remote sensing

MDPI

Technical Note

LACC2.0: Improving the LACC Algorithm for Reconstructing Satellite-Derived Time Series of Vegetation Biochemical Parameters

Mingzhu Xu [1], Rong Shang [1,*], Jing M. Chen [1,2] and Lingfang Zeng [3,4]

[1] Key Laboratory for Humid Subtropical Ecogeographical Processes of the Ministry of Education, School of Geographical Sciences, Fujian Normal University, Fuzhou 350007, China; xumz@fjnu.edu.cn (M.X.); jing.chen@utoronto.ca (J.M.C.)
[2] Department of Geography and Planning, University of Toronto, Toronto, ON M5S 3G3, Canada
[3] Shanghai Surveying and Mapping Institute, Shanghai 200063, China; lfzeng@shsmi.cn
[4] Key Laboratory of Spatial-Temporal Big Data Analysis and Application of Natural Resources in Megacities, Shanghai 200333, China
* Correspondence: shangrong@fjnu.edu.cn

Abstract: The locally adjusted cubic-spline capping (LACC) algorithm is well recognized for its effectiveness in the global time series reconstruction of vegetation biophysical and biochemical parameters. However, in its application, we often encounter issues, such as identifying positively biased outliers for vegetation biochemical parameters and reducing the influence of long consecutive gaps. In this study, we improved the LACC algorithm to address the above two issues by (1) incorporating a procedure to remove outliers and (2) integrating the spatial information of neighboring pixels for large data gap filling. To evaluate the performance of the new version of LACC (namely LACC2.0), leaf chlorophyll content (LCC) was taken as an example. A reference LCC curve was generated for each pixel of the global map as the true value for global evaluation, and a time series of LCC with real gaps in the original data for each pixel was created by adding Gaussian noises into observations for testing the effectiveness of time series reconstruction algorithms. Results showed that the percentage of pixels with an RMSE smaller than 5 $\mu g/cm^2$ was improved from 81.2% in LACC to 96.4% in LACC2.0, demonstrating that LACC2.0 had the potential to provide a better reconstruction of global daily satellite-derived vegetation biochemical parameters. This finding highlights the significance of outlier removal and spatial-temporal fusion to enhance the accuracy and reliability of time series reconstruction.

Keywords: time series reconstruction; gap filling; leaf chlorophyll content; LACC; noise removal; spatio-temporal interpolation

check for updates

Citation: Xu, M.; Shang, R.; Chen, J.M.; Zeng, L. LACC2.0: Improving the LACC Algorithm for Reconstructing Satellite-Derived Time Series of Vegetation Biochemical Parameters. *Remote Sens.* **2023**, *15*, 3277. https://doi.org/10.3390/rs15133277

Academic Editors: Gaofei Yin, Jean-Philippe Gastellu-Etchegorry, Baodong Xu and Shengbiao Wu

Received: 13 May 2023
Revised: 22 June 2023
Accepted: 24 June 2023
Published: 26 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A long-term time series of satellite-derived vegetation biophysical and biochemical parameters, such as leaf area index (LAI), clumping index (CI), and leaf chlorophyll content (LCC), are of great importance for global carbon cycle modeling [1–5]. Unfortunately, cloud, cloud shadow, snow cover, and thick aerosol on satellite images often hinder the retrieval of vegetation parameters, and result in gaps and noises in the derived time series. These missing data and noises of vegetation biophysical and biochemical parameters further hinder the simulation of seasonal and inter-annual variations of the vegetation carbon sink using carbon cycle models [6–8]. It is essential to perform time series reconstruction to recover missing data and maintain the temporal continuity and integrity of these parameters for facilitating accurate carbon cycle modeling.

Many time series reconstruction algorithms were proposed to process satellite-derived surface parameters [6,9–12]. They can be classified into four types [9]: (1) threshold-based algorithms, such as the mean-value iteration (MVI) [13], and the iterative interpolation

for data reconstruction (IDR) [14]; (2) filter-based algorithms, such as the Savitzky-Golay (SG) filter [15] and the changing-weight filter [16]; (3) curve-fitting algorithms, such as the double logistic function (DL) [17], the asymmetric Gaussian function (AG) [18], the Fourier-based approach (Fourier) [19,20], and the locally adjusted cubic spline capping approach (LACC) [6]; (4) and other algorithms, such as the Whittaker smoother (Whit) [21]. The LACC algorithm [6], which is a well-recognized curve-fitting algorithm [9], has shown good performances in reconstructing the time series of LAI, CI, and LCC at the global scale [1–3,22]. LACC has a comprehensive ability to consider rapid vegetation growth and senescence curvatures, to protect key points of a time series, to resist data noise, and to achieve curve stability [9]. It was found to be the most effective algorithm among six gap-filling algorithms (Fourier, DL, IDR, Whit, SG, and LACC) in reconstructing a time series of normalized difference vegetative index (NDVI) at a global scale [9]. The LACC algorithm attains this ability by automatically determining the seasonal variation of the curvature of a parameter curve from a noisy time series [6]. However, issues in identifying positively biased outliers for vegetation biochemical parameters and reducing the influence of long consecutive gaps still exist in LACC and could be addressed to improve the time series reconstruction for carbon cycle modeling.

First, the current LACC has a poor ability to distinguish positively biased outliers. Outliers commonly exist in the time series of satellite-derived vegetation biophysical and biochemical parameters due to residual clouds, cloud shadow, etc. that are omitted by the regular quality control of products. They correspond to unreliable values of vegetation biophysical and biochemical parameters and influence subsequent applications. In most cases, vegetation biophysical parameters tend to be negatively biased due to the influence of cloud and cloud shadow. Purposely, most negatively biased outliers could be identified and replaced by optimum values by LACC [6]. However, positively biased retrieval is also possible for vegetation biochemical parameters at a leaf scale (e.g., LCC). This could frequently happen when a vegetation biophysical parameter (e.g., LAI or CI), used as an input to the retrieval algorithm of a vegetation biochemical parameter, is negatively biased, e.g., when LAI is too small the retrieved LCC would be too large [3,22,23]. These positively biased outliers of vegetation biochemical parameters tend to be recognized as valid data by LACC, leading to false-peak fluctuations in the reconstructed time series of LCC [3]. Integrating an outlier removal procedure into LACC could be helpful to minimize the influence of remaining outliers, especially for positively biased outliers of vegetation biochemical parameters.

Second, although the temporal information of a time series is fully used, the current LACC cannot reconstruct a seasonal trajectory well when a time series has long consecutive gaps in the growing season of vegetation [9]. Spatial information has been effectively used in several studies of spatio-temporal reconstruction of remote sensing data [24–27]. Compared with a single pixel, a neighboring area has a much higher probability of obtaining valid observations. As neighboring pixels usually have similar climates, vegetation types, and vegetation growth conditions [26,28], they have the potential to provide valid information during long consecutive gaps, and help split long consecutive gaps into smaller gaps during time series reconstruction [15,24–27,29]. It was found that the Spatial-Temporal Savitzky-Golay (STSG) method performed significantly better than SG, AG, DL, and Fourier in addressing the problem of temporally continuous NDVI gaps [30]. Therefore, it was crucial to integrate spatial information into the LACC algorithm to reduce the influence of long consecutive gaps.

In this study, we aimed to improve the LACC algorithm (named LACC2.0) for time series reconstruction of satellite-derived vegetation biochemical parameters in two ways: (1) to incorporate a procedure to screen additional outliers omitted by the regular quality control of the algorithm; and (2) to make use of spatial information for temporal gap filling. To compare and evaluate the performance of LACC and LACC2.0, a global LCC product was processed as an example. Reference LCC curves were generated as true values for

global evaluation, and the modeled time series of LCC with real gaps and Gaussian noises were created for global reconstruction.

2. Data and Methods

2.1. The LCC, Surface Reflectance, and Land Cover Data

The global MODIS leaf chlorophyll content (LCC) product [3] was used to evaluate the performance of the LACC2.0 time series reconstruction algorithm. It provides eight-day LCC data at the 500-m spatial resolution from 2000 to 2021. This product was generated from the MOD09A1 surface reflectance product. The inflection-based cloud detection (IBCD) algorithm [31] was used to screen cloud and snow cover, and the StateQA band of MOD09A1 was used to screen thick aerosol in the surface reflectance data. The screening of invalid observations resulted in plenty of gaps in the derived LCC data. To provide a seasonally continuous version of the LCC product, the LACC algorithm was used to fill the gaps and smooth the time series of LCC [3]. The original version of LCC with gaps was used to evaluate the performance of LACC and LACC2.0 in this study. The reconstructed LCC product generated using the original LACC algorithm was compared to the newly reconstructed LCC data generated using LACC2.0.

The MOD09A1 surface reflectance product was used for the additional outlier removal in LACC2.0, as it contains more information on outliers than the single time series of LCC. The MOD09A1 provides eight-day composited surface reflectance at 500-m spatial resolution since 2000. The surface reflectance in Red, Green, Blue, near-infrared (NIR), short wavelength infrared 1 (SWIR1), and 2 (SWIR2) bands were selected to screen the remaining outliers omitted by the StateQA band of the MOD09A1 product and the IBCD cloud detection algorithm.

The ESA CCI land cover data [32] was used to integrate spatial information during time series reconstruction using LACC. It was resampled from 300 m spatial resolution into 500 m spatial resolution and projected into the MODIS sinusoidal projection.

2.2. The LACC2.0 Algorithm

The LACC algorithm was designed for daily time series reconstruction of vegetation indices (VIs) and vegetation biophysical and biochemical parameters using piecewise cubic-spline curve fitting (Equation (1)) based on the principle of local weighting [6]. It was flexible for reconstructing the vegetation parameters with a wide range of seasonal patterns [9]. There were three iterations in LACC to automatically determine the seasonal variation of curvature: the initial curve fitting was used to determine the weights for each point, the second curve fitting was used to adjust the weights for each point, and the final curve fitting used the adjusted weights to output the reconstructed daily time series.

$$M_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (1)$$

where, x is the day-of-year (DOY) that ranges from x_i to x_{i+1}, x_i is the x-axis coordinate for the i-th knot, $M_i(x)$ is the cubic-spline model built between i-th and $(i+1)$-th knots, $i = 1, 2, \ldots, n-1$ (n is the total number of valid observations in the time series), and a_i, b_i, c_i, d_i are the model parameters.

The improved LACC (named LACC2.0) algorithm was modified in two main ways: first, to incorporate a procedure to screen outliers; second, to integrate the spatial information for time series reconstruction. Additionally, the current implementation of LACC reconstructed the time series data for each year separately, which could result in obvious mutations of the reconstructed parameters at the junction of two years, especially for the Southern Hemisphere and Indian monsoon region when vegetation growth is rapid [33]. In LACC2.0, we conducted continuous reconstruction for all years to avoid these mutations. These modifications were labeled with different colors in the flowchart of the LACC2.0 algorithm (Figure 1). There were three parts (rounded rectangle with dotted

lines) in LACC2.0: QA screening and outlier removal, spatial interpolation, and iterative temporal interpolation.

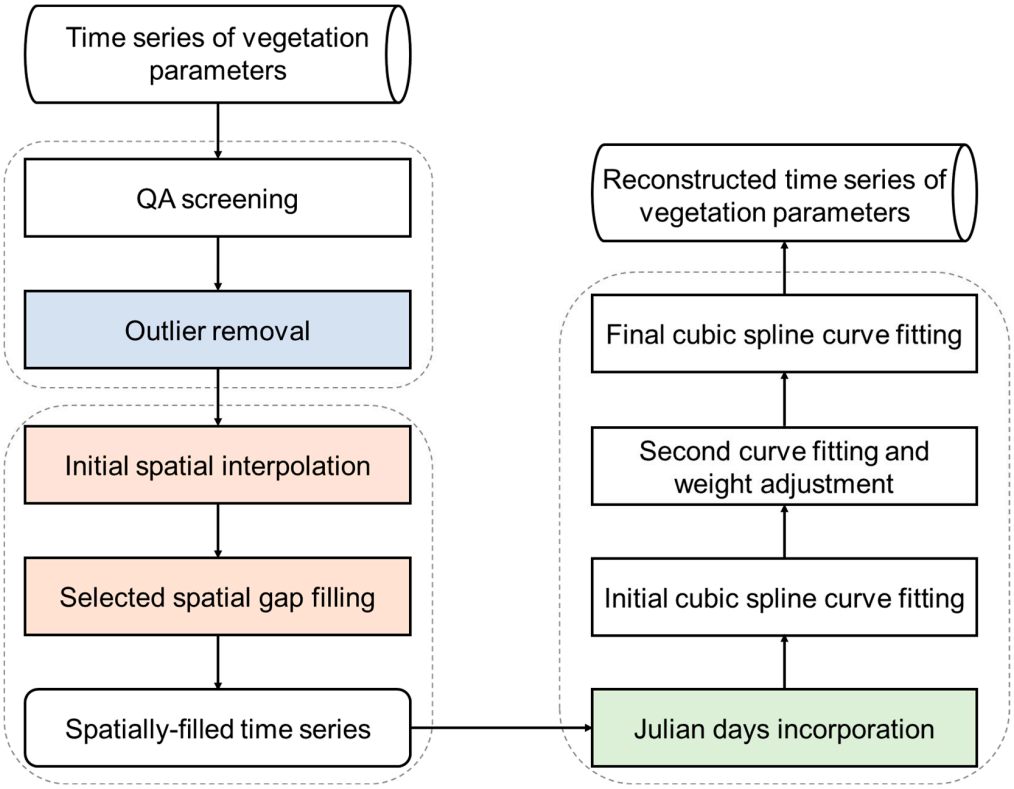


Figure 1. Flowchart of the LACC2.0 algorithm. The rounded rectangles with dotted lines indicate three parts of LACC2.0: QA screening and outlier removal, spatial interpolation, and iterative temporal interpolation. The colored rectangles are the improvements of LACC2.0 compared to the original LACC.

Outlier removal was conducted right after quality control of the LCC product. As outliers usually alter values in the time series abruptly and ephemerally [34], the time series filter proposed by Shang et al. [35] was used to remove the remaining outliers in the time series of LCC:

$$\sum_{\Lambda=1}^{m}\left[\frac{r_\Lambda(i)-0.5\times(r_\Lambda(i-1)+r_\Lambda(i+1))}{Dif_\Lambda(i)}\right]^2>\chi^2(m,Op) \quad (2)$$

$$Dif_\Lambda(i)=\sqrt{\sum_{i=2}^{N-1}[r_\Lambda(i)-0.5\times(r_\Lambda(i-1)+r_\Lambda(i+1))]^2/(n-2)} \quad (3)$$

where, m is the number of MOD09A1 bands (m = 6 as the red, green, blue, NIR, SWIR1, and SWIR2 bands of MODIS were used), Λ is the band index (from 1 to m), $r_\Lambda(i)$ is the ith observation in a time series, n is the number of observations of a time series, $\chi^2(k,Op)$ is the threshold used for the outlier removal derived from the chi-squared distribution, and Op is the outlier probability (determined to be 0.99 through sensitivity analysis in Table A1). For a given observation, if the normalized square difference between its surface reflectance and the mean surface reflectance of its previous and next observations at six bands exceeds the threshold obtained from the chi-squared distribution, it will be identified as an outlier. Both positively and negatively biased outliers will be identified. This filter has been successfully used for screening the outliers in the time series of harmonized Landsats 7–8 and Sentinel-2 surface reflectance [35]. It should be noted that we used the MOD09A1 surface reflectance to screen the outliers directly. For other parameters, if there is no available input surface

reflectance to aid in identifying the outliers, the parameter itself could be used with a value of m equal to 1.

After outlier removal, the spatial information was integrated for those observations with no available values in the original time series of LCC. The data-adapted iterative steering kernel regression (DISKR) algorithm [36] was used to incorporate the spatial information for time series reconstruction. By using a nonlinear data-adapted procedure, DISKR could overcome the limitations caused by the linear properties of classic kernel regression [36]. The algorithm considered not only the pixel location and density of raster data but also on the image features such as edges and orientations [36]. DISKR was structured as an optimization problem:

$$\min_{\{\beta_n\}} \sum_{i=1}^{P} \left[y_i - \beta_0 - \beta^T (x_i - x) - \beta_2^T vech(x_i - x)(x_i - x)^T - \cdots \right]^2 \cdot K_{adapt}(x_i - x, y_i - y) \quad (4)$$

where, P is the number of points in the 11 × 11 window (11 × 11 was determined as the optimal window size by the sensitivity analysis in Table A2), x_i is the location coordinates of the i-th point, y_i is the pixel value of the i-th point, x is the location coordinates of the pixel to be processed, y is the predicted pixel value of the pixel to be processed, $K_{adapt}(x_i - x, y_i - y)$ is the adaptive kernel function used to assign weights to each point to indicate the degree of influence that each sample point has on the pixel to be processed, β_n is a regression coefficient vector that includes an intercept (β_0) and coefficients for several independent variables, and $vech(x_i - x)$ is a column vector obtained by vertically stacking the upper triangular elements of the matrix $(x_i - x)$. The initial spatial interpolation by DISKR could fill all the gaps, but not all spatially-filled values are reliable. For a certain observation of a pixel, only when the valid observations in the same land cover types were more than 20% in its nearby 11 × 11 window after outlier removal, the spatially-filled value will be used.

Finally, the implementation of the LACC2.0 was modified by reconstructing the time series in each year separately to continuous reconstruction for multiple years. The input variable of doy_{obs} (day-of-year for each input observation, 366 for maximum) used in the cubic-spline curve fitting was replaced by $date_{obs}$, which records the observing date (Julian days) of each input observation in the time series. With this modification, the observations for all years could be stacked together, and LACC2.0 could reconstruct the time series of vegetation parameters continuously.

2.3. Validation Methods

Pixel-level reference LCC curves and a time series of LCC with realistic gaps and noises were generated for validation at a global scale. For a specific pixel, its reference LCC curve was first calculated as the multi-year average of each day-of-year (DOY) and then smoothed by the Savitzky-Golay filter [15] with a smoothing polynomial of four and a half-width of five [9]. Reference LCC curves (Figure 2) created this way were used as true values for evaluating time series reconstruction algorithms. Based on the reference LCC curve, a modeled five-year time series of LCC with gaps and noises (Figure 3) was generated as input for time series reconstruction. The real gaps in the MODIS LCC product from 2015 to 2019 were first added to the reference LCC curve by setting the LCC values of gaps as zero. Random noise was then added into 20% of randomly-selected positions of the time series where LCC > 0, and the maximum magnitude of noise was constrained to 40% of the selected values. It should be noted that both negatively biased noises and positively biased noises were simulated in this study, which was different from global evaluations and comparisons of six time series reconstruction methods where only negatively biased noises were simulated [9].

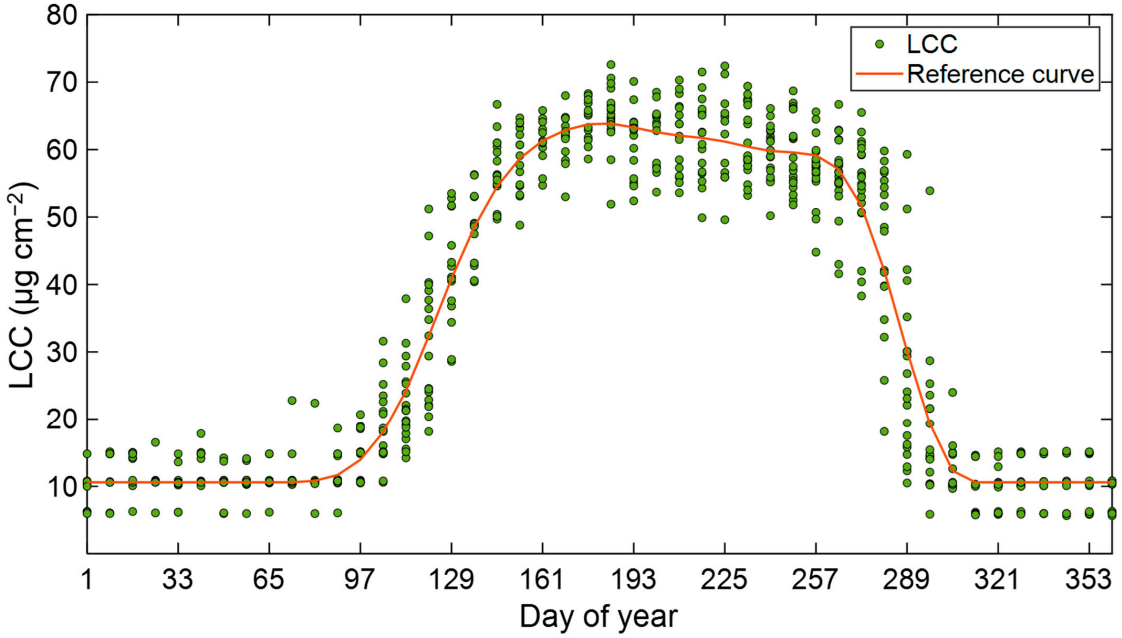


Figure 2. Example of a reference LCC curve of a pixel.

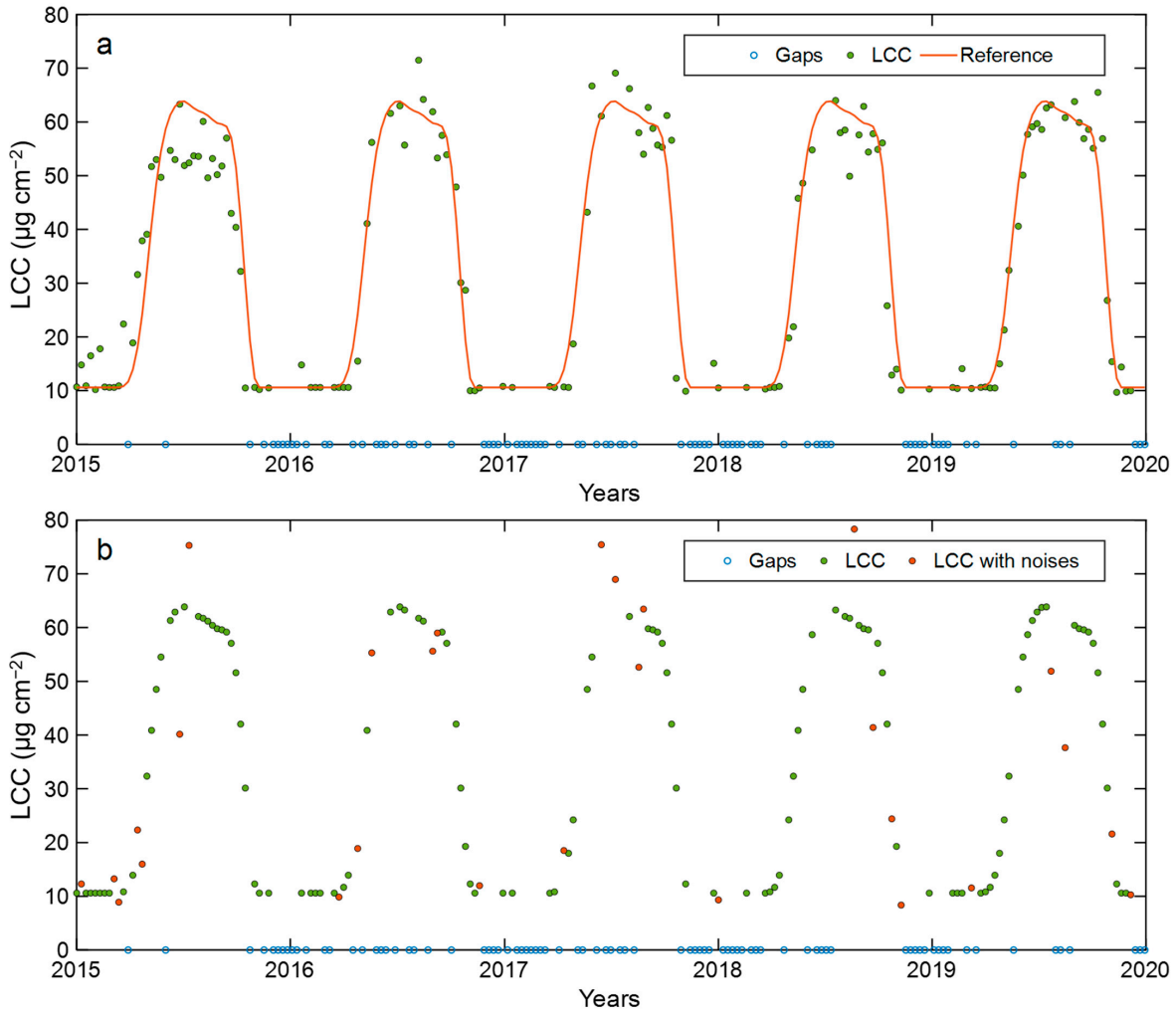


Figure 3. Example of gaps from original LCC data (**a**) and generation of the modeled time series of LCC by adding gaps and noise into the reference LCC (**b**).

The root mean squared error (RMSE) calculated between the reference LCC curve and the reconstructed time series, derived from the modeled time series of LCC with gaps and noises, was used as the indicator of evaluation. A lower value of RMSE indicates a better performance of the time series reconstruction. Both LACC and LACC2.0 were used to reconstruct the modeled five-year time series of LCC pixel by pixel at the global scale. Additionally, to indicate the importance of incorporating spatial interpolation, LACC with the processes of outlier removal of the time series but without incorporating spatial interpolation (shortened as to $LACC_O$), was also tested to reconstruct modeled global LCC. Three global maps of RMSE were then calculated between the reference and reconstructed time series using the three algorithms (LACC, $LACC_O$, and LACC2.0). The spatial variation of the three RMSE maps was analyzed. The percentage of pixels with an RMSE smaller than 5 $\mu g/cm^2$ was used to indicate their overall performance.

3. Results

3.1. Global Gap Percentages of the MODIS LCC Product

The filled values and identified outliers in the MODIS LCC product were regarded as gaps. The gaps that appeared in rapid vegetation growth seasons had a larger impact on the time series reconstruction than other seasons [9]. Figure 4 shows global statistics of the gap percentages in the MODIS LCC product from 2000 to 2021 for different seasons. The boreal regions had a very high gap percentage of LCC in winter (>90%), spring (30–90%), and autumn (30–80%), which could be explained by the long-time snowfall and snow cover. Tropical regions such as Amazon forests, the Malay Archipelago, the Gulf of Guinea, and the Congo Basin also had high gap percentages (over 50%) of LCC in most seasons. The south of Asia had a high gap percentage of LCC in summer (>70%) due to the frequent cloud covers under the Asian monsoon climate. Figure 5 shows global statistics of the maximum consecutive gaps in the MODIS LCC product, which had a large impact on the time series reconstruction [9]. The large consecutive gaps were also mainly distributed in the boreal regions, tropical forest regions, and south of Asia. The consecutive gaps were the largest in boreal regions (80–280 days), followed by some tropic regions, especially regions along the Andes, the east of Amazon forests, and the Gulf of Guinea (typically over 120 days).

3.2. Pixel-Level Reconstruction of Time Series of LCC

The LACC algorithm was modified by incorporating a procedure of outlier removal, reconstructing time series data in all years together, and incorporating spatial information into time series reconstruction. Figure 6 shows an example of time series reconstruction with and without the outlier removal at a pixel with a cover type of grassland. Using LACC, the positively biased outliers at the end of 2009, 2012, and 2013 caused obviously positive fluctuations in the time series reconstruction. In comparison, the negatively biased outliers at the beginning of 2009 did not cause obvious negative fluctuation. By incorporating the procedure of outlier removal, these false fluctuations were avoided and the seasonal trajectories of LCC were well captured by the time series reconstruction from the LACC2.0.

Figure 7 shows the difference in reconstructing the LCC in each year separately (Figure 7a) and reconstructing the LCC in all years together (Figure 7b) using LACC, taking a pixel covered by savanna in the Southern Hemisphere as an example. Obvious mutations of the reconstructed LCC were observed at the junction of different years (especially at the end of the years 2009, 2011, and 2013) when the LCC in each year was reconstructed by LACC separately. These mutations were effectively avoided when the LCC in all years were reconstructed together.

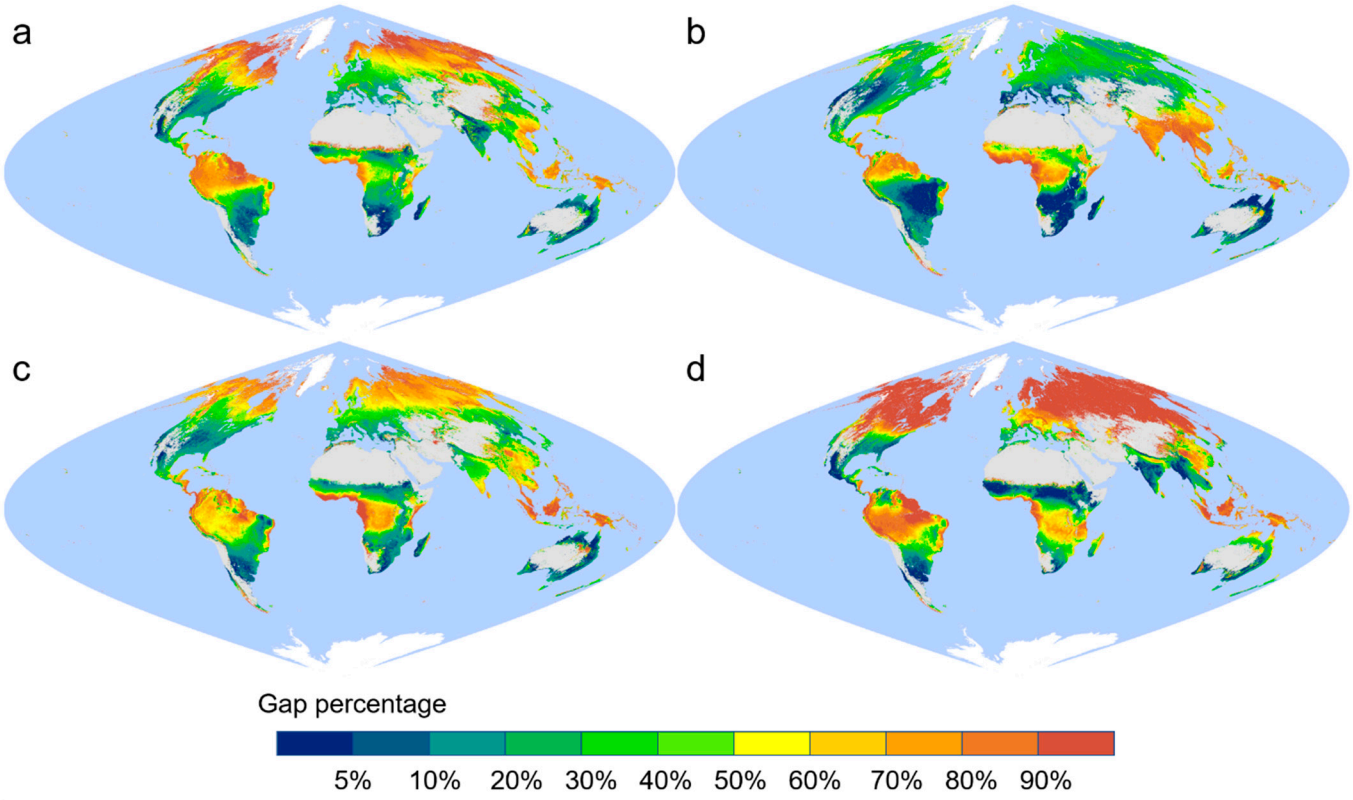


Figure 4. Global statistics of the gap percentages in the MODIS LCC product from 2000 to 2021 for different seasons: (**a**) March to May, (**b**) June to August, (**c**) September to November, and (**d**) December to February.

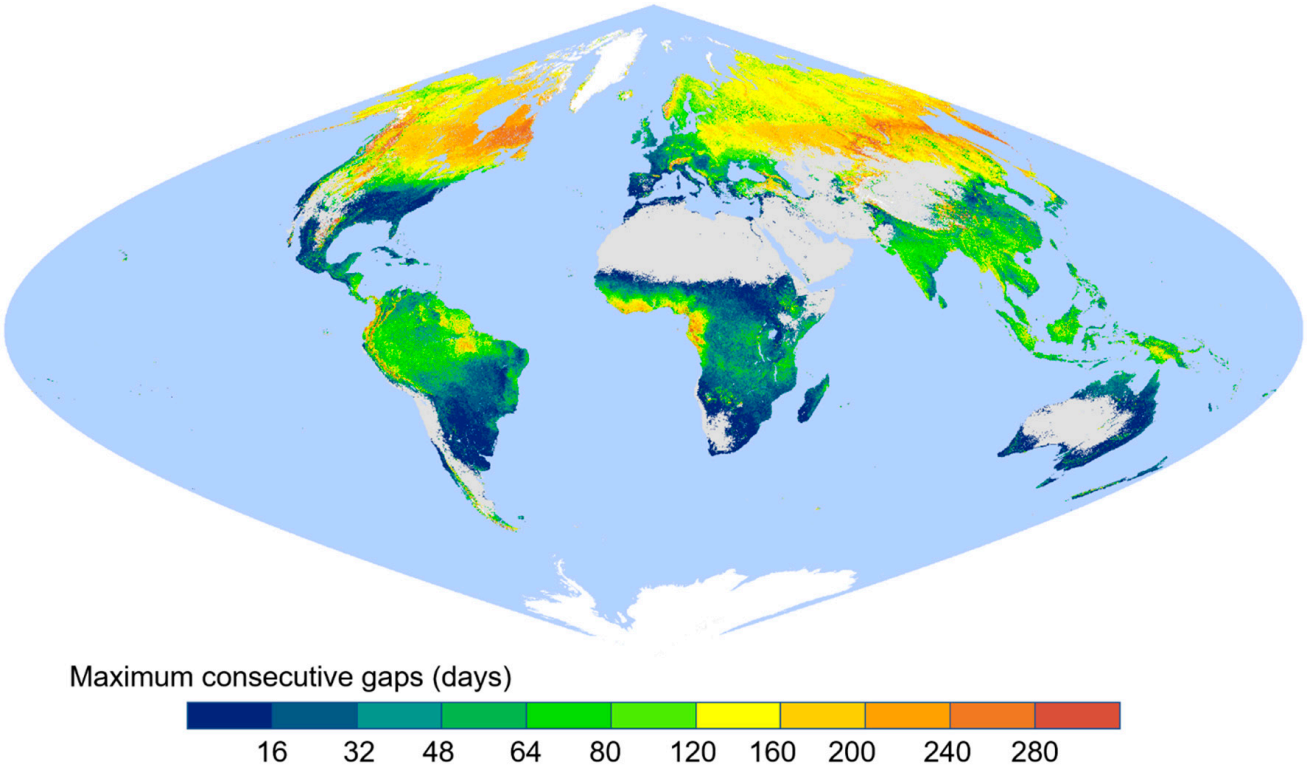


Figure 5. Global statistics of the maximum consecutive gaps in the MODIS LCC product.

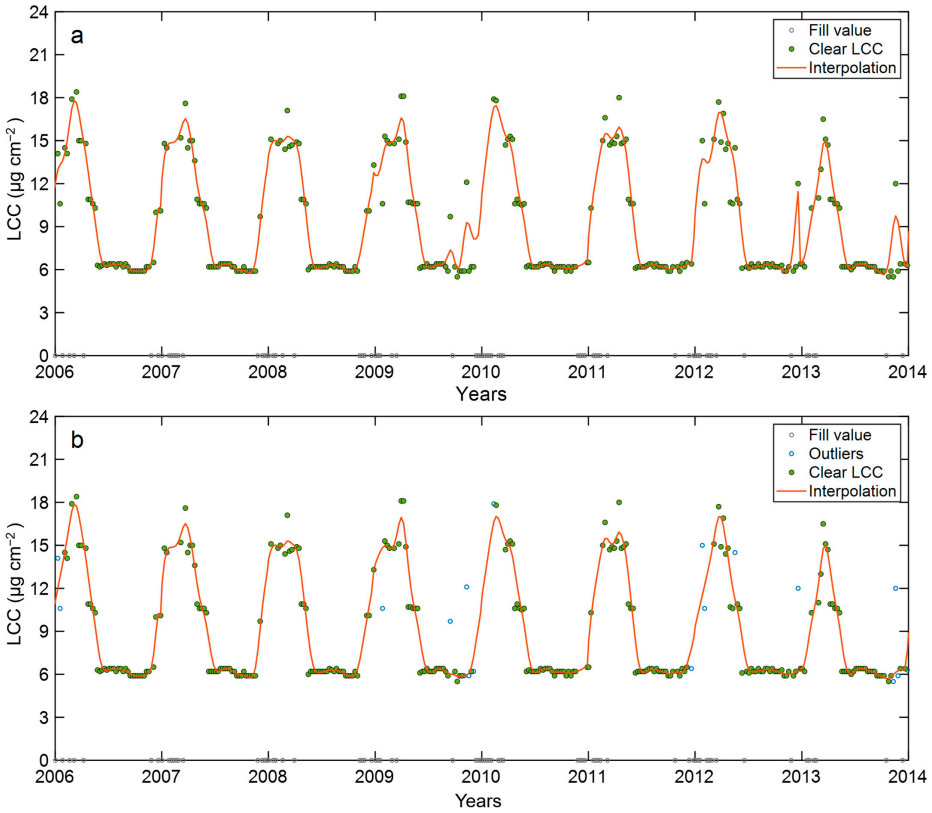


Figure 6. Comparison of time series reconstruction without the outlier removal using LACC (**a**) and with the outlier removal using LACC2.0 (**b**).

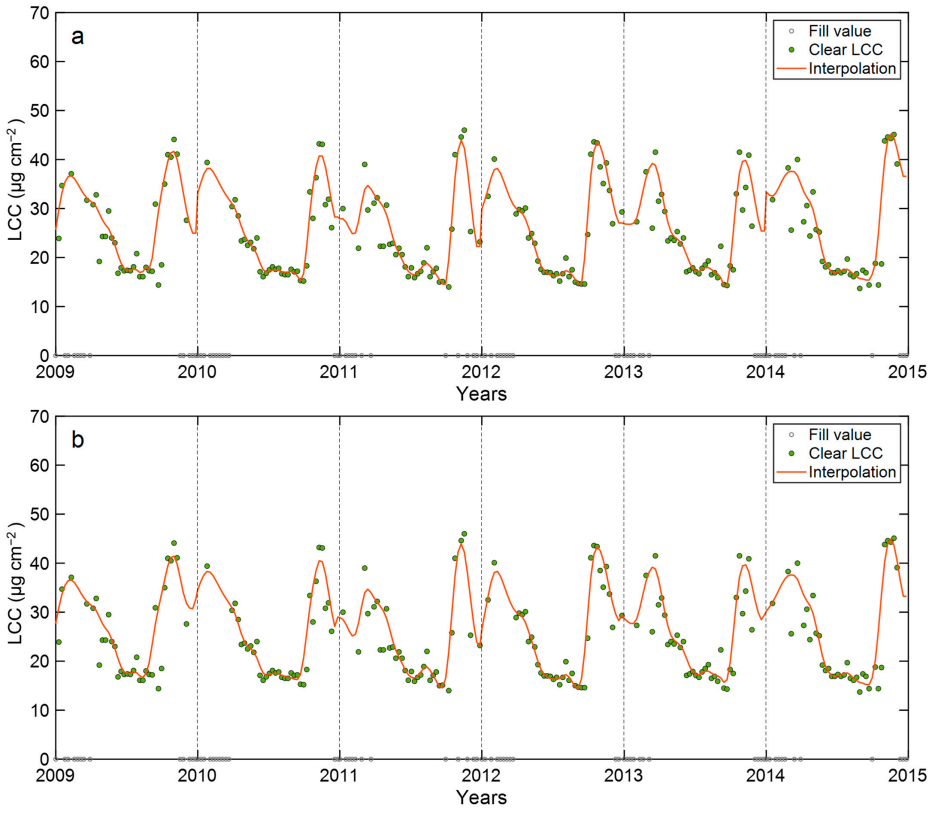


Figure 7. Comparison of the outputted time series by reconstructing the LCC in each year separately (**a**) and reconstructing the LCC in all years together (**b**) using LACC.

Figure 8 shows four site examples of reconstructed time series of LCC in different plant functional types, using the original LACC and LACC2.0 algorithms. LACC2.0 achieved obvious improvements in reconstructing the vegetation growth cycles of LCC in the four plant functional types. In the evergreen broadleaf forest site, using LACC2.0 with the outlier removal and spatial interpolation, the vegetation growth cycles of LCC in 2012, 2014, and 2015 were better reconstructed compared with those reconstructed using LACC. In the evergreen needleleaf forest site, large consecutive gaps appeared in the start and end periods of the vegetation growth cycle in 2016, 2017, 2019, and 2020, causing a significant overestimation of LCC during these periods. By incorporating spatial interpolation before LACC reconstruction, several extra LCC values were derived from neighbor pixels during these gap periods, separating the original large consecutive gaps into smaller gaps. With these spatially-filled LCC data as initial values, the overestimations of the results using the original LACC algorithm were avoided and the vegetation growth cycles were well reconstructed. Similarly, in the shrubland and grassland sites, extra LCC data in vegetation growth seasons derived from neighbor pixels helped improve the reconstruction of the time series of LCC.

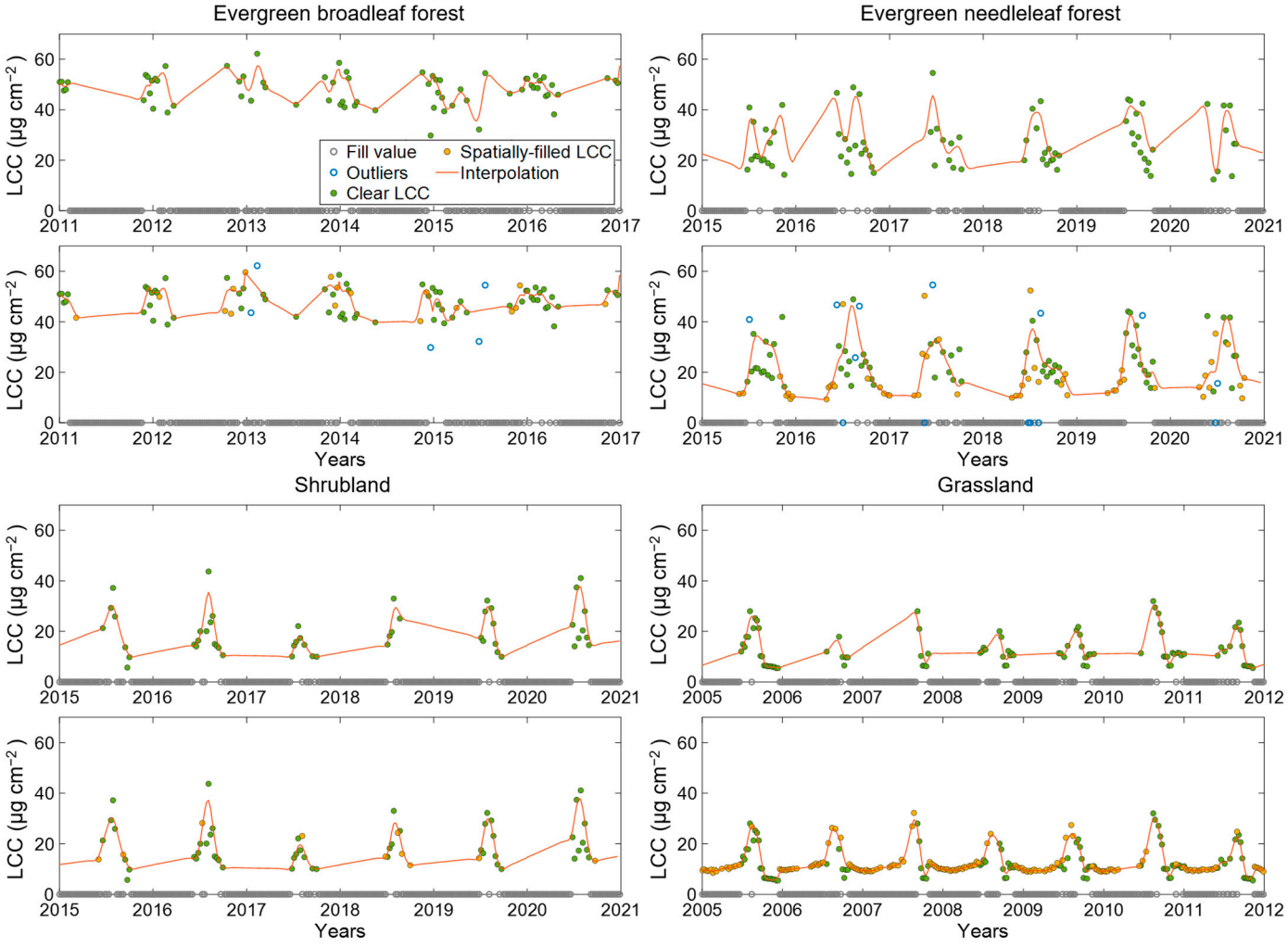


Figure 8. Comparing the reconstruction of the time series of LCC using the original LACC (**upper** figures) and LACC2.0 (**lower** figures) algorithms at sites of four different plant functional types.

3.3. Global Performance of the LACC2.0

The global time series reconstruction using LACC2.0 was evaluated by comparing the modeled time series of LCC and the reference LCC curve pixel by pixel. Figure 9 shows the global RMSE of the time series reconstruction using the original LACC, $LACC_O$, and

LACC2.0. For the original LACC, most regions had RMSE values smaller than 5 μg/cm^2. Large RMSE values (>5 μg/cm^2) were mainly located in the tropical forest area, boreal area, and south Asia, which had larger consecutive gaps as shown in Figure 2. The relative RMSE of LCC was larger than that from the simulated noise evaluation in global time series reconstruction of NDVI [9], because both the negatively biased and positively biased noises were simulated into the test time series of LCC. The original LACC performed well for reducing the negatively biased noises [9] but poorer for reducing the positively biased noises. By incorporating an outlier removal procedure, most noises were screened and their impacts on the reconstruction were reduced, resulting in global smaller RMSEs in LACC$_O$ than that in LACC (Figure 9e). The percentage of pixels with RMSE < 5 μg/cm^2 was improved from 81.2% to 84.5%. By further incorporating spatial interpolation, global RMSEs in the time series reconstruction were largely reduced (Figure 9f) and the percentage of pixels with RMSE < 5 μg/cm^2 was improved to 96.4%. The improvement was most obvious in the tropical area, boreal area, and south Asia, where the gap percentages of LCC were high and consecutive gaps were large (Figures 1 and 2). Despite this, these regions still had larger RMSE (>2 μg/cm^2 mostly) than the other regions using LACC2.0, indicating the residual influence of consecutive gaps.

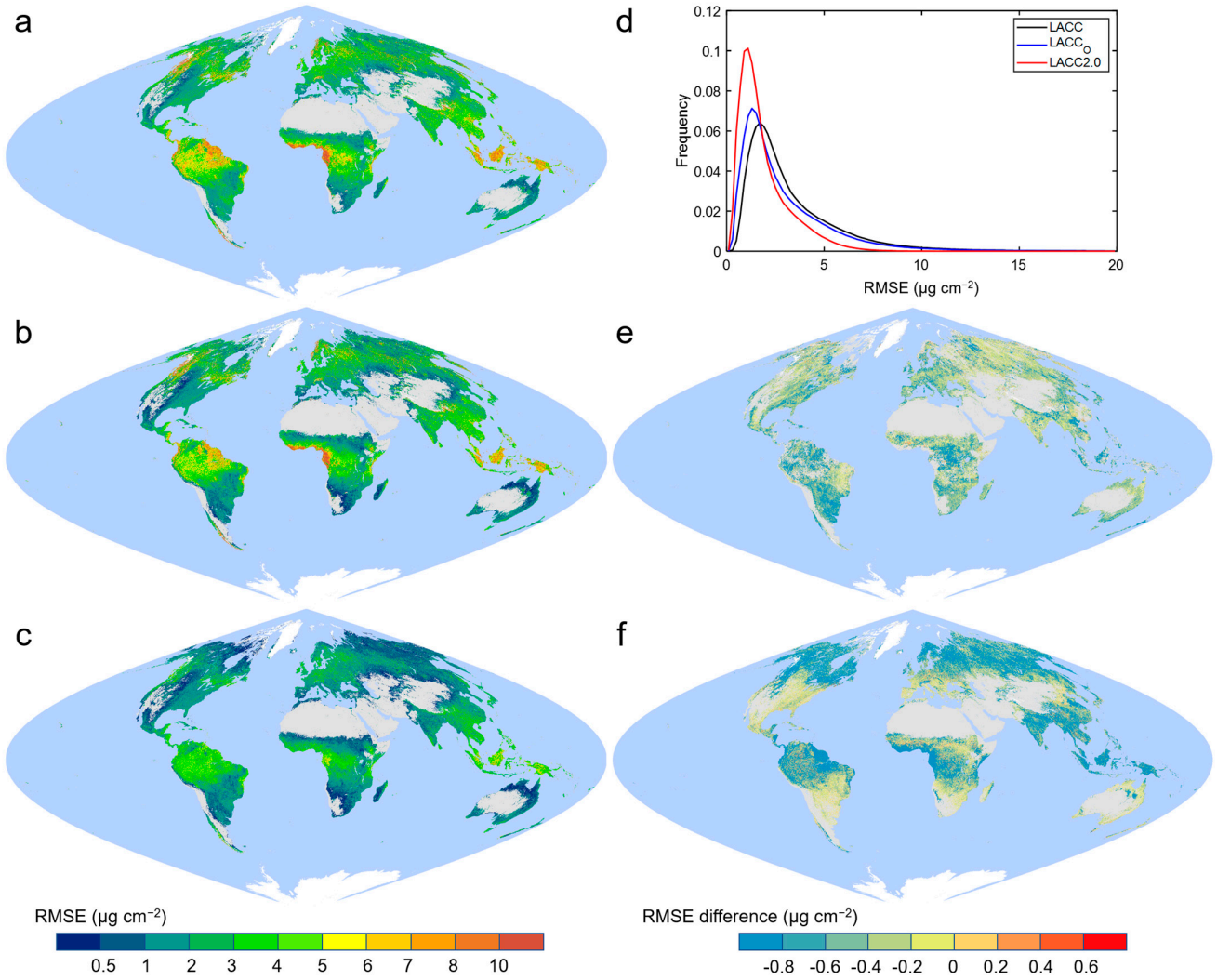


Figure 9. (**a**–**c**) are the global performances of LCC time series reconstruction (indicated by RMSE) using LACC, LACC$_O$, and LACC2.0, respectively. (**d**) shows the frequency of global RMSE in (**a**–**c**). (**e**) is the RMSE difference between LACC$_O$ and LACC, and (**f**) is the RMSE difference between LACC2.0 and LACC$_O$.

4. Discussion

As a commonly-used time series reconstruction algorithm for deriving continuous vegetation biophysical parameters, the LACC algorithm performed the best among six algorithms in the global reconstruction of vegetation parameters by considering vegetation growth trajectory, protection of key points in the seasonal curve, noise resistance, and curve stability [9]. In this study, the original LACC algorithm was modified by incorporating an outlier removal procedure and integrating spatial information for gap filling in order to achieve better reconstruction of vegetation biochemical parameters. Using a global LCC product as an example for validation, the percentage of LCC pixels with an RMSE smaller than 5 $\mu g/cm^2$ was improved from 81.2% in the original LACC to 96.4% in the improved LACC (LACC2.0), demonstrating that LACC2.0 had the potential to provide a better reconstruction of global daily vegetation biochemical parameters.

The presence of outliers or noises in time series data can significantly affect the accuracy and reliability of the gap-filling process [37,38], and it is crucial to develop robust techniques to mitigate the influence of outliers in order to fill the gaps accurately. In this study, a time series filter proposed by Shang et al. [35] to identify abrupt and ephemeral data in a time series was used to screen outliers. This filter had shown good performance in improving the time series reconstruction. However, a short and abrupt peak in the growing season in some vegetation (e.g., some grasslands and croplands) may be mistakenly identified as an outlier, especially when data before and after the peak are also missing. This could be improved by adjusting the threshold according to the vegetation phenology [33]. Additionally, the MOD09A1 surface reflectance data was utilized to eliminate outliers instead of relying solely on the LCC. This choice was made because the MOD09A1 product provides a more comprehensive set of outlier information compared to the single time series of the LCC. In cases where corresponding surface reflectance data is not available, the vegetation biochemical parameter can be employed for outlier removal. In such situations, the accuracy of the final time series reconstruction may not be compromised. There are also other algorithms [39–41] that could identify outliers. They can be compared, and an optimal method can be chosen for future modifications.

Long consecutive gaps strongly affect the temporal continuity of a time series, and are particularly challenging for temporal gap filling [9]. In this study, the DISKR algorithm was used to incorporate the spatial information for time series reconstruction in LACC2.0 and achieved high validation accuracy. However, new outliers can be introduced using spatial information, adding uncertainty to the reconstructed time series, because even neighboring pixels with the same land cover type can have different curves of vegetation biochemical parameters. For example, a cropland area may be mixed with different crop types, which obviously have a different time series of LCC. In this situation, the neighboring information may be misleading for time series reconstruction. Additionally, despite the success of the DISKR algorithm in reducing the percentage of long consecutive gaps, limitations remain in some areas due to the persistent cloud and snow covers [42,43]. These remaining long consecutive gaps can impede our ability to accurately capture complete trajectories of vegetation growth, particularly when the gaps occur during periods of rapid vegetation growth [9]. In future modifications, generating an initial curve of vegetation growth trajectory considering the vegetation phenology [33,44] and land changes [34,35] may be helpful to minimize the influence of long consecutive gaps.

5. Conclusions

In this study, we presented LACC2.0, an improved version of the locally adjusted cubic-spline capping (LACC) algorithm for global time series reconstruction of satellite-derived vegetation biochemical parameters by incorporating the outlier removal procedure and integrating the spatial information for filling large gaps. The improved algorithm with these modifications was applied to a global product of leaf chlorophyll content (LCC) as an example, and it is shown that the improved algorithm performed obviously better than the original algorithm in reconstructing the time series of LCC at typical sites. For global

evaluation, a reference LCC curve was generated as the true value, and the modeled time series of LCC with real gaps and Gaussian noises was created for reconstruction. Results showed that the percentage of pixels with an RMSE smaller than 5 $\mu g/cm^2$ was improved from 81.2% in LACC to 96.4% in LACC2.0, demonstrating that LACC2.0 had the potential to provide a better reconstruction of global daily satellite-derived vegetation biochemical parameters. This finding highlights the significance of outlier removal and spatial-temporal fusion to enhance the accuracy and reliability of time series reconstruction.

Author Contributions: Conceptualization, M.X. and R.S.; Methodology, M.X., R.S. and L.Z.; Validation, M.X.; Formal analysis, M.X.; Writing—original draft, M.X. and R.S.; Writing—review & editing, M.X., R.S., J.M.C. and L.Z.; Funding acquisition, M.X. and R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant number 42201360 and 42101367), China Postdoctoral Science Foundation (grant number 2021M690638), Natural Science Foundation of Fujian Province (grant number 2021J05041), Fujian Forestry Science and Technology Key Project (grant number 2022FKJ03), and Open Fund Project of the Academy of Carbon Neutrality of Fujian Normal University (grant number TZH2022-02).

Data Availability Statement: Not applicable.

Acknowledgments: All authors are grateful to the anonymous reviewers and editors for appraising our manuscript and for offering instructive comments.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The influence of using different values of outlier probability (Op) in the chi-squared distribution on the reconstruction of LCC time series at 212 FLUXNET2015 sites. The mean RMSE between the reference LCC curve and reconstructed time series was calculated for all 212 sites. The Op value with the lowest mean RMSE was regarded as the optimal value.

Op	0.9	0.95	0.99	0.995	0.999
Mean RMSE ($\mu g\ cm^{-2}$)	2.96	2.44	2.03	2.21	2.57

Table A2. The influence of using different window sizes in integrating the spatial information on the reconstruction of LCC time series at 212 FLUXNET2015 sites. The mean RMSE between the reference LCC curve and reconstructed time series was calculated for all 212 sites. The window size with the lowest mean RMSE was regarded as the optimal value.

Window size	5 × 5	7 × 7	9 × 9	11 × 11	13 × 13
Mean RMSE ($\mu g\ cm^{-2}$)	2.52	2.37	2.14	2.03	2.09

References

1. Chen, J.M.; Ju, W.; Ciais, P.; Viovy, N.; Liu, R.; Liu, Y.; Lu, X. Vegetation Structural Change since 1981 Significantly Enhanced the Terrestrial Carbon Sink. *Nat. Commun.* **2019**, *10*, 4259. [CrossRef]
2. Liu, Y.; Liu, R.; Chen, J.M. Retrospective Retrieval of Long-Term Consistent Global Leaf Area Index (1981–2011) from Combined AVHRR and MODIS Data. *J. Geophys. Res.* **2012**, *117*, 14. [CrossRef]
3. Xu, M.; Liu, R.; Chen, J.M.; Liu, Y.; Wolanin, A.; Croft, H.; He, L.; Shang, R.; Ju, W.; Zhang, Y.; et al. A 21-Year Time-Series of Global Leaf Chlorophyll Content Maps from MODIS Imagery. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4413513. [CrossRef]
4. Xu, M.; Liu, R.; Chen, J.M.; Shang, R.; Liu, Y.; Qi, L.; Croft, H.; Ju, W.; Zhang, Y.; He, Y.; et al. Retrieving Global Leaf Chlorophyll Content from MERIS Data Using a Neural Network Method. *ISPRS J. Photogramm. Remote Sens.* **2022**, *192*, 66–82. [CrossRef]
5. Wei, S.; Fang, H. Estimation of Canopy Clumping Index from MISR and MODIS Sensors Using the Normalized Difference Hotspot and Darkspot (NDHD) Method: The Influence of BRDF Models and Solar Zenith Angle. *Remote Sens. Environ.* **2016**, *187*, 476–491. [CrossRef]

6. Chen, J.M.; Deng, F.; Chen, M.Z. Locally Adjusted Cubic-Spline Capping for Reconstructing Seasonal Trajectories of a Satellite-Derived Surface Parameter. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2230–2238. [CrossRef]
7. Kutsch, W.L.; Kolari, P. Data Quality and the Role of Nutrients in Forest Carbon-Use Efficiency. *Nat. Clim. Chang.* **2015**, *6*, 959–960. [CrossRef]
8. Liu, R.; Chen, J.M.; Liu, J.; Deng, F.; Sun, R. Application of a New Leaf Area Index Algorithm to China's Landmass Using MODIS Data for Carbon Cycle Research. *J. Environ. Manag.* **2007**, *85*, 649–658. [CrossRef]
9. Liu, R.; Shang, R.; Liu, Y.; Lu, X. 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.* **2017**, *189*, 164–179. [CrossRef]
10. Moreno-Martínez, Á.; Izquierdo-Verdiguier, E.; Maneta, M.P.; Camps-Valls, G.; Robinson, N.; Muñoz-Marí, J.; Sedano, F.; Clinton, N.; Running, S.W. Multispectral High Resolution Sensor Fusion for Smoothing and Gap-Filling in the Cloud. *Remote Sens. Environ.* **2020**, *247*, 111901. [CrossRef]
11. Pipia, L.; Muñoz-Marí, J.; Amin, E.; Belda, S.; Camps-Valls, G.; Verrelst, J. Fusing Optical and SAR Time Series for LAI Gap Fillingwith Multioutput Gaussian Processes. *Remote Sens. Environ.* **2019**, *235*, 111452. [CrossRef]
12. Wang, J.; Lee, C.K.F.; Zhu, X.; Cao, R.; Gu, Y.; Wu, S.; Wu, J. A New Object-Class Based Gap-Filling Method for PlanetScope Satellite Image Time Series. *Remote Sens. Environ.* **2022**, *280*, 113136. [CrossRef]
13. Ma, M.; Veroustraete, F. Reconstructing Pathfinder AVHRR Land NDVI Time-Series Data for the Northwest of China. *Adv. Space Res.* **2006**, *37*, 835–840. [CrossRef]
14. Julien, Y.; Sobrino, J.A. Comparison of Cloud-Reconstruction Methods for Time Series of Composite NDVI Data. *Remote Sens. Environ.* **2010**, *114*, 618–625. [CrossRef]
15. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A Simple Method for Reconstructing a High-Quality NDVI Time-Series Data Set Based on the Savitzky-Golay Filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [CrossRef]
16. Zhu, W.; Pan, Y.; He, H.; Wang, L.; Mou, M.; Liu, J. A Changing-Weight Filter Method for Reconstructing a High-Quality NDVI Time Series to Preserve the Integrity of Vegetation Phenology. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 1085–1094. [CrossRef]
17. Beck, P.S.A.; Atzberger, C.; Høgda, K.A.; Johansen, B.; Skidmore, A.K. Improved Monitoring of Vegetation Dynamics at Very High Latitudes: A New Method Using MODIS NDVI. *Remote Sens. Environ.* **2006**, *100*, 321–334. [CrossRef]
18. Jönsson, P.; Eklundh, L. Seasonality Extraction by Function Fitting to Time-Series of Satellite Sensor Data. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1824–1832. [CrossRef]
19. Dash, J.; Jeganathan, C.; Atkinson, P.M. The Use of MERIS Terrestrial Chlorophyll Index to Study Spatio-Temporal Variation in Vegetation Phenology over India. *Remote Sens. Environ.* **2010**, *114*, 1388–1402. [CrossRef]
20. Zhou, J.; Jia, L.; Menenti, M. Reconstruction of Global MODIS NDVI Time Series: Performance of Harmonic ANalysis of Time Series (HANTS). *Remote Sens. Environ.* **2015**, *163*, 217–228. [CrossRef]
21. Atzberger, C.; Eilers, P.H.C. A Time Series for Monitoring Vegetation Activity and Phenology at 10-Daily Time Steps Covering Large Parts of South America. *Int. J. Digit. Earth* **2011**, *4*, 365–386. [CrossRef]
22. Croft, H.; Chen, J.M.; Wang, R.; Mo, G.; Luo, S.; Luo, X.; He, L.; Gonsamo, A.; Arabian, J.; Zhang, Y.; et al. The Global Distribution of Leaf Chlorophyll Content. *Remote Sens. Environ.* **2020**, *236*, 111479. [CrossRef]
23. Xu, M.; Liu, R.; Chen, J.M.; Liu, Y.; Shang, R.; Ju, W.; Wu, C.; Huang, W. Retrieving Leaf Chlorophyll Content Using a Matrix-Based Vegetation Index Combination Approach. *Remote Sens. Environ.* **2019**, *224*, 60–73. [CrossRef]
24. Chu, D.; Shen, H.; Guan, X.; Chen, J.M.; Li, X.; Li, J.; Zhang, L. Long Time-Series NDVI Reconstruction in Cloud-Prone Regions via Spatio-Temporal Tensor Completion. *Remote Sens. Environ.* **2021**, *264*, 112632. [CrossRef]
25. Poggio, L.; Gimona, A.; Brown, I. Spatio-Temporal MODIS EVI Gap Filling under Cloud Cover: An Example in Scotland. *ISPRS J. Photogramm. Remote Sens.* **2012**, *72*, 56–72. [CrossRef]
26. Wang, Q.; Atkinson, P.M. Spatio-Temporal Fusion for Daily Sentinel-2 Images. *Remote Sens. Environ.* **2018**, *204*, 31–42. [CrossRef]
27. Yan, L.; Roy, D.P. Large-Area Gap Filling of Landsat Reflectance Time Series by Spectral-Angle-Mapper Based Spatio-Temporal Similarity (SAMSTS). *Remote Sens.* **2018**, *10*, 609. [CrossRef]
28. Shi, W.; Guo, D.; Zhang, H. A Reliable and Adaptive Spatiotemporal Data Fusion Method for Blending Multi-Spatiotemporal-Resolution Satellite Images. *Remote Sens. Environ.* **2022**, *268*, 112770. [CrossRef]
29. Yang, X.; Chen, J.; Guan, Q.; Gao, H.; Xia, W. Enhanced Spatial-Temporal Savitzky-Golay Method for Reconstructing High-Quality NDVI Time Series: Reduced Sensitivity to Quality Flags and Improved Computational Efficiency. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4412317. [CrossRef]
30. Cao, R.; Chen, Y.; Shen, M.; Chen, J.; Zhou, J.; Wang, C.; Yang, W. A Simple Method to Improve the Quality of NDVI Time-Series Data by Integrating Spatiotemporal Information with the Savitzky-Golay Filter. *Remote Sens. Environ.* **2018**, *217*, 244–257. [CrossRef]
31. Liu, R.; Liu, Y. Generation of New Cloud Masks from MODIS Land Surface Reflectance Products. *Remote Sens. Environ.* **2013**, *133*, 21–37. [CrossRef]
32. Bontemps, S.; Defourny, P.; Radoux, J.; Van Bogaert, E.; Lamarche, C.; Achard, F.; Mayaux, P.; Boettcher, M.; Brockmann, C.; Kirches, G. Consistent Global Land Cover Maps for Climate Modelling Communities: Current Achievements of the ESA's Land Cover CCI. In Proceedings of the ESA Living Planet Symposium, Edinburgh, UK, 9–13 September 2013.

33. Shang, R.; Liu, R.; Xu, M.; Liu, Y.; Dash, J.; Ge, Q. Determining the Start of the Growing Season from MODIS Data in the Indian Monsoon Region: Identifying Available Data in the Rainy Season and Modeling the Varied Vegetation Growth Trajectories. *Remote Sens.* **2018**, *10*, 122. [CrossRef]
34. Zhu, Z.; Zhang, J.; Yang, Z.; Aljaddani, A.H.; Cohen, W.B.; Qiu, S.; Zhou, C. Continuous Monitoring of Land Disturbance Based on Landsat Time Series. *Remote Sens. Environ.* **2020**, *238*, 111116. [CrossRef]
35. Shang, R.; Zhu, Z.; Zhang, J.; Qiu, S.; Yang, Z.; Li, T.; Yang, X. Near-Real-Time Monitoring of Land Disturbance with Harmonized Landsats 7–8 and Sentinel-2 Data. *Remote Sens. Environ.* **2022**, *278*, 113073. [CrossRef]
36. Takeda, H.; Farsiu, S.; Milanfar, P. Kernel Regression for Image Processing and Reconstruction. *IEEE Trans. Image Process.* **2007**, *16*, 349–366. [CrossRef]
37. Hird, J.N.; McDermid, G.J. Noise Reduction of NDVI Time Series: An Empirical Comparison of Selected Techniques. *Remote Sens. Environ.* **2009**, *113*, 248–258. [CrossRef]
38. Moreno, Á.; García-Haro, F.J.; Martínez, B.; Gilabert, M.A. Noise Reduction and Gap Filling of FAPAR Time Series Using an Adapted Local Regression Filter. *Remote Sens.* **2014**, *6*, 8238–8260. [CrossRef]
39. Cilli, R.; Monaco, A.; Amoroso, N.; Tateo, A.; Tangaro, S.; Bellotti, R. Machine Learning for Cloud Detection of Globally Distributed Sentinel-2 Images. *Remote Sens.* **2020**, *12*, 2355. [CrossRef]
40. Foga, S.; Scaramuzza, P.L.; Guo, S.; Zhu, Z.; Dilley, R.D.; Beckmann, T.; Schmidt, G.L.; Dwyer, J.L.; Joseph Hughes, M.; Laue, B. Cloud Detection Algorithm Comparison and Validation for Operational Landsat Data Products. *Remote Sens. Environ.* **2017**, *194*, 379–390. [CrossRef]
41. Tarrio, K.; Tang, X.; Masek, J.G.; Claverie, M.; Ju, J.; Qiu, S.; Zhu, Z.; Woodcock, C.E. Comparison of Cloud Detection Algorithms for Sentinel-2 Imagery. *Sci. Remote Sens.* **2020**, *2*, 100010. [CrossRef]
42. Ding, C.; Liu, X.; Huang, F. Temporal Interpolation of Satellite-Derived Leaf Area Index Time Series by Introducing Spatial-Temporal Constraints for Heterogeneous Grasslands. *Remote Sens.* **2017**, *9*, 968. [CrossRef]
43. Liu, C.; Zhang, Q.; Luo, H.; Qi, S.; Tao, S.; Xu, H.; Yao, Y. An Efficient Approach to Capture Continuous Impervious Surface Dynamics Using Spatial-Temporal Rules and Dense Landsat Time Series Stacks. *Remote Sens. Environ.* **2019**, *229*, 114–132. [CrossRef]
44. Shang, R.; Liu, R.; Xu, M.; Liu, Y.; Zuo, L.; Ge, Q. The Relationship between Threshold-Based and Inflexion-Based Approaches for Extraction of Land Surface Phenology. *Remote Sens. Environ.* **2017**, *199*, 167–170. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.