



Reconstructed NDVI and EVI datasets in China (ReVIChina) generated by a spatial-interannual reconstruction method

Rui Yao, Yongjun Zhang, Lunche Wang, Jiayi Li & Qiquan Yang

To cite this article: Rui Yao, Yongjun Zhang, Lunche Wang, Jiayi Li & Qiquan Yang (2023) Reconstructed NDVI and EVI datasets in China (ReVIChina) generated by a spatial-interannual reconstruction method, International Journal of Digital Earth, 16:2, 4749-4768, DOI: [10.1080/17538947.2023.2283492](https://doi.org/10.1080/17538947.2023.2283492)

To link to this article: <https://doi.org/10.1080/17538947.2023.2283492>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 15 Nov 2023.



Submit your article to this journal [↗](#)



Article views: 621



View related articles [↗](#)



View Crossmark data [↗](#)



Reconstructed NDVI and EVI datasets in China (ReVICHina) generated by a spatial-interannual reconstruction method

Rui Yao^a, Yongjun Zhang^a, Lunche Wang^{b,c}, Jiayi Li^a and Qiquan Yang^{d,e}

^aSchool of Remote Sensing and Information Engineering, Wuhan University, Wuhan, People's Republic of China; ^bKey Laboratory of Regional Ecology and Environmental Change, School of Geography and Information Engineering, China University of Geosciences, Wuhan, People's Republic of China; ^cHubei Key Laboratory of Critical Zone Evolution, School of Geography and Information Engineering, China University of Geosciences, Wuhan, People's Republic of China; ^dCollege of Surveying and Geo-Informatics, Tongji University, Shanghai, People's Republic of China; ^eState Key Laboratory of Lunar and Planetary Sciences, Macau University of Science and Technology, Macau, People's Republic of China

ABSTRACT

Remote sensing-based vegetation index (VI) data are significantly impacted by cloud contamination. Spatiotemporal reconstruction methods demonstrate higher accuracy than temporal reconstruction methods. However, the computing time and random access memory (RAM) consumption of these spatiotemporal reconstruction methods for large-scale reconstruction remains unclear. In this study, a method called spatial-interannual reconstruction (SIR) was proposed to reconstruct cloud-contaminated pixels in MODIS normalized difference VI (NDVI) and enhanced VI (EVI) data. SIR has four major advantages: (1) High accuracy. The average mean absolute error of SIR was 0.0338, which was 20.2% and 23.4% lower than that of two state-of-the-art spatiotemporal reconstruction methods (i.e. interpolation of the mean anomalies (IMA) and Gapfill). (2) High computing speed. The average computing time of SIR was 99.7% and 98.8% lower than IMA and Gapfill, respectively. (3) Low RAM consumption. (4) Simultaneous reconstruction of all invalid values. Reconstructed 250 m spatial resolution and 16-day composite NDVI and EVI datasets in China from 2000 to 2022 (written as ReVICHina) were developed based on the SIR method and MODIS MOD13Q1 data. Spatiotemporal analyses revealed that the reconstructed datasets were more reliable than the original product and a similar dataset.

ARTICLE HISTORY

Received 21 September 2023
Accepted 8 November 2023

KEYWORDS

Remote sensing;
reconstruction; vegetation;
NDVI; EVI

1. Introduction

Vegetation plays a critical role in the Earth's system. Existing research indicates a global greening trend in vegetation, which profoundly impacts vegetation productivity and the carbon cycle, water cycle and distribution, climate change and surface energy balance (Chen et al. 2019a; 2019b; Forzieri et al. 2020; Piao et al. 2019; Zeng et al. 2018). Remote sensing is the most effective method for monitoring vegetation dynamics. Some remote sensing vegetation index (VI) products, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) normalized difference VI (NDVI)

CONTACT Rui Yao ✉ yaorui123@cug.edu.cn 📧 School of Remote Sensing and Information Engineering, Wuhan University, No.299 Bayi Road, Wuchang District, Wuhan City, Wuhan 430079, People's Republic of China

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and enhanced VI (EVI), have been widely utilized in vegetation dynamics and related research (Cazorla et al. 2023; Dalagnol et al. 2023; Ims et al. 2019).

However, the quality of satellite-derived VI data is significantly affected by cloud contamination. Despite the use of the maximum value composite method in many VI products (e.g. MODIS and Global Inventory Monitoring and Modeling Studies (GIMMS)) to mitigate the impact of cloud contamination, residual effects persist widely, particularly in areas with frequent cloud cover (Li et al. 2021; Liu et al. 2017). For example, Liu et al. (2017) discovered a significant number of abnormally low values in MODIS NDVI products in humid areas due to cloud contamination.

To address this issue, various methods have been developed to reconstruct cloud-contaminated VI. Early methods primarily utilized temporal information (i.e. information from neighbouring dates), such as the Savitzky–Golay filter (Chen et al. 2004), the Double Logistic function (Beck et al. 2006), the Fourier-based method (Jakubauskas, Legates, and Kastens 2001), the Asymmetric Gaussian function (Jönsson and Eklundh 2002) and others (Kong et al. 2019; Yang et al. 2015). These methods are relatively simple and have been widely employed in vegetation dynamics and related research (Choler et al. 2021; Hilker et al. 2014; Tian et al. 2021). More recently, spatiotemporal reconstruction methods that incorporate both spatial and temporal information (i.e. information from neighbouring pixels and dates) have been widely developed (Cao et al. 2018; Chen et al. 2021; Chu et al. 2021; Gerber et al. 2018; Huang et al. 2021; Li et al. 2023; Militino et al. 2019; Yang et al. 2022). These methods have demonstrated higher accuracy compared to methods that solely rely on temporal information. For example, Cao et al. (2018) proposed a method that integrates spatiotemporal information with the Savitzky–Golay filter, achieving higher accuracy than the Savitzky–Golay filter, the Asymmetric Gaussian function, the Double Logistic function and the Fourier-based method. Chu et al. (2021) introduced a method that utilizes spatial, intra-annual and interannual information to reconstruct NDVI. The results showed that this method performs better than four temporal filtering methods and the method developed by Cao et al. (2018). Li et al. (2023) developed a spatiotemporal prefill method with harmonic analysis of time series (ST-HANTS) to reconstruct NDVI. ST-HANTS significantly outperformed the Savitzky–Golay filter, wavelet transform, HANTS and data assimilation.

Despite significant progress being made in reconstructing cloud-contaminated VI, two limitations still exist. First, the computing time and random access memory (RAM) consumption of these spatiotemporal reconstruction methods for large-scale reconstruction remains unclear. Previous studies have not assessed the computing time and RAM consumption (Cao et al. 2018; Chen et al. 2021; Gerber et al. 2018; Huang et al. 2021; Li et al. 2023) or assessed the computing time in small regions (Chu et al. 2021; Militino et al. 2019; Yang et al. 2022). Second, there is a lack of large-scale, long time series reconstructed VI datasets produced using spatiotemporal reconstruction methods with high-precision. Gao et al. (2023) developed a monthly reconstructed NDVI dataset with 250 m resolution in China. However, Gao et al. (2023) first used a spatial reconstruction method to reconstruct NDVI and then used the Savitzky–Golay filter to smooth the reconstructed NDVI. Both the spatial reconstruction method and the Savitzky–Golay filter have been proven to have lower accuracy than spatiotemporal reconstruction methods (Cao et al. 2018; Gerber et al. 2018; Li et al. 2023; Liu et al. 2019; Militino et al. 2019). Using low-quality datasets produced by low-precision methods may result in unreliable results.

Considering the aforementioned limitations, this study proposes a reconstruction method that simultaneously has high accuracy, high computing speed and low RAM consumption. Subsequently, 250 m spatial resolution and 16-day composite reconstructed NDVI and EVI datasets for China from 2000 to 2022 were developed using this method and MODIS MOD13Q1 data. In the following text, Section 2 introduces the study area and data. Section 3 details the preprocessing and reconstruction methods. Section 4 presents the results, including the accuracy and computing costs of the proposed method, as well as the developed reconstructed NDVI and EVI datasets.

Section 5 discusses the proposed method and developed datasets. Finally, Section 6 summarizes this study.

2. Study area and data

China, the largest developing country in the world (Figure 1), ranks second in terms of population and third in terms of land area. Across China, from the southeast to the northwest, a general decrease in NDVI can be observed (Figure 1(a)). This is accompanied by a gradual transformation

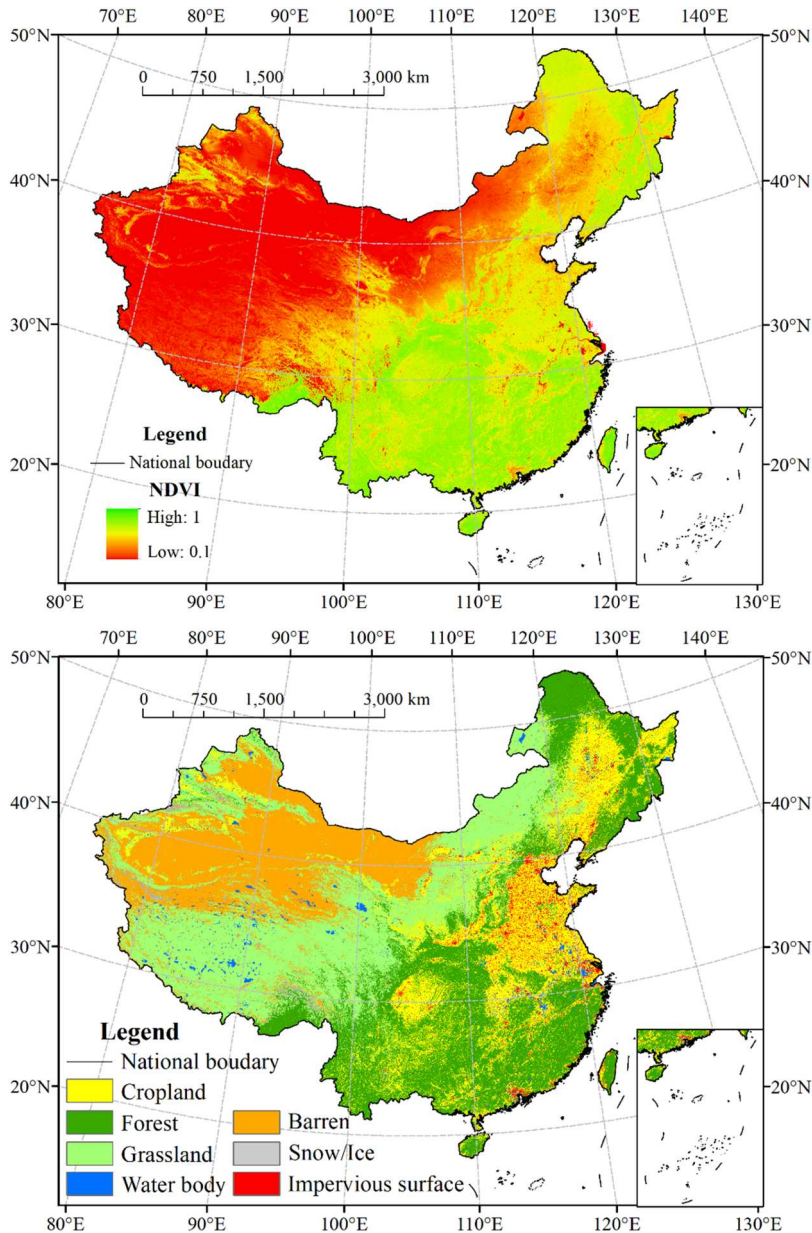


Figure 1. The study area of this study. The background maps display the 23-year average growing season reconstructed normalised difference vegetation index (NDVI) and the China land cover dataset (CLCD) in 2021 (Yang and Huang 2021). Note that the lower limit of the NDVI was set to 0.1 in the postprocessing steps.

of climate types from hot and humid to cold and dry, and a shift in dominant land cover types from forests to deserts. China is a hot research area for vegetation dynamics (Song, Feng, and Wang 2022; Yao et al. 2019; Zheng et al. 2019). Over the past few decades, China has exhibited the most prominent greening trend globally (Chen et al. 2019a; Piao et al. 2019; Zhang et al. 2017).

MODIS MOD13Q1 NDVI and EVI data (version 6.1, 250 m spatial resolution and 16-day composite) from February 2000 to December 2022 were used in this study. The NDVI is one of the most widely used vegetation indices and is sensitive to chlorophyll variations. The EVI was designed to enhance sensitivity in densely vegetated areas and mitigate atmospheric influences (Huete et al. 2002; Xue and Su 2017). The MOD13Q1 product incorporates quality assurance flags that provide information about the quality of each pixel (Huete et al. 2002).

3. Methods

3.1. Preprocessing

Pixels with quality other than 'good' in the MOD13Q1 NDVI and EVI data were set as invalid according to previous studies (Cao et al. 2018; Gerber et al. 2018). However, this filtering approach may unintentionally exclude some values that appear to be correct, and lead to an increase in computing time. To address this issue, the following preprocessing steps were employed to retain more pixels.

- (1) Pixels with a multiyear average growing season (from Apr. to Oct.) NDVI lower than 0.1 were considered non-vegetated areas, primarily consisting of deserts and water bodies. The NDVI values of these pixels were directly set to 0.1, irrespective of their quality. The threshold of 0.1 aligns with previous studies (Liu et al. 2020; Zhao, Dai, and Dong 2018). The threshold for EVI was set to 0.067, which corresponds to the percentile of 0.1 in NDVI.
- (2) Pixels with a multiyear average NDVI (EVI) on the same date lower than 0.1 (0.067) (mainly representing withered vegetation in winter) were directly set as 0.1 (0.067), regardless of their quality.
- (3) In the MODIS VI products, there are pixels labelled as other quality, but their values appear to be correct (Figure 2). Therefore, the pixels labelled as other quality but with a VI higher than 80% of the multiyear average VI on the same date, were considered good quality and retained.
- (4) Pixels with an NDVI (EVI) lower than 0.1 (0.067) were directly set as 0.1 (0.067). The reason for using this step is to prevent the impact of extremely low values (e.g. negative values in some water bodies) on the reconstruction of the VI.

3.2. Reconstruction of the VI

In this study, a method was proposed to reconstruct cloud-contaminated VI. Because this method uses spatial and interannual information, it is called spatial-interannual reconstruction (SIR). A schematic of the SIR is shown in Figure 3, consisting of the following three steps.

- (1) Calculating a multiyear average image. To reconstruct a target image, a multiyear average image on the same date was calculated. Only VI data with good quality or those retained in the preprocessing steps were considered for the calculation of the multiyear average image. If all the VI data were invalid (i.e. labelled as other quality and not retained in the preprocessing steps), the average of all the invalid VI data was used. Consequently, the multiyear average image did not contain any missing or invalid values.
- (2) Defining a spatial window. For a target pixel in the target image to be reconstructed, a spatial window with an initial size of 11×11 pixels was established. If there was more than one valid

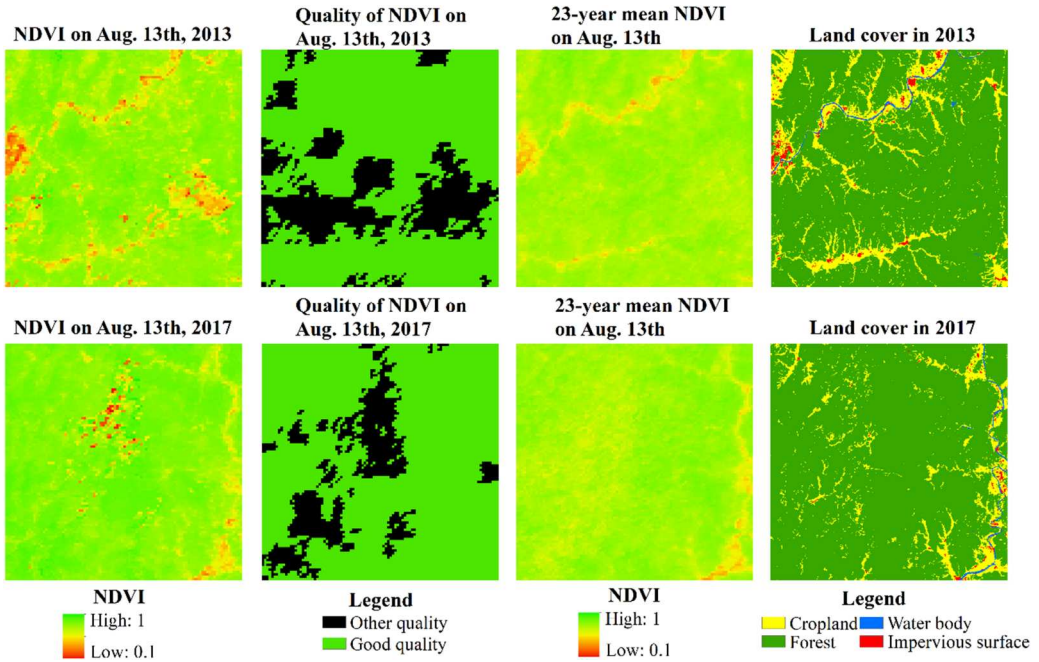


Figure 2. The NDVI, the quality of the NDVI, the 23-year mean NDVI and the land cover in an area of the Wuyi Mountains (WYM). Note that the lower limit of the NDVI was set to 0.1 in the preprocessing steps.

VI data point within the spatial window in the target image, the VI data from both the target and multiyear average images within the spatial window were extracted to reconstruct the target pixel. Otherwise, the size of the spatial window was increased from 11×11 – 31×31 , 111×111 , 431×431 , and so on until the aforementioned condition was met. The interval of size of the latter special window was set as four times that of the previous special window, to reduce the computing time needed for filling large gaps.

- (3) Reconstructing the target pixel. The VI of the target pixel was reconstructed using Equation (1):

$$VI_{x,t,1} = VI_{x,m} + VI_{y,t} - VI_{y,m} \quad (1)$$

where $VI_{x,t,1}$ is the reconstructed VI of the target pixel. x and y are the target pixel and a pixel in the spatial window, respectively. t and m represent the target image and multiyear average image, respectively. $VI_{x,m}$ and $VI_{y,m}$ are always valid, because the multiyear average image does not contain any invalid values. Therefore, if $VI_{y,t}$ is valid, $VI_{x,t,1}$ can be reconstructed. This equation is based on the hypothesis that the difference in VI between x and y in the target image is equivalent to the multiyear average image. Within the spatial window, each valid VI can be utilized to reconstruct a VI value for the target pixel. The VI of the target pixel was finally computed as the weighted average of all the reconstructed VI values. The weight was calculated as:

$$w_n = \frac{1}{D_{x,n}^2 \times (|VI_{x,m} - VI_{n,m}| + 1)} \quad (2)$$

where $D_{x,n}^2$ is the square of the distance between the target pixel x and pixel n . The power of $D_{x,n}$ was set as 2, to reduce the computing time. Distant pixels were given small weights, because closer pixels tend to be more similar. The formula on the right side of the multiplication sign in the denominator represents the difference in VI between the target pixel and pixel n

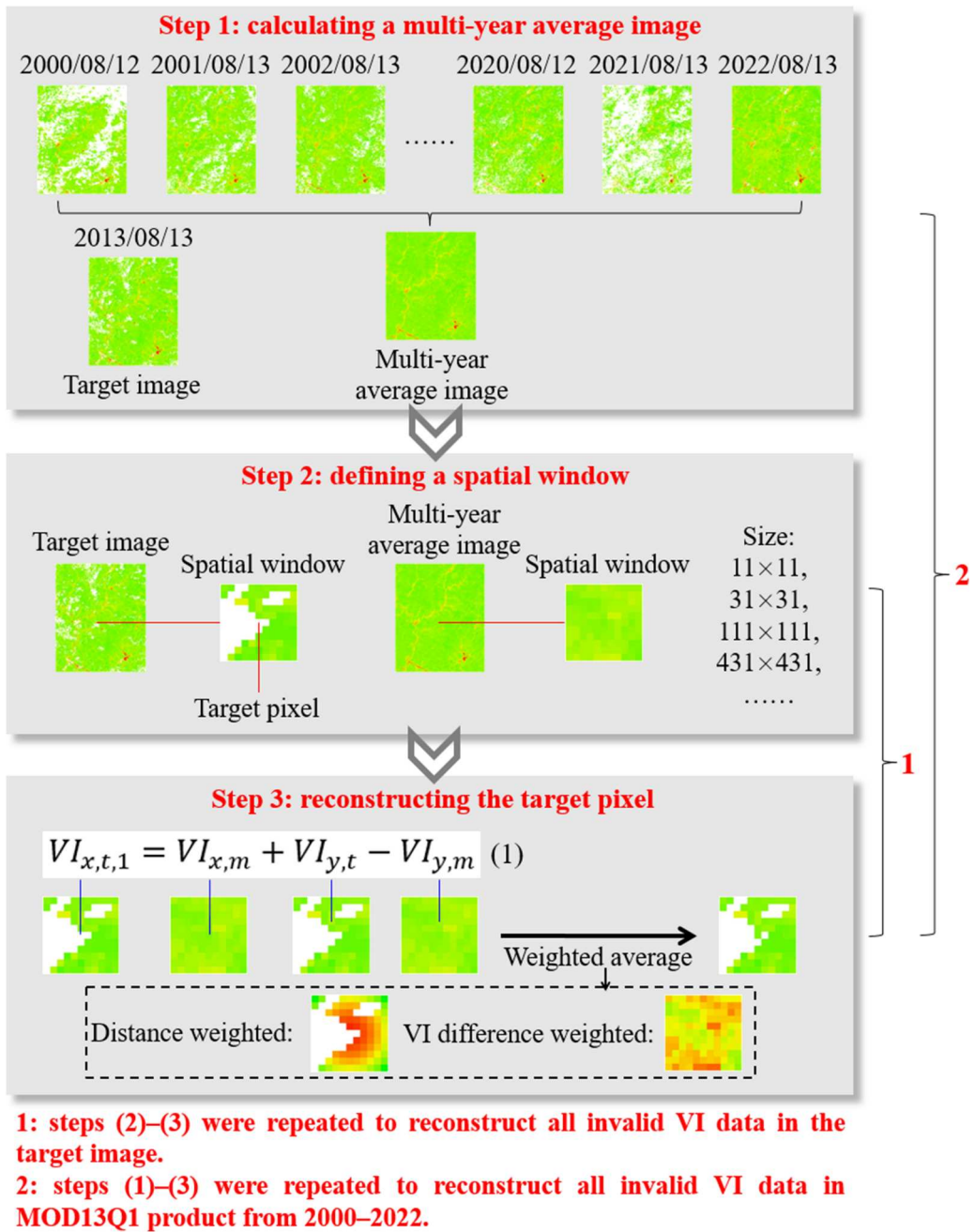


Figure 3. Schematic diagrams of the SIR method.

in the multiyear average image. This indicator is utilized because pixels with similar VI values generally correspond to similar land cover types. w_n represents the weight of pixel n .

Steps (2)–(3) were repeated to reconstruct all invalid VI data in the target image. Subsequently, steps (1)–(3) were repeated to reconstruct all invalid VI data in the MOD13Q1 product from 2000 to 2022. It is important to note that VI reconstruction may generate outliers, such as VI values exceeding 1. Therefore, NDVI and EVI values exceeding 1 were directly set to

1. NDVI (EVI) values below 0.1 (0.067) were directly set to 0.1 (0.067). To further refine the time series of the VI data, a Savitzky–Golay filter (Savitzky and Golay 1964) was applied for smoothing. The use of an upper envelope was not necessary, as low outliers were already reconstructed. The R code for the SIR method is available from the authors.

3.3. Accuracy validation

The accuracy of VI reconstruction was validated in the North China Plain (NCP), Qinghai-Tibet Plateau (QTP) and Wuyi Mountains (WYM) (Figure 4). Each area is 200×150 km (or each image had 800 rows and 600 columns). These three areas were chosen to represent cropland/impervious surface mosaics, grassland and forest, respectively. Artificial gaps were introduced into the images during the growing season of 2015. These gaps were created with dimensions of 30×30 and 120×120 pixels, representing small and large cloud-contaminated areas, respectively. Subsequently, the SIR method was employed to fill these artificially generated gaps. Finally, the reconstructed VI data were compared with the original VI data. The mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2) were utilized to quantify the accuracy.

The performance of the SIR was compared with interpolation of the mean anomalies (IMA) (Militino et al. 2019) and Gapfill (Gerber et al. 2018). These methods were selected because: (1) they are spatiotemporal reconstruction methods and have shown higher accuracy compared to other reconstruction methods (Gerber et al. 2018; Militino et al. 2019). (2) They can be easily implemented using R add-on packages (IMA: ‘RGISTool’; Gapfill: ‘gapfill’). The accuracy validation and comparison were conducted using the preprocessed images.

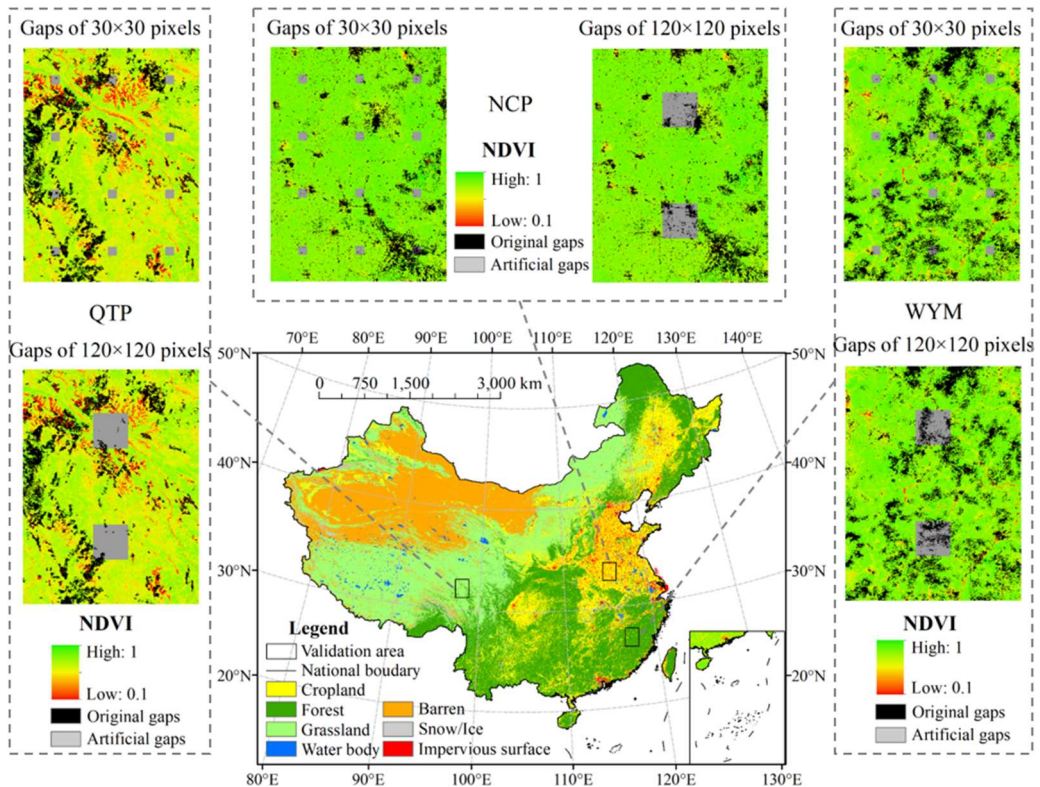


Figure 4. Schematic diagrams of the accuracy validation. NCP: North China Plain. QTP: Qinghai-Tibet Plateau.

3.4. Computational costs

The computing time of the three methods was compared in the WYM from 2 February 2003 to 10 June 2003 (a total of nine images). These specific images were chosen because the proportions of invalid values in them were evenly distributed between 0% and 100%. This selection allowed us to evaluate the performance of the different methods under varying proportions of invalid values. Subsequently, the computing time and RAM consumption of the SIR and Gapfill methods were tested for the entire region of China (the image has 16,179 rows and 19,381 columns) for 2015. The IMA method was not tested for the entire region of China because its RAM consumption approached the upper limit of the available RAM (128 gigabytes (GB)) on our computer when the image size exceeded 1100 rows and 1100 columns. These experiments were conducted using a computer equipped with an Intel Core i7-12700 Central Processing Unit (CPU) and 128 GB of RAM.

4. Results

4.1. Accuracy of the VI reconstruction

The SIR method significantly outperformed IMA and Gapfill (Figures 5 and 6). For all cases, the SIR method consistently exhibited lower MAE and RMSE values, as well as higher R^2 values compared

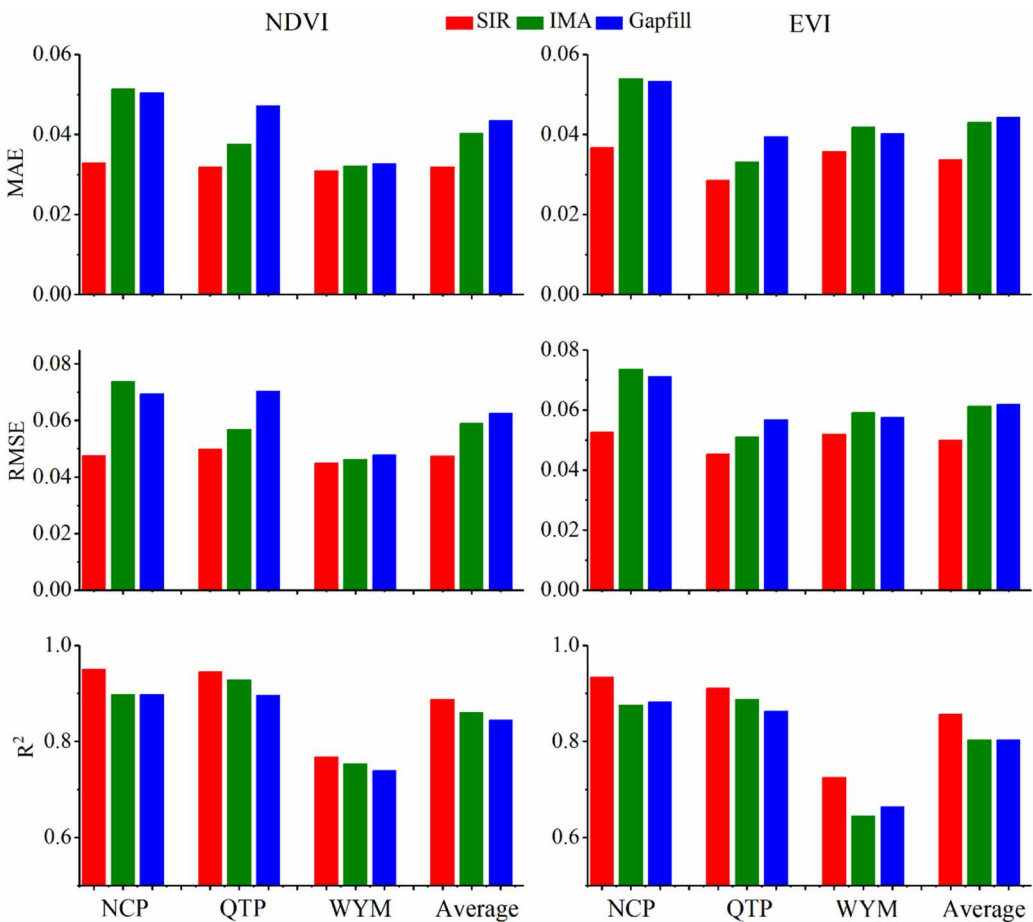


Figure 5. The accuracy of the three methods in filling the gaps of 30×30 pixels. IMA: Interpolation of the mean anomalies. MAE: mean absolute error. RMSE: root mean square error. R^2 : coefficient of determination.

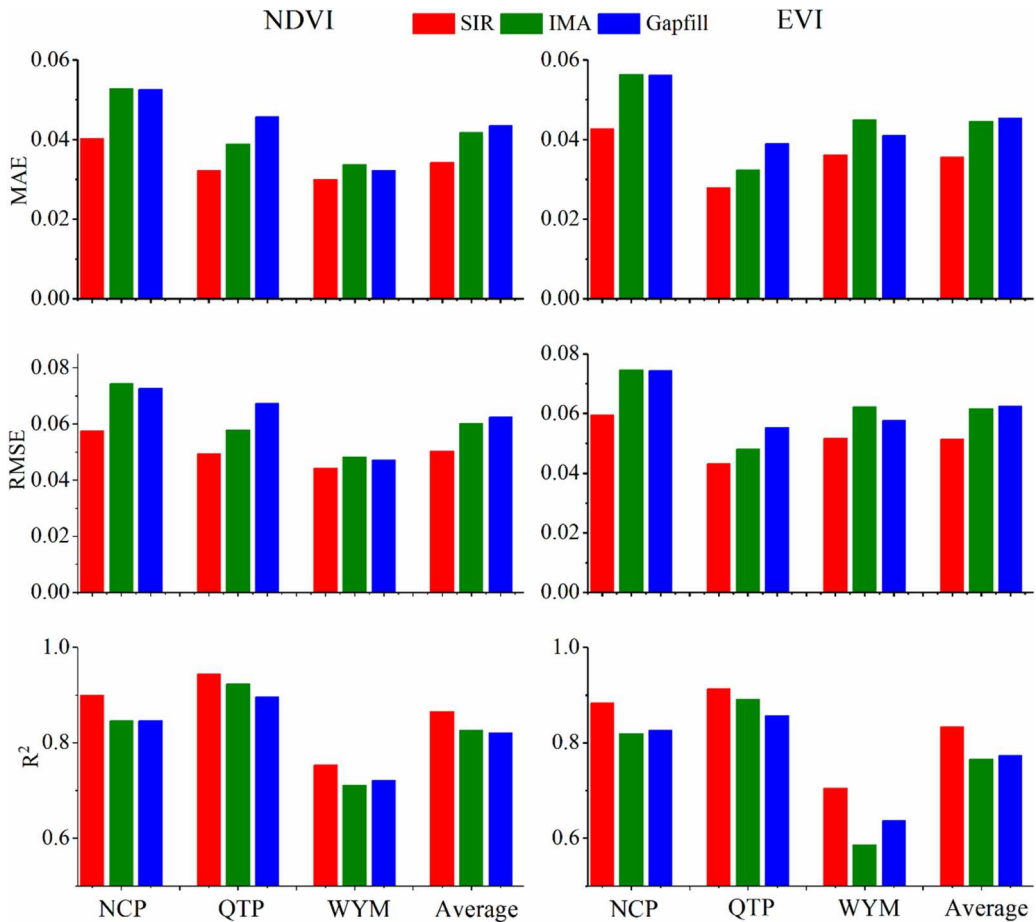


Figure 6. The accuracy of the three methods in filling the gaps of 120×120 pixels.

to the IMA and Gapfill methods. On average, the SIR method achieved an MAE of 0.0338, an RMSE of 0.0498, and an R^2 of 0.861. The average MAE of the SIR method was 20.2% and 23.4% lower than that of the IMA and Gapfill methods, respectively. The improvement in accuracy was particularly pronounced in the NCP region. To uncover the underlying reasons, we analyzed the intra-annual variations in the regional average NDVI and the MAE in the NCP region (Figure 7). The NDVI curves exhibited significant fluctuations, with two prominent peaks and valleys in each curve. This is because the dominant land cover type in the NCP region is cropland. Human activities, such as planting and harvesting, have a substantial impact on the NDVI values. The IMA and Gapfill methods, which rely on information from neighbouring dates, may introduce uncertainties. The MAEs of the IMA and Gapfill methods were relatively low on dates with drastic changes in NDVI. In contrast, the NDVI curves in different years exhibited similarity, and the NDVI values on the same dates were close. Consequently, utilizing interannual information proved to be more reliable than relying on intra-annual information in this particular case. The fluctuation of MAE in the SIR method was smaller than that of the IMA and Gapfill methods. In summary, the utilization of interannual information in the SIR method resulted in high accuracy within the NCP region.

The MAE and RMSE values for VI reconstruction were higher in the NCP than in the QTP and WYM (Figures 5 and 6). The average MAEs of the SIR method in the NCP, QTP and WYM were

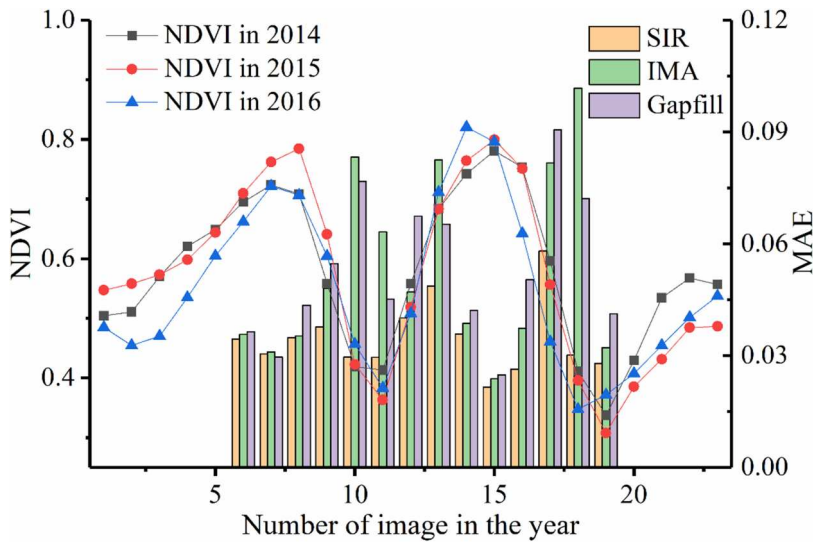


Figure 7. Intra-annual variations in the regional average NDVI from 2014 to 2016 and the MAE in 2015 in the NCP region.

0.0382, 0.0301 and 0.0332, respectively. This discrepancy can be attributed to the heterogeneity of the land cover types and the VI within the NCP region, making VI reconstruction more challenging. Additionally, the accuracy in filling the gaps of the 120×120 pixels was much higher than that for the gaps of the 30×30 pixels in the NCP region (Figures 5 and 6). For example, the average MAEs of the SIR method for filling the gaps of the 30×30 and 120×120 pixels in the NCP region were 0.0348 and 0.0415, respectively. This result is understandable, because distant pixels were used when filling the gaps of the 120×120 pixels. Distant pixels generally exhibit weaker correlations with the target pixel, thereby introducing higher errors in VI reconstruction, particularly in areas characterized by heterogeneous VIs and land cover types. In comparison, the average MAEs of the three methods for filling the gaps of the 30×30 pixels were similar to those for filling the gaps of the 120×120 pixels in the WYM and QTP regions. This can be attributed to the homogeneity of the VI and land cover type in these two regions.

4.2. Computing costs of the VI reconstruction

The SIR method exhibited a high computing speed (Table 1). On average, the computing time for the SIR method to reconstruct the images in the WYM was 13.89 s, which was 99.7% and 98.8% less than IMA and Gapfill, respectively. Most reconstructions using the SIR method were completed within 20 s. In comparison, Gapfill and IMA often took more than 5 min and 1 h to reconstruct images, respectively. In addition, when reconstructing images for all of China, the average computing time for the SIR method was 1433.22 s, with most reconstructions completed within 1 h (Table 2). In comparison, Gapfill often took more than 1 d to reconstruct images for the entire region of China. Due to the extensive computing time and RAM consumption, IMA was not tested for the entire region of China. It is important to note that the results were obtained using a single core of the CPU. Utilizing multiple cores in parallel could significantly increase the computing speed. For example, R software can support parallel computation of multiple windows, therefore, VI reconstruction can be easily run using multiple cores of the CPU in parallel. The reasons behind the high computing speed of the SIR method are analyzed in the discussion section.

Table 1. The computing time (unit: second) for the NDVI reconstruction and the proportion of valid values before and after the reconstruction in the WYM (800 rows and 600 columns).

Date	SIR			IMA			Gapfill		
	Proportion of valid values before reconstruction	Computing time	Proportion of valid values after reconstruction	Computing time	Proportion of valid values after reconstruction	Computing time	Proportion of valid values after reconstruction	Computing time	Proportion of valid values after reconstruction
2 February 2003	82.1%	5	100%	5484	99.4%	328	100%		
18 February 2003	92.9%	3	100%	8144	99.8%	119	100%		
6 March 2003	29.2%	18	100%	2238	97.4%	1518	100%		
22 March 2003	61.4%	19	100%	4332	98.4%	2426	100%		
7 April 2003	42.8%	16	100%	4201	95.4%	1499	100%		
23 April 2003	73.6%	6	100%	8960	97.8%	446	100%		
9 May 2003	35.2%	17	100%	4111	92.6%	1178	100%		
25 May 2003	57.6%	8	100%	8200	97.0%	532	100%		
10 June 2003	14.8%	33	100%	809	98.2%	2157	100%		
Average	54.4%	13.89	100%	5164.33	97.3%	1133.667	100%		

Table 2. The computing time (unit: seconds) and random access memory (RAM) consumption (unit: GB) of the SIR method for the NDVI reconstruction in all of China (16,179 rows and 19,381 columns) in 2015.

Date	SIR			Gapfill		
	Proportion of valid values before reconstruction	Computing time	RAM consumption	Proportion of valid values after reconstruction	Computing time	RAM consumption
1 January 2015	81.5%	1737	6.39	100%	*	33.87
17 January 2015	79.9%	1291	6.41	100%	*	33.67
2 February 2015	83.9%	1044	6.38	100%	*	33.58
18 February 2015	76.1%	2035	6.38	100%	*	33.78
6 March 2015	81.4%	2231	6.46	100%	*	33.64
22 March 2015	86.2%	1136	6.80	100%	*	33.71
7 April 2015	91.3%	590	6.54	100%	56017	33.36
23 April 2015	87.4%	850	6.37	100%	*	33.47
9 May 2015	89.6%	894	6.36	100%	*	33.40
25 May 2015	84.8%	1556	6.37	100%	*	33.53
10 June 2015	81.4%	1403	6.38	100%	*	33.67
26 June 2015	84.9%	1309	6.36	100%	*	33.53
12 July 2015	85.5%	967	6.46	100%	*	33.51
28 July 2015	90.2%	677	6.39	100%	59323	33.48
13 August 2015	86.7%	1472	6.38	100%	*	33.69
29 August 2015	84.6%	1422	6.36	100%	*	33.89
14 September 2015	87.4%	1326	6.38	100%	*	33.73
30 September 2015	95.4%	370	6.43	100%	32345	33.26
16 October 2015	92.6%	553	6.59	100%	60880	33.28
1 November 2015	84.3%	1275	6.51	100%	*	33.55
17 November 2015	76.9%	5407	6.44	100%	*	33.76
3 December 2015	83.2%	1768	6.53	100%	*	33.59
19 December 2015	78.9%	1651	6.38	100%	*	34.15
Average	85.0%	1433.22	6.44	100%	*	33.61

Note: * indicates that the computing time was more than 1 d.

The RAM consumption of the SIR method was low (Table 2). The average RAM consumption for reconstructing NDVI images for the entire region of China in 2015 using the SIR method was 6.44 GB, which was 80.8% lower than that of Gapfill (33.61 GB). The differences in RAM consumption for reconstructing different images were small for both the SIR and Gapfill methods. With its low RAM consumption, the SIR method can easily be run using multiple cores of CPU. For example, with our computer equipped with 12 cores of CPU and 128 GB RAM, the SIR method could be efficiently run using 9 cores of CPU in parallel, with the CPU utilization rate approaching 100% and the RAM utilization rate below 70%.

The SIR method simultaneously reconstructed all of the cloud-contaminated pixels, similar to Gapfill (Tables 1 and 2). In contrast, IMA and many previous methods are unable to reconstruct all missing values simultaneously (Cao et al. 2018; Militino et al. 2019; Yang et al. 2022), or they require numerous complex steps for complete reconstruction (Li et al. 2018; Yao et al. 2023c). This suggests that the SIR method is a convenient and efficient approach.

4.3. Reconstructed NDVI and EVI datasets

The original MODIS MOD13Q1 NDVI and EVI data were preprocessed using the preprocessing methods described in Section 3.1. Subsequently, 250 m spatial resolution and 16-day composite reconstructed NDVI and EVI datasets were developed for China (written as ReVICHina) from 2000 to 2022 using the SIR method and preprocessed images. The ReVICHina NDVI data were compared to the original MOD13Q1 NDVI products and a monthly reconstructed NDVI product in Gao et al. (2023). Because the reconstructed NDVI product in Gao et al. (2023) is monthly composite, the original and ReVICHina NDVI data were averaged into a month for comparison. The WYM region was selected for comparison, because: (1) it has a humid climate, and the NDVI is strongly affected by cloud contamination. (2) The dominant vegetation type in the WYM region is forests, which are less affected by human activities.

Figure 8 presents the comparisons of the spatial distributions of the three NDVI datasets in the WYM region in 2002. It can be clearly seen that the original NDVI data were significantly affected by cloud contamination. The NDVI significantly decreased from April to May. The effect of cloud contamination on the NDVI was not completely removed by Gao et al. (2023). In contrast, the spatial distributions of ReVICHina NDVI in different months were similar. Low NDVI values were generally observed in impervious surfaces and rivers. These results indicate that the impact of cloud contamination on VI has been effectively eliminated in ReVICHina.

Figure 9 illustrates the intra-annual comparisons of the three NDVI datasets in the WYM region in 2002, 2005, 2007 and 2015. It is evident that the original VI curves exhibited numerous abnormally low outliers caused by cloud contamination. The impact of the cloud contamination on the regional average NDVI exceeded 0.3. The effect of the cloud contamination on the NDVI was also evident in Gao et al. (2023). In comparison, the ReVICHina NDVI curves were relatively smooth, with few extreme fluctuations. Additionally, there were strong correlations between the ReVICHina NDVI data in different years. For example, the Pearson's correlation coefficient between the ReVICHina NDVI in 2002 and 2005 was 0.955. In contrast, the Pearson's correlation coefficients for the original NDVI and the NDVI in Gao et al. (2023) were 0.622 and 0.919, respectively. These findings indicate that the quality of the VI data was significantly improved by the SIR method.

Figure 10 shows the reconstructed 23-year average growing season VI and the difference between the original and reconstructed 23-year average growing season VI. In the forests in southern China, the original VI was significantly lower than the reconstructed VI, indicating that the cloud contamination had a substantial impact in these areas. This can be attributed to two factors: (1) Southern China has a humid climate. VI is more frequently affected by cloud contamination. (2) Forests generally have high

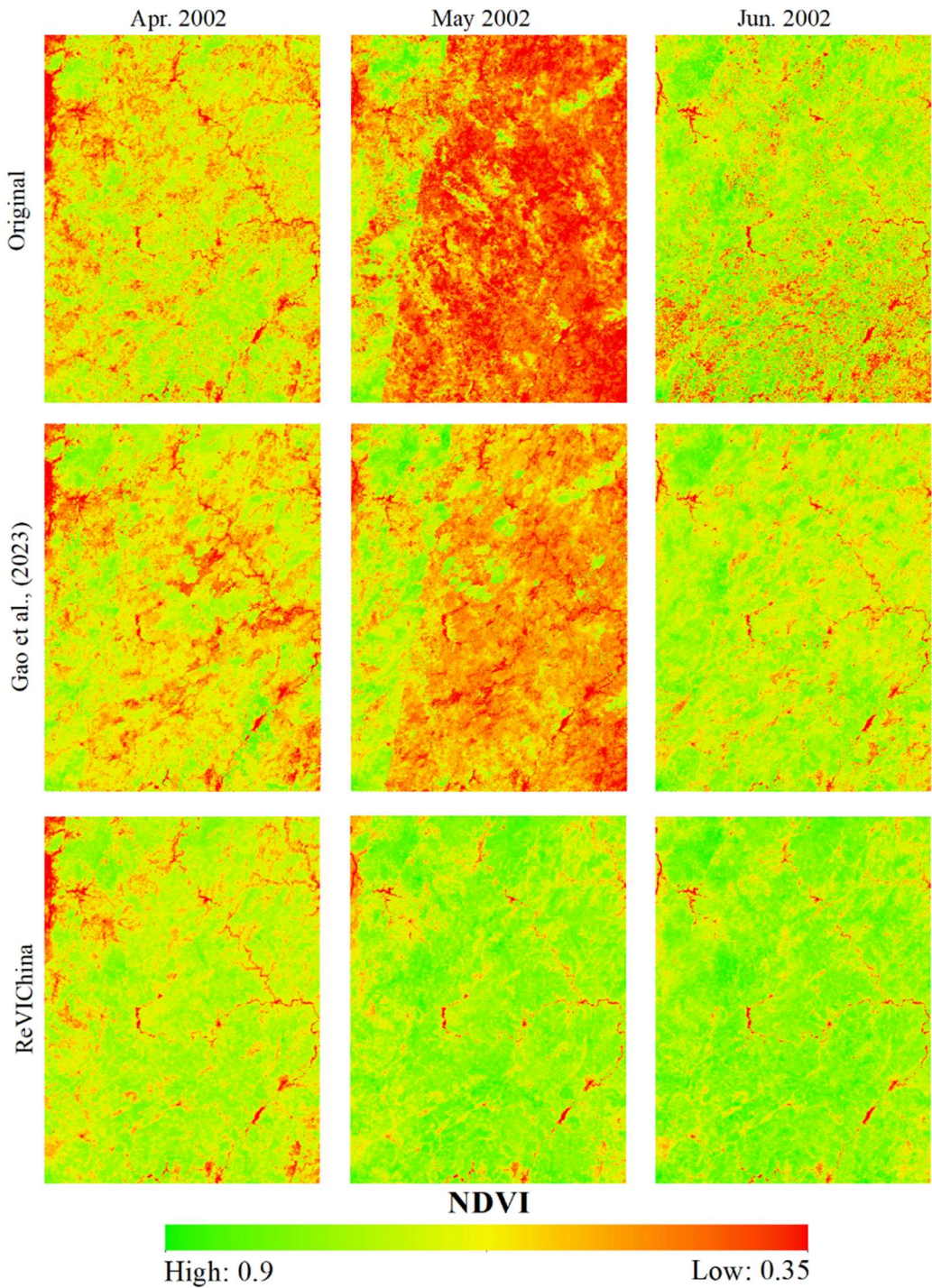


Figure 8. Comparisons of the spatial distributions of the original NDVI, the reconstructed NDVI in Gao et al. (2023) and the ReVI-China NDVI in the WYM.

VI, thus leading to a large discrepancy between the cloud-contaminated and reconstructed VI values. In contrast, the difference between the original and reconstructed VIs was minimal in northern China due to its arid climate and generally low VI values.

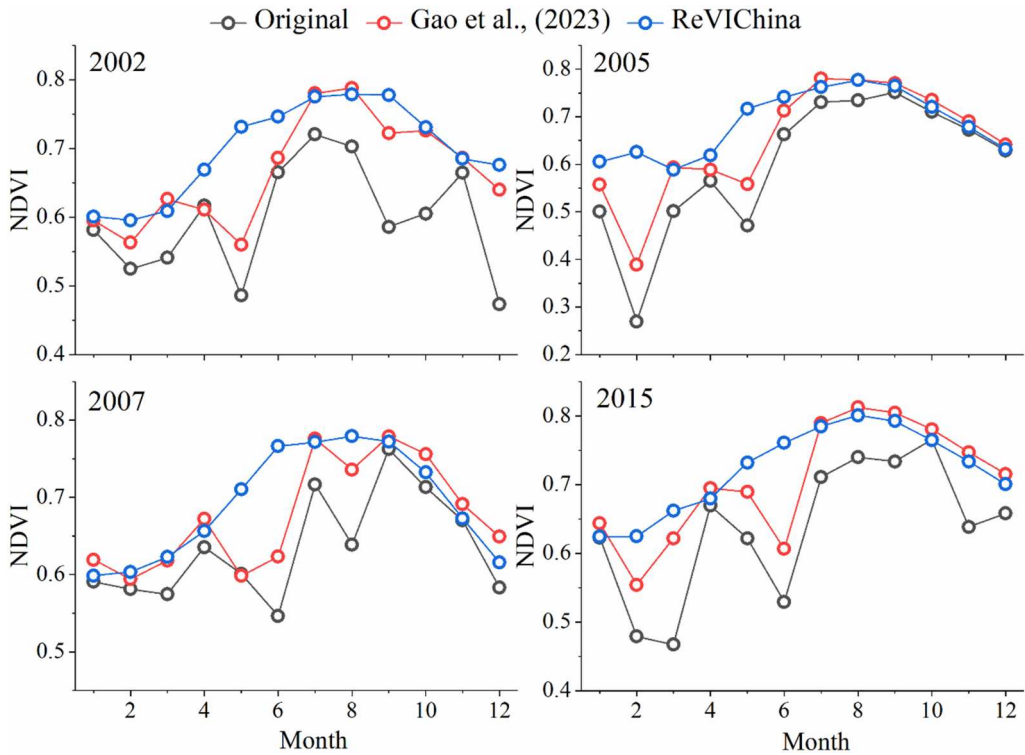


Figure 9. Comparisons in the intra-annual variations of the original NDVI, the reconstructed NDVI in Gao et al. (2023) and the ReViChina NDVI in the WYM.

5. Discussion

5.1. The SIR method

In this study, a spatiotemporal reconstruction method called SIR was proposed to reconstruct cloud-contaminated VI data. The SIR method was shown to have high accuracy, high computing speed, and low RAM consumption, and could simultaneously reconstruct all the cloud-contaminated pixels. These four major advantages of the SIR method are discussed in detail below.

First, the SIR method had high accuracy. One of the reasons for its high accuracy is the utilization of interannual information, which enhances the reconstruction accuracy in the NCP (as discussed in Section 4.1). Additionally, SIR employs a weighted average of all reconstructed VI values as the final reconstructed VI (as described in step 3 of Section 3.2). This approach reduces the impact of outliers and improves the accuracy of VI reconstruction.

Second, the SIR method demonstrated high computing speed due to its well-designed framework. The main reasons for the high computing speed of the SIR include the following: (1) the interval of the size of the spatial window increases significantly (as explained in step 2 in Section 3.2) instead of remaining constant as seen in methods such as Gapfill and some previous approaches (Gerber et al. 2018; Li et al. 2018; Zeng et al. 2015). This strategy reduces the computing time needed for filling large gaps. (2) SIR utilizes only one related image (the multiyear average image), which is significantly fewer compared to previous approaches (Gerber et al. 2018; Li et al. 2018; Sun et al. 2017; Yao et al. 2021). By reducing the number of related images, the computing time is decreased. (3) All the steps in the SIR method employ simple and fast mathematical operations, such as addition, subtraction, division, and averaging. SIR avoids complex and time-consuming operations. The lengthy computation time of the IMA method is attributed to its use of a thin plate spline method for interpolation,

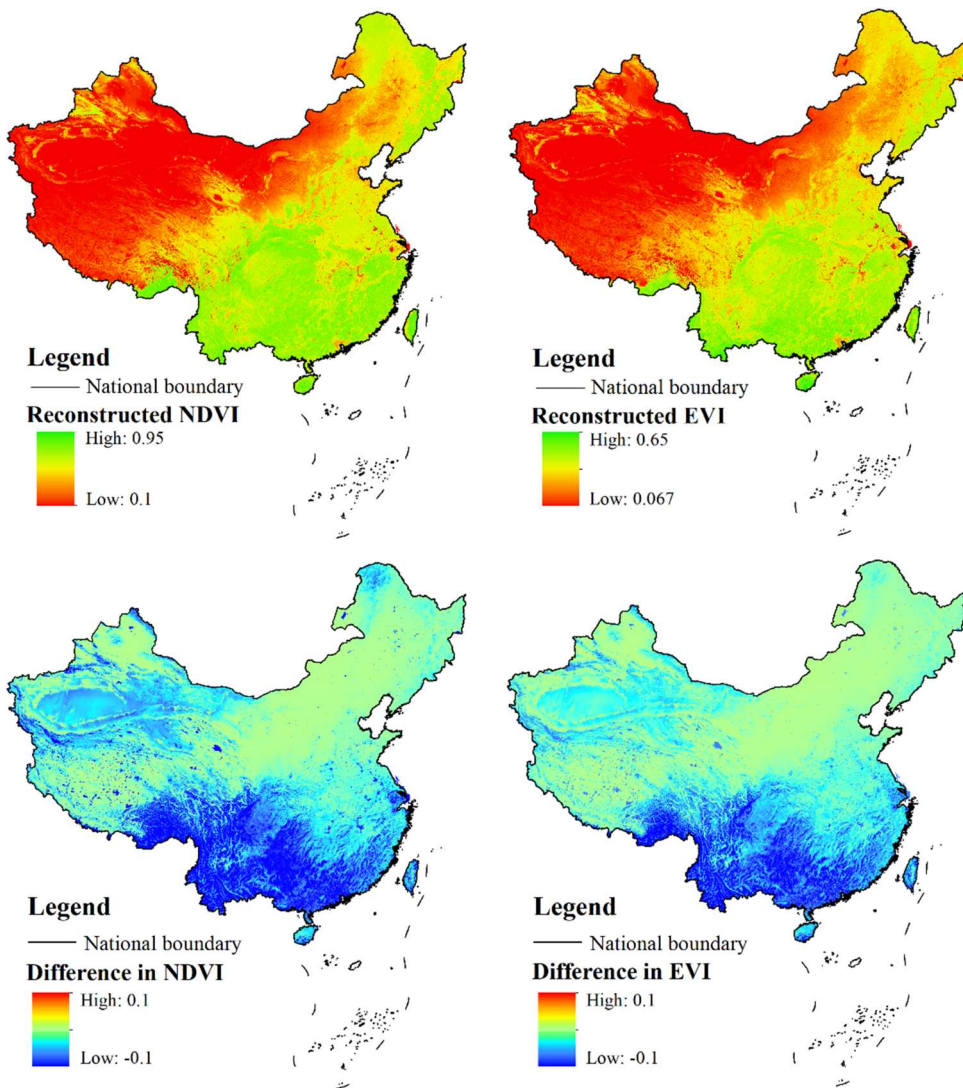


Figure 10. The reconstructed 23-year average growing season VI (first row) and the difference between the original and reconstructed 23-year average growing season VI (second row).

which is time-consuming (Militino et al. 2019). The computing time of the Gapfill is long because it ranks each image based on valuewise comparisons between all images (Gerber et al. 2018).

Third, the SIR method exhibited low RAM consumption. The reasons for the low RAM consumption of SIR are as follows: (1) SIR utilizes only one related image (the multiyear average image) to reconstruct the target image. Consequently, only two images (the target image and the multiyear average image) need to be read during the reconstruction process. (2) All the steps in the SIR method involve simple mathematical operations that do not require additional RAM allocation.

Fourth, the SIR method can simultaneously reconstruct all cloud-contaminated pixels, as can the Gapfill (Tables 1 and 2). This is because: (1) the related image does not contain any invalid values, and (2) there is no upper limit set for the size of the spatial window. In comparison, IMA and many previous methods cannot simultaneously reconstruct all missing values (Cao et al. 2018; Militino et al. 2019; Yang et al. 2022), or require many complex steps (Li et al. 2018; Yao et al. 2023c). The main reason for this is that the related images used in these methods contain invalid values.

Overall, SIR is a robust reconstruction method with high accuracy and low computing costs. The main difference between SIR and previous methods is that SIR has been proven suitable for large-scale reconstruction.

It is important to note that the SIR method is specifically designed for reconstructing VI data. It operates based on the hypothesis that the VI difference between closely spaced pixels in the target image is consistent with the multiyear average image (refer to step 3 in Section 3.2). However, this hypothesis may not hold true when reconstructing other types of data, such as land surface temperature and snow cover. This is because the spatial distribution of land surface temperature and snow cover can be influenced by various factors and can vary significantly across different years. For instance, land surface temperature can be affected by soil moisture, while snow cover can be influenced by snowfall. Therefore, SIR cannot be directly applied to reconstruct these types of data. Some modifications, such as adjusting the calculation method of related images, would be necessary when utilizing the SIR method for reconstructing such data.

5.2. The reconstructed NDVI and EVI datasets

This study developed 250 m spatial resolution and 16-day composite NDVI and EVI datasets in China from 2000 to 2022. The cloud-contaminated pixels were reconstructed using a high-precision spatiotemporal reconstruction method called SIR. The reconstructed NDVI and EVI datasets exhibited significantly higher quality compared to the original MODIS MOD13Q1 product (refer to Section 4.3).

Remote sensing VI data play a crucial role in various fields. First, it is extensively used to analyze vegetation dynamics and explore the relationships between vegetation and factors such as climate, hydrology, and human health (Geng et al. 2020; Liu et al. 2022; Yang et al. 2021). The improved quality of the developed VI datasets can enhance the reliability of conclusions drawn from these studies. Second, remote sensing VI data serve as an essential input for estimating crop yields, drought indices, groundwater recharge, aboveground biomass, and other parameters (Li et al. 2019; Parizi et al. 2020; Xie and Fan 2021; Zhang et al. 2021; 2023).

6. Conclusions

In this study, a method called SIR was proposed to reconstruct cloud-contaminated VI data. SIR has high accuracy, high computing speed, and low RAM consumption, and can simultaneously reconstruct all cloud-contaminated pixels.

The accuracy of SIR was validated using simulated cloud-contaminated areas. The average MAE of SIR was 0.0338, which was 20.2% and 23.4% lower than the MAE of the other two methods tested. The average computing time of SIR was 99.7% and 98.8% lower than that of the other two reconstruction methods tested in the WYM. For reconstructing NDVI images across all of China, the average computing time and RAM consumption of SIR were 1433.22 s and 6.44 GB, respectively. The SIR method, due to its low computing costs, is suitable for large-scale reconstruction. The low computing costs of the SIR method can be attributed to its well-designed framework such as a reduced number of related images and simple mathematical operations. The 250 m spatial resolution and 16-day composite reconstructed NDVI and EVI datasets in China from 2000 to 2022 were developed based on the SIR method (available at <https://doi.org/10.5281/zenodo.7988989> and <https://doi.org/10.5281/zenodo.7979989>, approximately 220 GB) (Yao et al. 2023a; 2023b). Spatiotemporal analyses demonstrated that the reconstructed datasets were more reliable than the original product and a similar dataset.

The developed reconstructed VI dataset can serve as essential input data in various fields such as ecology, hydrology, meteorology, agriculture, and forestry. Future studies can explore its applicability for reconstructing other types of data, such as land surface temperature and snow cover.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

MOD13Q1 NDVI and EVI data are accessible at: <https://ladsweb.modaps.eosdis.nasa.gov/>.

Funding

This work was financially supported by the National Natural Science Foundation of China (Grants No. 41971295 and 42271328), and the Special Fund of Hubei LuoJia Laboratory (grant number 220100031).

References

- Beck, P. S. A., C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore. 2006. "Improved Monitoring of Vegetation Dynamics at Very High Latitudes: A new Method Using MODIS NDVI." *Remote Sensing of Environment* 100 (3): 321–334. <https://doi.org/10.1016/j.rse.2005.10.021>.
- Cao, Ruyin, Yang Chen, Miaogen Shen, Jin Chen, Ji Zhou, Cong Wang, and Wei Yang. 2018. "A Simple Method to Improve the Quality of NDVI Time-series Data by Integrating Spatiotemporal Information with the Savitzky-Golay Filter." *Remote Sensing of Environment* 217:244–257. <https://doi.org/10.1016/j.rse.2018.08.022>.
- Cazorla, Beatriz P., Javier Cabello, Andrés Reyes, Emilio Guirado, Julio Peñas, Antonio J. Pérez-Luque, and Domingo Alcaraz-Segura. 2023. "A Remote-sensing-based Dataset to Characterize the Ecosystem Functioning and Functional Diversity in the Biosphere Reserve of the Sierra Nevada (Southeastern Spain)." *Earth System Science Data* 15:1871–1887. <https://doi.org/10.5194/essd-15-1871-2023>.
- Chen, Yang, Ruyin Cao, Jin Chen, Licong Liu, and Bunkei Matsushita. 2021. "A Practical Approach to Reconstruct High-Quality Landsat NDVI Time-series Data by gap Filling and the Savitzky-Golay Filter." *ISPRS Journal of Photogrammetry and Remote Sensing* 180:174–190. <https://doi.org/10.1016/j.isprsjprs.2021.08.015>.
- Chen, J., P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh. 2004. "A Simple Method for Reconstructing a High-Quality NDVI Time-series Data Set Based on the Savitzky-Golay Filter." *Remote Sensing of Environment* 91 (3–4): 332–344. <https://doi.org/10.1016/j.rse.2004.03.014>.
- Chen, Jing M., Weimin Ju, Philippe Ciais, Nicolas Viovy, Ronggao Liu, Yang Liu, and Xuehe Lu. 2019b. "Vegetation Structural Change Since 1981 Significantly Enhanced the Terrestrial Carbon Sink." *Nature Communications* 10:4259. <https://doi.org/10.1038/s41467-019-12257-8>.
- Chen, Chi, Taejin Park, Xuhui Wang, Shilong Piao, Baodong Xu, Rajiv K. Chaturvedi, Richard Fuchs, et al. 2019a. "China and India Lead in Greening of the World Through Land-use Management." *Nature Sustainability* 2:122–129. <https://doi.org/10.1038/s41893-019-0220-7>.
- Choler, Philippe, Arthur Bayle, Bradley Z. Carlson, Christophe Randin, Gianluca Filippa, and Edoardo Cremonese. 2021. "The Tempo of Greening in the European Alps: Spatial Variations on a Common Theme." *Global Change Biology* 27. <https://doi.org/10.1111/gcb.15820>.
- Chu, Dong, Huanfeng Shen, Xiaobin Guan, Jing M. Chen, Xinghua Li, Jie Li, and Liangpei Zhang. 2021. "Long Time-series NDVI Reconstruction in Cloud-prone Regions via Spatio-temporal Tensor Completion." *Remote Sensing of Environment* 264:112632. <https://doi.org/10.1016/j.rse.2021.112632>.
- Dalagnol, Ricardo, Lênio Soares Galvão, Fabien Hubert Wagner, Yhasmin Mendes de Moura, Nathan Gonçalves, Yujie Wang, Alexei Lyapustin, Yan Yang, Sassan Saatchi, and Luiz Eduardo Oliveira Cruz Aragão. 2023. "AnisoVeg: Anisotropy and Nadir-normalized MODIS Multi-angle Implementation Atmospheric Correction (MAIAC) Datasets for Satellite Vegetation Studies in South America." *Earth System Science Data* 15:345–358. <https://doi.org/10.5194/essd-15-345-2023>.
- Forzieri, Giovanni, Diego G. Miralles, Philippe Ciais, Ramdane Alkama, Youngryel Ryu, Gregory Duveiller, Ke Zhang, et al. 2020. "Increased Control of Vegetation on Global Terrestrial Energy Fluxes." *Nature Climate Change* 10 (4): 356–362. <https://doi.org/10.1038/s41558-020-0717-0>.
- Gao, Jixi, Yuanli Shi, Hongwei Zhang, Xuhui Chen, Wenguo Zhang, Wenming Shen, Tong Xiao, and Yuhuan Zhang. 2023. "China Regional 250m Normalized Difference Vegetation Index Data Set (2000–2022)." *National Tibetan Plateau/Third Pole Environment Data Center*. <https://doi.org/10.11888/Terre.tpd.c.300328>.
- Geng, Xiaojun, Xuancheng Zhou, Guodong Yin, Fanghua Hao, Xuan Zhang, Zengchao Hao, Vijay P. Singh, and Yongshuo H. Fu. 2020. "Extended Growing Season Reduced River Runoff in Luanhe River Basin." *Journal of Hydrology*. <https://doi.org/10.1016/j.jhydrol.2019.124538>.

- Gerber, Florian, Rogier de Jong, Michael E. Schaepman, Gabriela Schaepman-Strub, and Reinhard Furrer. 2018. "Predicting Missing Values in Spatio-temporal Remote Sensing Data." *IEEE Transactions on Geoscience and Remote Sensing* 56:2841–2853. <https://doi.org/10.1109/TGRS.2017.2785240>.
- Hilker, Thomas, Alexei I. Lyapustin, Compton J. Tucker, Forrest G. Hall, Ranga B. Myneni, Yujie Wang, Jian Bi, Yhasmin Mendes de Moura, and Piers J. Sellers. 2014. "Vegetation Dynamics and Rainfall Sensitivity of the Amazon." *Proceedings of the National Academy of Sciences of the United States of America* 111 (45): 16041–16046. <https://doi.org/10.1073/pnas.1404870111>.
- Huang, Anqi, Rumping Shen, Wenli Di, and Huimin Han. 2021. "A Methodology to Reconstruct LAI Time Series Data Based on Generative Adversarial Network and Improved Savitzky-Golay Filter." *International Journal of Applied Earth Observation and Geoinformation* 105:102633. <https://doi.org/10.1016/j.jag.2021.102633>.
- Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. "Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices." *Remote Sensing of Environment* 83 (1–2): 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2).
- Ims, Rolf A., John-Andre Henden, Marita A. Strømeng, Anders V. Thingnes, Mari J. Garmo, and Jane U. Jepsen. 2019. "Arctic Greening and Bird Nest Predation Risk Across Tundra Ecotones." *Nature Climate Change* 9 (8): 607–610. <https://doi.org/10.1038/s41558-019-0514-9>.
- Jakubauskas, Mark E., David R. Legates, and Jude H. Kastens. 2001. "Harmonic Analysis of Time-series AVHRR NDVI Data." *Photogrammetric Engineering and Remote Sensing* 67:461–470. https://www.asprs.org/wp-content/uploads/pers/2001journal/april/2001_apr_461-470.pdf.
- Jönsson, Per, and Lars Eklundh. 2002. "Seasonality Extraction by Function Fitting to Time-series of Satellite Sensor Data." *IEEE Transactions on Geoscience and Remote Sensing* 40 (8): 1824–1832. <https://doi.org/10.1109/TGRS.2002.802519>.
- Kong, Dongdong, Yongqiang Zhang, Xihui Gu, and Dagang Wang. 2019. "A Robust Method for Reconstructing Global MODIS EVI Time Series on the Google Earth Engine." *ISPRS Journal of Photogrammetry and Remote Sensing* 155:13–24. <https://doi.org/10.1016/j.isprsjprs.2019.06.014>.
- Li, Shuang, Liang Xu, Jiajia Chen, Yazhen Jiang, Shuying Sun, Shaohuai Yu, Zhenyu Tan, and Xinghua Li. 2023. "Monitoring Vegetation Dynamics (2010–2020) in Shengnongjia Forestry District with Cloud-removed MODIS NDVI Series by a Spatio-temporal Reconstruction Method." *The Egyptian Journal of Remote Sensing and Space Sciences* 26 (3): 527–543. <https://doi.org/10.1016/j.ejrs.2023.06.010>.
- Li, Shuang, Liang Xu, Yinghong Jing, Hang Yin, Xinghua Li, and Xiaobin Guan. 2021. "High-quality Vegetation Index Product Generation: A Review of NDVI Time Series Reconstruction Techniques." *International Journal of Applied Earth Observation and Geoinformation* 105:102640. <https://doi.org/10.1016/j.jag.2021.102640>.
- Li, Xiaoma, Yuyu Zhou, Ghassem R. Asrar, and Zhengyuan Zhu. 2018. "Creating a Seamless 1 km Resolution Daily Land Surface Temperature Dataset for Urban and Surrounding Areas in the Conterminous United States." *Remote Sensing of Environment* 206:84–97. <https://doi.org/10.1016/j.rse.2017.12.010>.
- Li, Xuecao, Yuyu Zhou, Lin Meng, Ghassem R. Asrar, Chaqun Lu, and Qisheng Wu. 2019. "A Dataset of 30 m Annual Vegetation Phenology Indicators (1985–2015) in Urban Areas of the Conterminous United States." *Earth System Science Data* 11:881–894. <https://doi.org/10.5194/essd-11-881-2019>.
- Liu, Huiyu, Fusheng Jiao, Jingqiu Yin, Tingyou Li, Haibo Gong, Zhaoyue Wang, and Zhenshan Lin. 2020. "Nonlinear Relationship of Vegetation Greening with Nature and Human Factors and its Forecast – A Case Study of Southwest China." *Ecological Indicators* 111:106009. <https://doi.org/10.1016/j.ecolind.2019.106009>.
- Liu, Yongchang, Zhi Li, Yaning Chen, Patient Mindje Kayumba, Xuanxuan Wang, Chuanxiu Liu, Yunxia Long, and Fan Sun. 2022. "Biophysical Impacts of Vegetation Dynamics Largely Contribute to Climate Mitigation in High Mountain Asia." *Agricultural and Forest Meteorology* 327:109233. <https://doi.org/10.1016/j.agrformet.2022.109233>.
- Liu, Ronggao, Rong Shang, Yang Liu, and Xiaoliang Lu. 2017. "Global Evaluation of Gap-filling Approaches for Seasonal NDVI with Considering Vegetation Growth Trajectory, Protection of Key Point, Noise Resistance and Curve Stability." *Remote Sensing of Environment* 189:164–179. <https://doi.org/10.1016/j.rse.2016.11.023>.
- Liu, Zihan, Wenfeng Zhan, Jiameng Lai, Falu Hong, Jinling Quan, Benjamin Bechtel, Fan Huang, and Zhaoxu Zou. 2019. "Balancing Prediction Accuracy and Generalization Ability: A Hybrid Framework for Modelling the Annual Dynamics of Satellite-Derived Land Surface Temperatures." *ISPRS Journal of Photogrammetry and Remote Sensing* 151:189–206. <https://doi.org/10.1016/j.isprsjprs.2019.03.013>.
- Militino, Ana F., M. Dolores Ugarte, Unai Perez-Goya, and Marc G. Genton. 2019. "Interpolation of the Mean Anomalies for Cloud Filling in Land Surface Temperature and Normalized Difference Vegetation Index." *IEEE Transactions on Geoscience and Remote Sensing* 57:6068–6078. <https://doi.org/10.1109/TGRS.2019.2904193>.
- Parizi, Esmaeel, Seiyed Mossa Hosseini, Behzad Ataie-Ashtiani, and Craig T. Simmons. 2020. "Normalized Difference Vegetation Index as the Dominant Predicting Factor of Groundwater Recharge in Phreatic Aquifers: Case Studies Across Iran." *Scientific Reports* 10:17473. <https://doi.org/10.1038/s41598-020-74561-4>.
- Piao, Shilong, Xuhui Wang, Taejin Park, Chi Chen, Xu Lian, Yue He, Jarle W. Bjerke, et al. 2019. "Characteristics, Drivers and Feedbacks of Global Greening." *Nature Reviews Earth & Environment* 1:14–27. <https://doi.org/10.1038/s43017-019-0001-x>.
- Savitzky, Abraham, and M. J. E. Golay. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." *Analytical Chemistry* 36: 1627–1639. <https://doi.org/10.1021/ac60214a047>.

- Song, Wenqi, Yuhao Feng, and Zhiheng Wang. 2022. "Ecological Restoration Programs Dominate Vegetation Greening in China." *Science of the Total Environment* 848:157729. <https://doi.org/10.1016/j.scitotenv.2022.157729>.
- Sun, Liang, Zhongxin Chen, Feng Gao, Martha Anderson, Lisheng Song, Limin Wang, Bo Hu, and Yun Yang. 2017. "Reconstructing Daily Clear-sky Land Surface Temperature for Cloudy Regions from MODIS Data." *Computers & Geosciences* 105:10–20. <https://doi.org/10.1016/j.cageo.2017.04.007>.
- Tian, Jiaqi, Xiaolin Zhu, Jin Chen, Cong Wang, Miaogen Shen, Wei Yang, Xiaoyue Tan, Shuai Xu, and Zhilin Li. 2021. "Improving the Accuracy of Spring Phenology Detection by Optimally Smoothing Satellite Vegetation Index Time Series Based on Local Cloud Frequency." *ISPRS Journal of Photogrammetry and Remote Sensing* 180:29–44. <https://doi.org/10.1016/j.isprsjsprs.2021.08.003>.
- Xie, Fei, and Hui Fan. 2021. "Deriving Drought Indices from MODIS Vegetation Indices (NDVI/EVI) and Land Surface Temperature (LST): Is Data Reconstruction Necessary?" *International Journal of Applied Earth Observation and Geoinformation* 101:102352. <https://doi.org/10.1016/j.jag.2021.102352>.
- Xue, Jingru, and Baofeng Su. 2017. "Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications." *Journal of Sensors* 2017: 1353691. <https://doi.org/10.1155/2017/1353691>.
- Yang, Xue, Jin Chen, Qingfeng Guan, Huan Gao, and Wei Xia. 2022. "Enhanced Spatial–Temporal Savitzky–Golay Method for Reconstructing High-Quality NDVI Time Series: Reduced Sensitivity to Quality Flags and Improved Computational Efficiency." *IEEE Transactions on Geoscience and Remote Sensing* 60:1–17. <https://doi.org/10.1109/tgrs.2022.3190475>.
- Yang, Jie, and Xin Huang. 2021. "The 30m Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019." *Earth System Science Data* 13: 3907–3925. <https://doi.org/10.5194/essd-13-3907-2021>.
- Yang, Gang, Huanfeng Shen, Liangpei Zhang, Zongyi He, and Xinghua Li. 2015. "A Moving Weighted Harmonic Analysis Method for Reconstructing High-quality SPOT VEGETATION NDVI Time-Series Data." *IEEE Transactions on Geoscience and Remote Sensing* 53:6008–6021. <https://doi.org/10.1109/TGRS.2015.2431315>.
- Yang, Bo-Yi, Tianyu Zhao, Li-Xin Hu, Matthew H.E.M. Browning, Joachim Heinrich, Shyamali C. Dharmage, Bin Jalaludin, et al. 2021. "Greenspace and Human Health: An Umbrella Review." *The Innovation* 2:100164. <https://doi.org/10.1016/j.xinn.2021.100164>.
- Yao, Rui, Xin Huang, Yongjun Zhang, Lunche Wang, Jiayi Li, and Qiquan Yang. 2023a. "Reconstructed 250 m Resolution NDVI Dataset in China." <https://doi.org/10.5281/zenodo.7988989>.
- Yao, Rui, Xin Huang, Yongjun Zhang, Lunche Wang, Jiayi Li, and Qiquan Yang. 2023b. "Reconstructed 250 m Resolution EVI Dataset in China." <https://doi.org/10.5281/zenodo.7979989>.
- Yao, Rui, Lunche Wang, Xin Huang, Qian Cao, Jing Wei, Panxing He, Shaoqiang Wang, and Lizhe Wang. 2023c. "Global Seamless and High-Resolution Temperature Dataset (GSHTD), 2001–2020." *Remote Sensing of Environment* 286:113422. <https://doi.org/10.1016/j.rse.2022.113422>.
- Yao, Rui, Lunche Wang, Xin Huang, Xinxin Chen, and Zhengjia Liu. 2019. "Increased Spatial Heterogeneity in Vegetation Greenness Due to Vegetation Greening in Mainland China." *Ecological Indicators* 99:240–250. <https://doi.org/10.1016/j.ecolind.2018.12.039>.
- Yao, Rui, Lunche Wang, Xin Huang, Liang Sun, Ruiqing Chen, Xiaojun Wu, Wei Zhang, and Zigeng Niu. 2021. "A Robust Method for Filling the Gaps in MODIS and VIIRS Land Surface Temperature Data." *IEEE Transactions on Geoscience and Remote Sensing* 59:1–15. <https://doi.org/10.1109/tgrs.2021.3053284>.
- Zeng, Zhenzhong, Shilong Piao, Laurent Z. X. Li, Tao Wang, Philippe Ciais, Xu Lian, Yuting Yang, Jiafu Mao, Xiaoying Shi, and Ranga B. Myneni. 2018. "Impact of Earth Greening on the Terrestrial Water Cycle." *Journal of Climate* 31 (7): 2633–2650. <https://doi.org/10.1175/JCLI-D-17-0236.1>.
- Zeng, Chao, Huanfeng Shen, Mingliang Zhong, Liangpei Zhang, and Penghai Wu. 2015. "Reconstructing MODIS LST Based on Multitemporal Classification and Robust Regression." *IEEE Geoscience and Remote Sensing Letters* 12:512–516. <https://doi.org/10.1109/LGRS.2014.2348651>.
- Zhang, Yulong, Conghe Song, Lawrence E. Band, Ge Sun, and Junxiang Li. 2017. "Reanalysis of Global Terrestrial Vegetation Trends from MODIS Products: Browning or Greening?" *Remote Sensing of Environment* 191:145–155. <https://doi.org/10.1016/j.rse.2016.12.018>.
- Zhang, Huifang, Zhonggang Tang, Binyao Wang, Hongcheng Kan, Yi Sun, Yu Qin, Baoping Meng, et al. 2023. "A 250 m Annual Alpine Grassland AGB Dataset Over the Qinghai–Tibet Plateau (2000–2019) in China Based on in Situ Measurements, UAV Photos, and MODIS Data." *Earth System Science Data* 15:821–846. <https://doi.org/10.5194/essd-15-821-2023>.
- Zhang, Liangliang, Zhao Zhang, Yuchuan Luo, Juan Cao, Ruizhi Xie, and Shaokun Li. 2021. "Integrating Satellite-Derived Climatic and Vegetation Indices to Predict Smallholder Maize Yield Using Deep Learning." *Agricultural and Forest Meteorology* 311:108666. <https://doi.org/10.1016/j.agrformet.2021.108666>.
- Zhao, Lin, Aiguo Dai, and Bo Dong. 2018. "Changes in Global Vegetation Activity and its Driving Factors during 1982–2013." *Agricultural and Forest Meteorology* 249:198–209. <https://doi.org/10.1016/j.agrformet.2017.11.013>.
- Zheng, Lei, Guosong Zhao, Jinwei Dong, Quansheng Ge, Jian Tao, Xuezheng Zhang, Youcun Qi, Russell B. Doughty, and Xiangming Xiao. 2019. "Spatial, Temporal, and Spectral Variations in Albedo due to Vegetation Changes in China's Grasslands." *ISPRS Journal of Photogrammetry and Remote Sensing* 152:1–12. <https://doi.org/10.1016/j.isprsjsprs.2019.03.020>