



Technical Note Prediction of High-Quality MODIS-NPP Product Data

Zhenhua Liu¹, Ting Wang¹, Yonghua Qu^{2,3,4}, Huiming Liu¹, Xiaofang Wu¹ and Ya Wen^{1,*}

- ¹ College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, China; zhenhua@scau.edu.cn (Z.L.); wt@stu.scau.edu.cn (T.W.); huimingliu@scau.edu.cn (H.L.); wuxiaofang@scau.edu.cn (X.W.)
- ² State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications of CAS, Beijing 100875, China; qyh@bnu.edu.cn
- ³ Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing Normal University, Beijing 100875, China
- ⁴ School of Geography, Beijing Normal University, Beijing 100875, China
- * Correspondence: wenya26@scau.edu.cn; Tel.: +86-135-3343-2093

Received: 25 April 2019; Accepted: 18 June 2019; Published: 20 June 2019



Abstract: Net primary productivity (NPP) is a key vegetation parameter and ecological indicator for tracking natural environmental change. High-quality Moderate Resolution Imaging Spectroradiometer Net primary productivity (MODIS-NPP) products are critical for assuring the scientific rigor of NPP analyses. However, obtaining high-quality MODIS-NPP products consistently is challenged by factors such as cloud contamination, heavy aerosol pollution, and atmospheric variability. This paper proposes a method combining the discrete wavelet transform (DWT) with an extended Kalman filter (EKF) for generating high-quality MODIS-NPP data. In this method, the DWT is used to remove noise in the original MODIS-NPP data, and the EKF is applied to the de-noised images. The de-noised images are modeled as a triply modulated cosine function that predicts the NPP data values when excessive cloudiness is present. This study was conducted in South China. By comparing measured NPP data to original MODIS-NPP and NPP estimates derived from combining the DWT and EKF, we found that the accuracy of the NPP estimates was significantly improved. The MODIS-NPP estimates had a mean relative error (RE) of 13.96% and relative root mean square error (rRMSE) of 15.67%, while the original MODIS-NPP had a mean RE of 23.58% and an rRMSE of 24.98%. The method combining DWT and EKF provides a feasible approach for generating new, high-quality NPP data in the absence of high-quality original MODIS-NPP data.

Keywords: high-quality MODIS-NPP; EKF; discrete wavelet transform

1. Introduction

Net primary productivity (NPP) is a key variable in the carbon exchange between the biosphere and the atmosphere [1]. NPP is defined as the amount of above- and below-ground organic matter accumulated in vegetation per unit of area during a certain time period (usually a year), representing the productivity of an ecosystem [2]. NPP with 1 km spatial resolution and an 8 day interval has been produced operationally with the Moderate Resolution Imaging Spectroradiometer Net Primary Productivity product (MOD17) algorithm based on observations from the MODIS sensor (National Aeronautics and Space Administration, America) [3]. These MODIS-NPP data provide consistent spatial and temporal measures of crop yield, range forage, forest production, and other socio-economically significant products related to vegetation growth [4,5]. Thus, accurate NPP estimation is critical. The MOD17 algorithm was developed as a result of past achievements and lessons learned from a general ecosystem model, Biome-BGC [5]. A detailed description of the algorithm can be found elsewhere [6,7]. Although the MOD17 algorithm has undergone several improvements, MODIS-NPP

2 of 12

satellite data still include various sources of noise, such as clouds, atmospheric variability, and bidirectional reflectance distribution factors [8,9]. These distribution factors can affect the monitoring of crop yield, range forage, and forests. Therefore, it is crucial to reduce noise in images and obtain high-quality NPP data.

Several methods for reducing noise and constructing high-quality MODIS datasets have been proposed, applied, and evaluated in recent years [10–16]. Fast Fourier transform (FFT) and wavelet transform (WT) approaches are often used when evaluating MODIS time-series data [17–20]. Several mathematical filters have been applied to reduce noise and obtain high-quality time-series of remote sensing datasets [21,22]. Recently, data assimilation methods have also been used to reconstruct remote-sensing data dynamically. Gu et al. [23] adopted an optimal interpolation method to generate appropriate Moderate Resolution Imaging Spectroradiometer leaf area index (MODIS-LAI) data. Gu et al. [24] developed a simplified data assimilation method based on Normalized Difference Vegetation Index (NDVI) quality assurance (QA) datasets to reconstruct high-quality, spatially, and temporally continuous MODIS NDVI products. However, these methods also suffer from drawbacks. For example, they can only produce a high-quality output when the data contain little noise, so they are challenged when used with very cloudy images.

The objective of this study was to generate high-quality MODIS-NPP data by combining discrete wavelet transform (DWT) with an extended Kalman filter (EKF). Taking an area in South China as an example, we used the DWT to reduce cloud and atmospheric variability contamination in MODIS-NPP data. Then, the EKF was applied to the de-noised MODIS-NPP data to predict NPP data when the original data was not valid.

2. Materials

The study area was located in South China, and covered approximately 65,536 km² (256 km \times 256 km) within N22°36′30″–N24°44′0″ and E112°53′21.25″–E113°22′13.28″ (Figure 1). The area has a humid, subtropical, monsoon climate characterized by warm winters, hot summers, little frost and snow, and sufficient rain and sunshine, with an annual mean temperature of 22 °C. The annual average precipitation is 1400–2000 mm. The main vegetation types in the area are subtropical evergreen broadleaf forests, evergreen needleleaf forests, savannas, and common crops such as rice, peanut, and rape (Figure 1).



Figure 1. Location of the study area: (**a**) map of China with highlight of study area; (**b**) 500 m spatial resolution land cover map from the MODIS Land Cover Type Product (MCD12) showing the spatial distribution of sample plots (red solid plots are representative samples).

NPP measurements were taken from a total of 20 sample plots located in evergreen broadleaf and evergreen needle leaf forests within the study area in October 2013 (Figure 1). Each of the sample plots had an area of 1 km × 1 km, matching the spatial resolution of the MODIS products. Within each sample plot, the tree species, height, and diameter at breast height (DBH), with breast height being 130 cm, and age class (young, middle age, and mature) were measured. Estimates of above-ground tree biomass (AGB) were predicted with empirical regression models using given the following general formula:

$$AGB = a \times DBH^b, \tag{1}$$

where *a* and *b* are empirical coefficients that vary by tree species [25–28]. Estimates of tree biomass were calculated using the ratio of AGB to the below-ground tree biomass [29]. Estimates of tree biomass in the sample plots were obtained by adding the tree biomass values within each plot. The value of NPP for each sample plot was estimated using the plot biomass estimate divided by the average age, in years, of the trees within each plot [30]. The NPP within 1 km was an average of five sample NPPs. The average age of trees for each plot was estimated by calculating a weighted mean of the average age classes and the number of trees within each age class. According to previous studies [31], the six representative sample plots, including *Schima, Pinus massoniana, Acacia confusa, Camphora, Ficus,* and *Eucalyptus,* were selected to collect the measured NPP data, from 18 October 2013 to 27 December 2015, as shown in Figure 1.

To evaluate the performance of the method proposed here, we used 8 day composites of MOD17 data with a spatial resolution of 1 km from 9 January 2013 to 18 December 2015. The original MODIS-NPP data composites with heavy cloud cover do not provide useful data for predicting high-quality MODIS-NPP data during periods of active phenological change. Thus, the time-distance weight function was used to obtain these NPP data using NPP composites with no heavy clouds at adjoining time periods. The time-distance weight function can be expressed as

$$NPP_i = c \times NPP_k + d \times NPP_i, \tag{2}$$

where NPP_i is the *i*th day NPP, NPP_k represents the original MODIS-NPP value with no heavy clouds on the closest *k*th day preceding the *i*th day, and NPP_j represents the original MODIS-NPP value with no heavy clouds on the closest *j*th day after the *i*th day. The time distance weight, equal to (i - k)/(j - k) is *c*, and *d* is the time distance weight, equal to (j - i)/(j - k).

3. Methods

3.1. The Discrete Wavelet Transform for De-Noising MODIS-NPP Images

The purpose of de-noising original MODIS-NPP data is to reduce the noise in the DWT domain while preserving the NPP characteristics. The *N* NPP values of a noisy image are represented as a two-dimensional matrix:

$$\begin{bmatrix} y_{ij} \end{bmatrix}, \ i = 1 \cdots N, \ j = 1 \cdots M, \tag{3}$$

where *N* and *M* denote the number of rows and columns in the MODIS-NPP image. The noisy version of the image is modelled as [32]

$$y_{ij} = x_{ij} + n_{ij} \ i = 1 \cdots N, \ j = 1 \cdots M,$$
 (4)

where $[n_{ij}]$ are *iid* as $N(0, \sigma^2)$ and x_{ij} denotes the de-noised data value.

Wavelet de-noising involved three main steps: calculation of the DWT of the image, completion of a nonlinear thresholding step, and completion of a linear inverse DWT to get the de-noised estimate. One of the most important steps in the de-noising development process was finding the optimal threshold. Encouraged by results in recent studies [33–37], we tested six thresholding functions (Figure 2) to de-noise the original MODIS-NPP images in this study.



Figure 2. Threshold functions to estimate the wavelet coefficients: (**a**) soft threshold (ST); (**b**) hard threshold (HT); (**c**) half-soft threshold (HST); (**d**) improved semi-soft and half-soft threshold (ISHST); (**e**) improved soft and hard threshold (ISHT); and (**f**) improved threshold based on the ST and the ISHT (ISIT).

Taking a MODIS-NPP image from 9 January 2013 as an example, the signal-to-noise ratios (SNRs) and root mean square difference (RMSDs) from de-noising the image using the six thresholds varied from 9.02 to 17.20 Mg ha⁻¹ a⁻¹ and from 0.38 to 0.99, respectively (Table 1). The highest SNR and lowest RMSE were obtained using the improved soft and hard threshold (ISHT) threshold function, suggesting that this was the most suitable thresholding function for de-noising the images. Thus, the ISHT threshold function was selected as the optimal threshold in this study.

Threshold function	SNR	RMSD		
	10.11	0.40		
ST	12.11	0.69		
HT	9.02	0.99		
HST	10.86	0.82		
ISHST	9.95	0.89		
ISHT	17.20	0.38		
ISIT	11.38	0.75		

Table 1. Signal-to-noise ratios (SNRs) and root mean square difference (RMSDs) for the different threshold functions.

3.2. The Extended Kalman Filter for Predicting High-Quality MODIS-NPP Data

The formula to predict MODIS-NPP time series in a given pixel is written as [38]

$$y_k = \mu_k + \alpha_k \cos(\omega_k + \Phi_k) + v_k \tag{5}$$

where the subindex *k* indicates the time step, y_k represents the observed value of the predicted MODIS-NPP time series, v_k is the noise sample, ω_k denotes the angular frequency, α_k is the amplitude, μ_k is the mean, and Φ_k is the phase. The angular frequency can be calculated as $\omega_k = 2\pi f$, where *f* is determined by the annual vegetation growth cycle. Because MODIS-NPP images were calculated based on the MODIS combined Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) products (MOD15A2 products), *f* was set at 8/365. The values of μ_k , α_k and Φ_k are functions of time, and must be estimated given y_k for $k \in 1, ..., N$. The estimation of these parameters is nontrivial and

requires a non-linear estimator. Thus, the EKF was driven to estimate the three parameters for every increment of *k* [38].

The EKF is considered the standard framework to solve estimation problems in navigation [39]. The EKF prediction formula can be obtained by linearizing the process model. The observation model at time k + 1 is given by [39]

$$y_{k+1} = H \cdot x_{k+1} + v_k, \tag{6}$$

where y_{k+1} is the estimated NPP and x_{k+1} denotes the a priori estimate, defined as $x_{k+1} = [\mu_{k+1} \alpha_{k+1} \Phi_{k+1}]^T$. *H* is an observation operator used to relate model state variables to observations (it does not need be linear), and v_k is the zero-mean Gaussian observation noise $v_k \sim N(O, R)$. The discrete-time prediction EKF algorithm has two main stages, namely a prediction and an update (filtering) step as follows [39]:

Prediction equation

The predicted state estimate is defined as:

$$x_{k+1} = F \cdot \hat{x}_k + w_k. \tag{7}$$

The predicted covariance estimate is defined as

$$P_{k+1} = \nabla F_x \cdot \hat{P}_k \nabla F_x^T + Q_k, \tag{8}$$

where *F* denotes the state transition matrix, P_{k+1} is the a priori estimate error covariance matrix, \hat{x}_k represents the a posteriori estimate, \hat{P}_k is the a posteriori error covariance matrix, w_k is the zero-mean Gaussian process noise $w_k \sim N(O, Q_k)$, ∇F_x is the Jacobian of function *F* with respect to *x* evaluated at \hat{x}_k , and Q_k is the covariance matrix of the process noises.

Update equation

The updated state estimate is defined as

$$\hat{x}_{k+1} = x_k + K_{g(k)}(y_k - H_k x_k).$$
(9)

The near-optimal Kalman gain is defined as

$$K_g(k) = P_k \nabla H_k^T \left[\nabla H_k P_k \nabla H_k^T + R_k \right]^{-1}.$$
 (10)

The updated covariance estimate is defined as

$$\hat{P}_{k+1} = \left[I - K_g(k) \nabla H_k\right] P_k,\tag{11}$$

where $K_g(k)$ is the Kalman gain matrix at instant k, ∇H_k is the Jacobian of function H evaluated at x_k , R_k is the covariance matrix of the measurement noises, and I is the identity matrix.

The values of μ_k , α_k , and Φ_k are functions of time and must be estimated given y_k for $k \in 1$. This study used 137 time-series images so that M = 137. The a priori initial state parameters ($\mu_1 \alpha_1, \Phi_1$) in state vector (x_1) were set as

$$\mu_1 = \sum_{i=1}^N \frac{\text{NPP}_i}{N},\tag{12}$$

$$a_1 = \frac{\text{NPP}_{\text{max}} - \text{NPP}_{\text{min}}}{2},\tag{13}$$

where *N* is the number of remote-sensing pixels. NPP_{max} and NPP_{min} are the maximum and minimum NPP values in the studied area, respectively. An a priori initial estimate $\Phi_1 = 0^{\circ}$ was chosen assuming that the start date of 18 January 2013 represented a stage after vegetation was exchanged. Regarding parameters *R*, *P*₁, and *Q*₁, their values were first taken from the literature [39] and then tested and

adjusted using the EKF model and the layout of NPP samples. Thus, they were finally set as R = 0.01, $P_1 = \text{diag}[0.015, 0.015, 0.02963]$, $Q_1 = \text{diag}[0.01, 0.001, 0.03441]$, and $x_1 = [4.5612, 2.98425, 0]$.

3.3. Evaluation of the NPP Accuracy

The accuracy of the original MODIS-NPP data, de-noised MODIS-NPP data, and predicted MODIS-NPP data was evaluated by comparing the NPP in the 20 sample plots (Section 2) with the NPP in the pixels where the plots were located. The relative error (RE) was calculated as follows for each of the 20 plots in the dates when the ground measurements were performed:

$$RE = 100 \times \frac{|\hat{y} - y|}{y}, \tag{14}$$

where \hat{y} is the MODIS-NPP and y is the plot NPP. For the six representative plots, the RE was also analyzed across time. Then, the relative root mean square error (rRMSE) for the 20 plots as a whole was also calculated:

$$\mathrm{rRMSE} = 100 \times \frac{\sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}}{\overline{y}},\tag{15}$$

where \hat{y}_i is the MODIS-NPP, y_i is the plot NPP, \overline{y} is mean value of the plot NPP and *i* indicates the plot.

The NPP products elaborated in this study (de-noised MODIS-NPP data and predicted MODIS-NPP data) were also compared to the original MODIS-NPP data by means of the root mean square difference (RMSD) and, for the specific case of the de-noised MODIS-NPP data, through the signal-to-noise ratio (SNR):

RMSD =
$$\sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}$$
, (16)

where \hat{y}_i is the NPP product elaborated in this study, and y_i is the original MODIS-NPP.

SNR = 10 * log₁₀
$$\left[\frac{\sum_{i=1}^{M} \sum_{j=1}^{N} g(i,j)^{2}}{\sum_{i=1}^{M} \sum_{j=1}^{N} [g(i,j) - f(i,j)]^{2}} \right]$$
, (17)

where, g(i, j) represents original MODIS-NPP, f(i, j) is de-noised MODIS-NPP.

4. Results

4.1. Image De-Noising

In this study, we used the DWT algorithm to reduce the noise in the MODIS-NPP images and improve the data quality of MODIS imagery. The wavelet decomposition of an image by DWT is shown in Figure 3.



Figure 3. Wavelet decomposition of the MODIS image from 1 January 2014: (**a**) one-scale wavelet decomposition, (**b**) two-scale wavelet decomposition, and (**c**) three-scale wavelet decomposition.

The wavelet decomposition of a MODIS image using DWT (Figure 3) was completed as follows [31]: At the first level of decomposition, the image was divided into four sub-bands, HH_k , HL_k , LH_k , and LL_k , $k = 1 \cdots J$, where k is the scale, with J being the largest (or coarsest) scale in the decomposition. The HH_k sub-band represented the diagonal structures of the image, the HL_k sub-band represented the horizontal features, and the LH_k sub-band gave the vertical details. The LL_k sub-band was the coarser resolution residual consisting of lower frequency components and was further split into higher levels of decomposition. According to the ISHT threshold function, the DWT was driven to acquire de-noised MODIS-NPP images. For visual comparison, we selected two MODIS-NPP images from 1 January 2014 and 27 December 2015 (Figure 4) with different amounts of noise. The 27 December 2015 image was heavily contaminated by clouds, while the 1 January 2014 image only contained a little noise.



Figure 4. Comparison of original images and de-noised images including (**a**) the original image acquired from MOD17 on 1 January 2014, (**b**) the results of removing noise from the 1 January 2014 original image, (**c**) the original image acquired from MOD17 on 27 December 2015, (**d**) the results of removing noise from the 27 December 2015 original image, and (**e**)–(**h**) spatial subsets within the squares, delineated in red.

The white speckles in the original image from MOD17 on 1 January 2014 were notably suppressed in the de-noised results, suggesting that the DWT could reduce noise effectively. While comparing the original image with heavy cloud noise from 27 December 2015, the results of removing noise from the original image had little change, which indicated the de-noised (post DWT) results could not remove heavy clouds. To quantify the de-noising performance, RMSD and SNR were calculated between original the MODIS-NPP images and the de-noised ones. For 1 January 2014 and 27 December 2015, the respective measures were as follows: RMSD 0.36 Mg ha⁻¹ a⁻¹ and SNR 17.42, and RMSD 0 Mg ha⁻¹ a⁻¹ and SNR ∞ . Thus, the ISHT thresholding method achieved good speckle suppression along with improving the 1 January 2014 NPP image quality, while the values of SNR and RMSD were ineffective for MODIS-NPP images from 27 December 2015 with heavy cloud contamination.

4.2. Extended Kalman Filter Prediction

De-noised MODIS-NPP data were provided as input to the EKF model, and the resulting NPP data for the representative sample plots in the study area were plotted along with the original MODIS-NPP data, de-noised MODIS-NPP data, and measured NPP data (Figure 5).

Figure 5a shows that the predicted data (in blue) were close to the measured data (in green), while the original data (in black) were generally much lower than the others. Figure 5b shows that REs between predicted MODIS-NPP data and measured NPP data were smaller than REs between the original MODIS-NPP data and measured NPP data and between the de-noised MODIS-NPP data and measured NPP data. These lower REs between the predicted MODIS-NPP data and measured NPP data suggest that the proposed method has a greater potential to estimate NPP data with high accuracy compared to the results of DWT alone.



Figure 5. (a) Time series of the original MODIS-NPP data from MOD17, de-noised MODIS-NPP data, predicted MODIS-NPP data, and the measured NPP data from 2013 to 2015; (b) RE between the original MODIS-NPP data and the measured NPP data, RE between the de-noised MODIS-NPP data and the measured NPP data, and RE between the predicted MODIS-NPP data and measured NPP data based on the six representative sample plots.

Comparing the predicted MODIS-NPP data on the 361st day in 2015 with the de-noised MODIS-NPP data obtained from the DWT algorithm with the original MODIS-NPP data for the same date, the predicted NPP estimates showed better quality than the de-noised MODIS-NPP (Figure 6), indicating that the data quality of the de-noised data alone was too low in the presence of heavy clouds and heavy aerosols.



Figure 6. Comparison of spatial distributions of NPP on the 361st day from (**a**) the original MODIS-NPP (**b**) de-noised using the DWT, and (**c**) predicted using the method combining the DWT with EKF.

The results of comparing the different NPP products with the NPP in the 20 sample plots are shown in Table 2. The average NPP of the sample plots was 4.19 Mg ha⁻¹ a⁻¹. The average NPP was $3.59 \text{ Mg ha}^{-1} a^{-1}$ for the predicted NPP, $3.45 \text{ Mg ha}^{-1} a^{-1}$ for the de-noised NPP, and $3.19 \text{ Mg ha}^{-1} a^{-1}$ for the original MODIS-NPP data. The RE relative to plot NPP estimates varied from 2.84% to 28.02% for the predicted MODIS-NPP data and from 6.06% to 37.23% for the original MODIS-NPP data. The rRMSE was 15.67% for the predicted MODIS-NPP data, 19.5% for the de-noised MODIS-NPP data, and 24.98% for the original MODIS-NPP data, indicating that the method combining the DWT and the EKF notably increases the estimation accuracy of NPP in the study area. The increase in accuracy could be attributed to de-noised data from the DWT and effective prediction of the EKF. The results indicate that the predicted MODIS-NPP using the proposed method is more precise than the original MODIS-NPP and the de-noised MODIS-NPP when monitoring above- and below-ground organic accumulated vegetation matter.

Plot#	Species	Measured NPP	Original NPP	RE (%)	De-Noised NPP	^l RE (%)	Predicted NPP	RE (%)
1	Pinus massoniana	3.72	2.78	25.27	2.89	22.31	3.49	6.18
2	Bamboo	3.65	3.34	8.60	3.64	0.39	3.39	7.23
3	Eucalyptus	3.92	3.10	20.92	3.24	17.35	3.33	15.05
4	Acacia confusa	3.90	2.80	28.21	3.30	15.39	3.15	19.23
5	Ficus	3.88	2.90	25.31	3.31	14.75	3.27	15.78
6	Pinus massoniana	4.17	3.20	23.26	3.36	19.42	3.39	18.71
7	Mucuna Birdwoodiana	4.62	2.90	32.90	3.2	26.41	4.26	7.79
8	Mulberry	4.58	3.40	25.69	3.40	25.69	3.66	20.01
9	Pinus massoniana	4.64	3.37	27.37	3.45	25.65	3.34	28.02
10	Schima	4.07	3.12	28.10	3.57	12.39	3.18	21.96
11	Cheery blossom	3.70	2.92	21.13	3.68	0.60	3.40	8.16
12	Ficus	4.01	3.30	17.80	3.76	6.35	3.37	16.06
13	Camphor	4.70	3.25	30.85	3.18	32.34	3.76	20.00
14	Pinus massoniana	3.96	3.72	6.06	3.24	18.18	3.57	9.85
15	Camphor	4.04	3.30	18.29	3.30	18.29	3.35	17.05
16	Lychee	4.05	3.05	24.71	3.34	17.55	3.87	4.46
17	Camphor	4.50	3.17	29.56	3.50	22.22	3.85	14.44
18	Delonix regia	4.23	3.20	24.35	3.72	12.06	4.11	2.84
19	Tea tree	4.44	3.30	25.60	3.74	15.68	3.95	10.95
20	Alsophila spinulosa	4.96	3.80	23.38	4.02	18.94	4.19	15.52
	Mean	4.19	3.19	23.58	3.45	17.10	3.59	13.96
	SD	0.38	0.28		0.26		0.34	
	rRMSE (%)		24.98		19.50		15.67	

Table 2. Comparison of NPP data between the predicted MODIS-NPP, de-noised MODIS-NPP, original MODIS-NPP, and measured NPP based on field observations of sample plots (NPP units: Mg $ha^{-1}a^{-1}$).

5. Discussion

5.1. Comparison with Other Similar Studies

Previous research approaches for removing noise from images to construct high-quality MODIS NDVI and LAI time-series datasets [10,13,14,17,23,38] while providing valuable information have some limitations.

Firstly, these methods do not perform well under conditions of heavy cloud contamination, heavy aerosols, and atmospheric variability, among other factors. Figure 6 shows that the vegetation in the west of study area cannot be found due to heavy cloudiness, while the NPP estimates using the novel method presented in this study can be displayed clearly. To quantify the de-noising performance, we computed the REs between the original and predicted data, and the REs between the de-noised and measured data (Figure 5b). The results showed that big REs occurred during the rainy season (from April to September), which further indicates that de-noised results cannot improve the NPP quality under conditions of heavy cloud contamination, heavy aerosols, and atmospheric variability.

Secondly, these methods aim at constructing high-quality MODIS time-series datasets for existing MODIS data and do not predict the data for the next time period. The method proposed in this paper leverages an effective wavelet de-noising function and a robust EKF predicting technique, where the wavelet transform is used to remove noise in MODIS-NPP images, and the EKF is applied to deal with de-noised images modeled as a triply modulated cosine function to predict MODIS-NPP data for the next time period. Figure 7 shows that the high-quality MODIS-NPP data (the last blue point) was predicted using the novel method on the 361th day in 2015, which implies that the novel method can be used to predict the next eight-day NPP data in the future.



Figure 7. Comparison of original MODIS-NPP data from MOD17 and predicted data for Schima in 2015.

Thus, this novel method can also produce high-quality MODIS data in cases of heavy cloud contamination, heavy aerosol, and atmospheric variability, and can also predict MODIS-NPP values in a time series. Compared to original and de-noised NPP, the predicted values in the study area increased the estimation accuracy of forest NPP, reducing the rRMSE by 9.31% and 3.83%, respectively, which supports the novel method.

5.2. Prospects for Future Studies

In our research, a large amount of NPP data was needed to meet the high accuracy prediction requirements of the EKF algorithm. Some original MODIS-NPP data were invalid due to rain, heavy clouds, and heavy aerosol. Thus, we used the time-distance weighting approach to generate baseline NPP data before predicting the NPP data. However, this approach is not suitable for generating MODIS-NPP data during periods of rapid vegetation change, which may limit the generation of high-precision NPP products. Thus, the approach presented in this study should be used with caution. In addition, future work should include additional test sites to confirm the robustness of our proposed method.

6. Conclusions

In this study, we proposed a novel method combining the DWT with the EKF to accurately predict MODIS-NPP data when the original product is heavily contaminated by factors such as clouds or aerosols. For this purpose, the DWT algorithm of the wavelet transform successfully suppressed white speckles in the original MODIS-NPP data. The mean amplitude phase derived from EKF appeared to predict highly accurate NPP estimates based on the de-noised results. The results showed that compared to the original MODIS-NPP data, the method significantly decreased the mean RE and rRMSE values of the predicted MODIS-NPP data of the study area by about 9.62% and 9.31%, respectively, which indicates that this method has great potential for generating new high-quality MODIS-NPP data under conditions of invalid original MODIS-NPP values.

Even though the proposed method improved the accuracy effectively, some limitations still exist. One deficiency is that this method is not suitable for periods of rapid vegetation change, affecting the prediction accuracy of this period, and therefore the approach should be applied with caution. **Author Contributions:** Z.L. and Y.W. conceived and designed the experiments; Z.L., T.W., Y.Q., H.L. and X.W. performed the experiments; Z.L. analyzed the data; Z.L. wrote the paper.

Funding: This work was supported by the National Natural Science Foundation of China (41671333).

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

References

- 1. Ni, J. Net primary productivity in forests of China: Scaling-up of national inventory data and comparison with model predictions. *For. Ecol. Manag.* **2003**, *176*, 485–495. [CrossRef]
- 2. Golubyatnikov, L.L.; Denisenko, E.A. Modeling the values of net primary production for the zonal vegetation of European Russia. *Biol. Bull. Russ. Acad. Sci.* **2001**, *28*, 293–300. [CrossRef]
- Turner, D.P.; Ritts, W.D.; Cohen, W.B.; Gower, S.T.; Running, S.W.; Zhao, M.S.; Costa, M.H.; Kirschbaum, A.A.; Ham, J.M.; Saleska, S.R.; et al. Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sens. Environ.* 2006, 102, 282–292. [CrossRef]
- 4. Nemani, R.R.; Keeling, C.D.; Hashimoto, H.; Jolly, W.M.; Piper, S.C.; Tucker, C.J.; Myneni, R.B.; Running, S.W. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* 2003, 300, 1560–1563. [CrossRef] [PubMed]
- 5. Running, S.W.; Thornton, P.E.; Nemani, R.; Glassy, J.M. Global terrestrial gross and net primary productivity from the Earth Observing System. *Methods Ecosyst. Sci.* **2000**, *3*, 44–45.
- 6. Zhao, M.; Running, S.W. User's Guide Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Observing System MODIS Land Algorithm. *Version* **2015**, *3*, 1–28.
- Zhao, M.; Running, S.W.; Nemani, R.R. Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *J. Geophys. Res. Biogeosci.* 2006, 111, 1–13. [CrossRef]
- 8. Lu, X.; Liu, R.; Liu, J.; Liang, S. Removal of noise by wavelet method to generate high-quality temporal data of terrestrial MODIS products. Photogramm. *Eng. Remote Sens.* **2007**, *73*, 1129–1139.
- 9. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [CrossRef]
- 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]
- 11. Paul, S.; Willmes, S.; Gutjahr, O.; Preußer, A.; Heinemann, G. Spatial Feature Reconstruction of Cloud-Covered Areas in Daily MODIS Composites. *Remote Sens.* **2015**, *7*, 5042–5056. [CrossRef]
- 12. Gao, B.; Gong, H.; Wang, T.; Jia, L. Reconstruction of MODIS Spectral Reflectance under Cloudy-Sky Condition. *Remote Sens.* **2016**, *8*, 727. [CrossRef]
- 13. 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]
- 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]
- 15. Kang, J.; Tan, J.; Jin, R.; Li, X.; Zhang, Y. Reconstruction of MODIS Land Surface Temperature Products Based on Multi-Temporal Information. *Remote Sens.* **2018**, *10*, 1112. [CrossRef]
- 16. Zhou, J.; Jia, L.; Menenti, M.; Gorte, B. On the performance of remote sensing time series reconstruction methods A spatial comparison. *Remote Sens. Environ.* **2016**, *187*, 367–384. [CrossRef]
- 17. Xiao, Z.; Liang, S.; Wang, J.; Jiang, B.; Li, X. Real-time retrieval of Leaf Area Index from MODIS time series data. *Remote Sens. Environ.* **2011**, *115*, 97–106. [CrossRef]
- 18. Zhang, M.; Zhou, Q.; Chen, Z.; Jia, L.; Yong, Z.; Cai, C. Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 476–485.
- 19. Xu, J.; Hu, H.; Shu, H.; Hu, Z. Using compute unified device architecture-enabled graphic processing unit to accelerate fast Fourier transform-based regression Kriging interpolation on a MODIS land surface temperature image. *J. Appl. Remote Sens.* **2016**, *10*, 026036.

- Galford, G.L.; Mustard, J.F.; Melillo, J.; Gendrin, A.; Cerri, C.C.; Cerri, C.E.P. Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sens. Environ.* 2008, 112, 576–587. [CrossRef]
- 21. Westra, T.; De Wulf, R.R. Monitoring Sahelian floodplains using Fourier analysis of MODIS time-series data and artificial neural networks. *Int. J. Remote Sens.* **2007**, *28*, 1595–1610. [CrossRef]
- 22. 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]
- Gu, J.; Li, X.; Huang, C. Spatio-Temporal Reconstruction of MODIS NDVI Data Sets Based on Data Assimilation Methods. In Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Shanghai, China, 25–27 June 2008; pp. 242–247.
- 24. Gu, Y.; Bélair, S.; Mahfouf, J.F.; Deblonde, G. Optimal interpolation analysis of leaf area index using MODIS data. *Remote Sens. Environ.* **2006**, *104*, 283–296. [CrossRef]
- 25. Zeng, W.S. Methodology on Modeling of Single-Tree Biomass Equations for National Biomass Estimation in China. Ph.D. Thesis, Chinese Academy of Forestry, Beijing, China, 2011.
- 26. Xie, T.T.; Li, G.; Zhou, G.Y.; Wu, Z.M.; Zhao, H.B.; Qiu, Z.J.; Liang, R.Y. Aboveground biomass of natural Castanopsis carlesii-Schima superba community in Xiaokeng of Nanling Mountains, South China. *Chin. J. Appl. Ecol.* **2013**, *24*, 2399–2407.
- 27. Gen, L.I.; Zhou, G.; Wang, X.; Wu, Z.; Qiu, Z.; Zhao, H.; Liang, R. Aboveground biomass of natural Castanopsis fissa community at the Xiaokeng of NanLing Mountain, Southern China. *Acta Ecol. Sin.* **2011**, *31*, 3650–3658.
- 28. Chen, Z.H.; Zhang, H.D.; Wang, B.S.; Zhang, Z.Q. Studies on biomass and its allocation of the evergreen broadleaved forest in Heishiding, Guangdong. *Acta Phytoecol. Geobot. Sin.* **1993**, *17*, 289–298.
- 29. Fang, J.Y.; Liu, G.H.; Xu, S.L. Biomass and net production of forest vegetation in China. *Acta Ecol. Sin.* **1996**, 16, 497–508.
- 30. Yang, K.; Guan, D. Biomass and its distribution of forest in the Pearl River Delta. *Ecol. Environ.* **2006**, 15, 84–88.
- 31. Kang, J.; Liu, W.Q.; Yu, F.Q.; Zhai, Q.J.; Liao, W.B. Vegetation Types and Plant Community Characters in Bijiashan Park in Shenzhen, Guangdong Province. *Acta Sci. Nat. Univ. Sunyatseni* **2005**, *44*, 10–31.
- 32. Chang, S.G.; Yu, B.; Vetterli, M. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Process.* 2000, *9*, 1532–1546. [CrossRef]
- 33. Liu, H.B.; Han, S.Q.; Liu, J. Wavelet de-noising based on novel thresholding function and best decomposition scale. *Comput. Eng. Appl.* **2007**, *43*, 72–75.
- 34. Liu, W.D.; Liu, S.H.; Hu, X.F.; Wang, L. Analysis of Modified Methods of Wavelet Threshold De-Noising Functions. *High Volt. Eng.* **2007**, *33*, 59–63.
- 35. Zhang, X. The SURE-LET approach using hybrid thresholding function for image denoising. *Comput. Electr. Eng.* **2018**, *70*, 334–348. [CrossRef]
- Donoho, D.L.; Johnstone, I.M. Adapting to unknown smoothness via wavelet shrinkage. J. Am. Stat. Assoc. 1995, 90, 1200–1224. [CrossRef]
- 37. Luisier, F.; Blu, T.; Unser, M. A New SURE Approach to Image Denoising: Interscale Orthonormal Wavelet Thresholding. *IEEE Trans. Image Process.* 2007, *16*, 593–606. [CrossRef] [PubMed]
- Kleynhans, W.; Olivier, J.C.; Wessels, K.J.; Salmon, B.P.; Bergh, F.V.D.; Steenkamp, K. Detecting land cover change using an extended Kalman filter on MODIS NDVI time-series data. *IEEE Trans. Geosci. Remote Sens. Lett.* 2011, *8*, 507–511. [CrossRef]
- 39. Suwannachatkul, S.; Kasetkasem, T.; Chumkesornkulkit, K.; Rakwatin, P.; Chanwimaluang, T.; Kumazawa, I. Rice cultivation and harvest date identification based on a hidden Markov model. In Proceedings of the 11th IEEE International Conference on Electrical Engineering/electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Nakhon Ratchasima, Thailand, 14–17 May 2014; pp. 1–6.



© 2019 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 (http://creativecommons.org/licenses/by/4.0/).