 research articles, including more than 70 peer-reviewed articles in international journals, such as the *Remote Sensing of Environment*, *ISPRS Journal of Photogrammetry and Remote Sensing*, IEEE TRANSACTIONS IMAGE PROCESSING, and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. His research interests include image reconstruction, remote sensing image processing and application, and data fusion.

Dr. Yuan was a recipient of the Top-Ten Academic Star of Wuhan University in 2011, the Youth Talent Support Program of China in 2019, and the recognition of Best Reviewers of the IEEE GRSL in 2019. In 2014, he received the Hong Kong Scholar Award from the Society of Hong Kong Scholars and the China National Postdoctoral Council. He is an Associate Editor of five international journals and has frequently served as a Referee for more than 40 international journals for *Remote Sensing* and *Image Processing*.

Xiliang Lu received the B.S. degree in computational mathematics from Peking University, Beijing, China, in 2000, and the Ph.D. degree in computational mathematics from the National University of Singapore, Singapore, in 2006.

He is currently a Professor with the School of Mathematics and Statistics, Wuhan University, Wuhan, China. His research interest includes computational inverse problems and their applications.

Shutao Li (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively.

In 2001, he joined the College of Electrical and Information Engineering, Hunan University. From May 2001 to October 2001, he was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Hong Kong. From November 2002 to November 2003, he was a Post-Doctoral Fellow with the Royal Holloway College, University of London, London, U.K. From April 2005 to June 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 200 refereed articles. His research interests include image processing, pattern recognition, and artificial intelligence.

Dr. Li is a member of the Editorial Board of the *Information Fusion* and the *Sensing and Imaging*. He was a recipient of the two 2nd-Grade State Scientific and Technological Progress Awards of China in 2004 and 2006. He is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.