

Article

A Robust Method for Generating High-Spatiotemporal-Resolution Surface Reflectance by Fusing MODIS and Landsat Data

Junming Yang ¹, Yunjun Yao ^{1,*}, Yongxia Wei ², Yuhu Zhang ³, Kun Jia ¹, Xiaotong Zhang ¹, Ke Shang ¹, Xiangyi Bei ¹ and Xiaozheng Guo ¹

¹ State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; julming@mail.bnu.edu.cn (J.Y.); jiakun@bnu.edu.cn (K.J.); xtngzhang@bnu.edu.cn (X.Z.); shangke@mail.bnu.edu.cn (K.S.); xiangyibei@mail.bnu.edu.cn (X.B.); 201921051074@mail.bnu.edu.cn (X.G.)

² Key Laboratory of High Efficiency Utilization of Agricultural Water Resources, Ministry of Agriculture, School of Water Conservancy and Architecture, Northeast Agricultural University, Harbin 150030, China; weiyongxia@neau.edu.cn

³ College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China; yuhu.zhang@cnu.edu.cn

* Correspondence: yaoyunjun@bnu.edu.cn; Tel.: +86-10-5880-3002

Received: 20 June 2020; Accepted: 15 July 2020; Published: 18 July 2020

Abstract: The methods for accurately fusing medium- and high-spatial-resolution satellite reflectance are vital for monitoring vegetation biomass, agricultural irrigation, ecological processes and climate change. However, the currently existing fusion methods cannot accurately capture the temporal variation in reflectance for heterogeneous landscapes. In this study, we proposed a new method, the spatial and temporal reflectance fusion method based on the unmixing theory and a fuzzy C-clustering model (FCMSTRFM), to generate Landsat-like time-series surface reflectance. Unlike other data fusion models, the FCMSTRFM improved the similarity of pixels grouped together by combining land cover maps and time-series data cluster algorithms to define endmembers. The proposed method was tested over a 2000 km² study area in Heilongjiang Province, China, in 2017 and 2018 using ten images. The results show that the accuracy of the FCMSTRFM is better than that of the popular enhanced spatial and temporal adaptive reflectance fusion model (ESTARFM) (correlation coefficient (*R*): 0.8413 vs. 0.7589; root mean square error (RMSE): 0.0267 vs. 0.0401) and the spatial-temporal data fusion approach (STDFA) (*R*: 0.8413 vs. 0.7666; RMSE: 0.0267 vs. 0.0307). Importantly, the FCMSTRFM was able to maintain the details of temporal variations in complicated landscapes. The proposed method provides an alternative method to monitor the dynamics of land surface variables over complicated heterogeneous regions.

Keywords: Landsat; MODIS; FCMSTRFM; spatiotemporal data fusion; unmixing theory

1. Introduction

The surface reflectance, which characterizes the ability of the land surface to reflect solar and sky radiation pertaining to vegetation cover, soil moisture and surface roughness [1–3], is a valuable indicator for recognizing ground objects [4,5] and retrieving biophysical variables, e.g., vegetation indices (VIs) [6,7], leaf area index (LAI) [8,9] and biomass [10,11]. The requirement for reflectance data with both high spatial and temporal resolution is increasingly important to simulate the surface energy budget [12–14] and monitor ecosystem and hydrologic dynamics [15–18] at regional and

global scales. In satellite design, however, a trade-off must be made between temporal and spatial resolutions [19,20]. As a result, it is difficult to acquire remotely sensed data with both frequent coverage and high spatial resolution [21–23].

One technique for increasing the spatial resolution of frequent coverage satellite observations is the blending of images from sensors with complementary temporal and spatial characteristics with the aim of generating synthetic images with both high temporal and spatial resolutions [24–27]. Recently, several spatial and temporal data fusion methods (STFMs) have been developed to generate images with both high spatial and high temporal resolution. These methods can be classified into two categories: methods based on the spatial and temporal adaptive reflectance fusion model (STARFM) (hereafter referred to as STARFM-based models) and methods based on unmixing theory (hereafter referred to as UMX-based models). STARFM-based models are based on the assumption that the ratio of coarse pixel reflectance to neighboring similar pixels does not change over time [22,25,28]. These methods are particularly useful for preserving spatial detail information [20,28]. Previous scientists have substantially improved the STARFM [29–32] but have not addressed the following weaknesses: (i) the assumption of STARFM-based models is violated in complex heterogeneous regions due to the temporal variability of surface reflectance [33]; (ii) the models cannot predict short time change events if the changes are not recorded in at least one of the base fine-resolution images [25,28,29]; and (iii) land cover change may lead to low accuracy in STARFM-based methods [25,28,33,34].

For UMX-based methods, the abundances and endmembers are obtained by grouping high-resolution pixels, and the average reflectance of endmembers is unmixed from coarse-resolution images based on the assumption that the reflectance of each coarse spatial-resolution pixel is a linear combination of the responses of each endmember contributing to the mixture [19,33,35,36]. Generally, the accurate prediction of surface reflectance is mainly limited by grouping similar pixels and calculating the average reflectance [37,38]. The pixels grouped together should have similar spatial and temporal variability [39,40].

The main advantage of UMX-based methods is that, unlike STARFM-based methods, they define endmembers by clustering similar pixels, and the average reflectance of endmembers is unmixed from the coarse-spatial-resolution images. This allows for three additional possibilities. First, multiphase images can be used to group similar pixels. Second, auxiliary datasets such as land cover maps may supplement the grouping of similar pixels [41]. Third, frequent coverage coarse-resolution images can be used to capture phenological variations (i.e., NDVI profiles) [20,34]. However, most UMX-based methods employ only one or two (generally at the beginning and end of the observation period) high-resolution images to cluster similar pixels [33,35,38], which may fail to capture short temporal interval events that are not recorded in base high-resolution images.

To overcome the limitations of UMX-based models and to obtain the accurate prediction of reflectance, even in areas with large temporal and spatial variance, in this study, we have proposed a new data fusion method based on unmixing theory, as well as the spatial and temporal reflectance fusion model based on the fuzzy C-clustering model (FCMSTRFM). The FCMSTRFM improves the similarity of pixels that are grouped together by combining land cover maps and multiphase images to record temporal and spatial variation information from all available high-resolution images, and a new strategy was developed to calculate the average reflectance of each endmember. To evaluate the performances of the FCMSTRFM, we validated the FCMSTRFM with ten Landsat 8 Operational Land Imager (OLI) and MODIS datasets in the HePing Irrigated Area (HPIA), and we also compared our method with two other methods (the spatial-temporal data fusion approach (STDFA) and the enhanced spatial and temporal adaptive reflectance fusion model (ESTARFM)). The paper is structured as follows. Section 2 describes the FCMSTRFM method. Section 3 provides experimental data and data processing. Section 4 lists the fused reflectance results derived. Section 5 presents a discussion, and Section 6 explains our conclusion.

2. Methods

2.1. FCMSTRFM Logic

Figure 1 shows the flowchart of the processing steps implemented in the FCMSTRFM. The algorithm requires multiple pairs of images (a pair of images means fine- and coarse-resolution images obtained on the same date), time-series coarse-resolution images for the expected prediction dates and land cover mapping. Images of all input models are required to be radiometrically calibrated, atmospherically corrected and subjected to image registration. There are four main steps of the FCMSTRFM: (1) class and subclass definition; (2) sensor-bias adjustment; (3) class and subclass average reflectance calculation; and (4) pixel reflectance calculation.

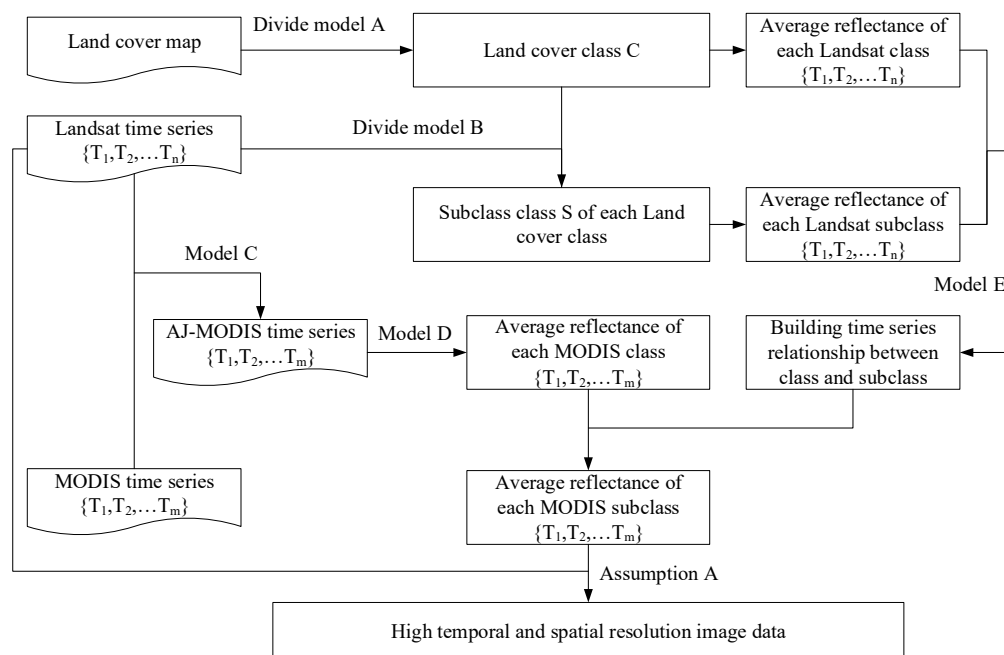


Figure 1. Flowchart of the method based on the spatial and temporal reflectance fusion model based on the fuzzy C-clustering model (FCMSTRFM). Cluster model A refers to time-series Landsat images clustered by land cover mapping; cluster model B refers to land cover class subclustered by the FCM; model C refers to sensor bias corrected by the method in Section 2.1.2; model D refers to unmixing of MODIS pixels; model E refers to the model in Section 2.1.3; and assumption A refers to the assumption that the temporal change in each fine-resolution pixel in the same subclass is constant. Parameter n is the number of Landsat images, and m is the number of MODIS images.

2.1.1. Class and Subclass Definition

To reduce the loss of temporal variability information, a method called pixel hierarchical clustering (PHC) is proposed for clustering pixels that have similar characteristics. PHC employs all available high-resolution images and land cover mapping to cluster similar pixels to ensure that the pixels clustered together have similar temporal variability. There are two steps in the PHC method: time-series Landsat images are clustered by land cover mapping to define the land cover class first, then each land cover class is further subclassified by the FCM to define the subclass.

Land cover data are mapped by dividing pixels in scene into several categories based on the pixels' reflectance, vegetation index and spectral characteristic [42–44], which is an effective classification of land surface characteristics. The pixels divided into one land cover type are similar in reflectance in some ways. For example, the vegetation spectral curve has specific peaks and valleys, and the soil reflectance increases with increasing wavelength. Therefore, we can use land cover mapping to coarsely cluster similarity pixels. However, it is possible that differences in the composition of land cover types (each land cover type contains multiple ground objects, e.g., the

different crops grown on farmland, impermeable layers containing roads and buildings) or environmental factors (e.g., soil type, altitude) can cause inconsistent temporal variability of pixels within the same land cover class. To eliminate the inconsistency, we further subclassify land cover classes.

Time-series data clustering algorithms (TSDCAs) cluster time-series data by minimizing the dissimilarity of time-series samples in the same cluster while maximizing the dissimilarity of different clusters [45–47]. Since the FCM, a TSDCA, has been shown to be effective in time-series data clusters [48–50], we employ the FCM to subclassify land cover classes. The FCM clustered the time-series dataset $X=\{X_{di}, di=2, \dots, n\}$ into c cluster centers $V=\{V_i, i=1, \dots, c\}$, $c \in \{2, 3, \dots, k, k < n-1\}$ by maximizing the following membership function:

$$U_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{c_{ij}}\right)^{2/(w-1)}} \quad (1)$$

where U_{ij} is the membership of X_{di} in cluster centers k , c is the number of cluster centers, w is the weighted index, d_{ij} is the Euclidean distance between data point i and the cluster center j , and c_{ij} is the total membership of each data point to the cluster center j . In this study, the FCM was used to subclassify each land cover type into 10 subclasses (the standard deviation of reflectance of subclass should less than 10% of the whole band), with a weighted index of 2.5 and termination error of 1×10^{-4} .

2.1.2. Sensor-bias Adjustment

Because of the differences in sensor systems, such as bidirectional reflectance distribution function (BRDF) effects, acquisition date, bandwidth and spectral response functions, there is a systematic reflectance bias between fine- and coarse-resolution images—that is, for a pure, homogeneous pixel covered by only one object, there is a bias in the reflectance among different sensors. To minimize the sensor bias, the MODIS pixel reflectance $R_M(x, y, t_i, B)$ (for ease of distinction, in this study, the subscripts M and L represent MODIS and Landsat images, respectively) must be adjusted to Landsat pixel reflectance $R_L(x, y, t_i, B)$ before blending them.

For a pure pixel, the difference in reflectance between different sensors only results from the sensor bias. Therefore, the relationship between MODIS and Landsat reflectance can be reasonably described by a linear model expressed as:

$$R_L(x, y, C, t_i, B) = a \times R_M(x, y, C, t_i, B) + b \quad (2)$$

where R_L and R_M denote the reflectance of Landsat and MODIS images, respectively; (x, y) is the location of both Landsat and MODIS pixels; t_i is the image acquired date; C is the land cover type; B is band of images and a and b are constants related to location and environment. To minimize error, in this study, we assume that different land cover types have different values of a and b .

Assuming the Landsat pixel number of land cover type C is n , then the relationship between all Landsat and MODIS pixel reflectance of land cover type C can be described as (3), and the relationship between Landsat and MODIS average reflectance of land cover type C can be described as (4).

$$\sum_1^k R_L(x, y, C, t_i, B) = a \times \sum_1^k R_M(x, y, C, t_i, B) + n \times b \quad (3)$$

$$\bar{R}_L(C, t_i, B) = a \times \bar{R}_M(C, t_i, B) + b \quad (4)$$

where \bar{R}_L and \bar{R}_M denote the average reflectance of Landsat and MODIS images, respectively. From (4), we know that there is also a linear relationship between Landsat and MODIS land cover class average reflectance. If there are multitemporal ($t=t_1, t_2, \dots, t_n$) fine- and coarse-resolution images, a linear regression model can be used to acquire a and b in (4). Equation (4) can be written as Equation (5) at date t_j .

$$\bar{R}_L(C, t_j, B) = a \times \bar{R}_M(C, t_j, B) + b \quad (5)$$

From Equations (4) and (5), we can obtain:

$$\bar{R}_L(C, t_i, B) - \bar{R}_L(C, t_j, B) = a(\bar{R}_M(C, t_i, B) - \bar{R}_M(C, t_j, B)) \quad (6)$$

Equation (6) shows that the change in fine-resolution land cover class average reflectance from t_i to t_j equals the scaled change in average reflectance t_i to t_j given by the coarse-resolution land cover class.

2.1.3. Class and Subclass Average Reflectance Calculation

Although it is mathematically feasible to use as many classes as the number of coarse-resolution images, too many endmembers may lead to ill-posed spectral unmixing [51]. Too few endmembers may lead to a decrease in the similarity of pixels within each endmember. To address this irreconcilable contradiction, we propose an efficient strategy for the unmixing problem that produces realistic estimated average reflectance of endmembers. The average reflectance of the land cover class was unmixed from the coarse-resolution images, and the average reflectance of the subclass was solved by building the time-series relationship between the subclass and land cover class.

In the FCMSTRFM, unmixing is performed by solving the linear mixing model. According to the linear mixing model, the reflectance $R_M(k, t_i, B)$ of coarse-resolution pixel k that consists of m discrete land cover class C weighted by fractional $A(k, C)$ can be described as Equation (7). If the number of MODIS pixels in the scenes is pn and the number of land cover classes is pm , the relationship between the reflectance of the coarse pixel and the average reflectance of the land cover class can be described as Equation (8).

$$R_M(k, t_i, B) = \sum_{C=1}^{pm} A(k, C) \times \bar{R}_M(C, t_i, B) + \varepsilon(C, t_i) \quad (7)$$

constraint condition:

$$\begin{bmatrix} R_M(1, t_i, B) \\ R_M(2, t_i, B) \\ \dots \\ R_M(pn, t_i, B) \end{bmatrix} = \begin{bmatrix} A(1,1) & A(1,2) & \dots & A(1, pm) \\ A(2,1) & A(2,2) & \dots & A(2, pm) \\ \dots & \dots & \dots & \dots \\ A(pn, .1) & A(pn, 2) & \dots & A(pn, pm) \end{bmatrix} \begin{bmatrix} \bar{R}_M(1, t_i, B) \\ \bar{R}_M(2, t_i, B) \\ \dots \\ \bar{R}_M(pm, t_i, B) \end{bmatrix} + \begin{bmatrix} \varepsilon(1, t_i) \\ \varepsilon(2, t_i) \\ \dots \\ \varepsilon(pm, t_i) \end{bmatrix} \quad (8)$$

where $A(k, C)$ is the abundance matrix of mixed pixel k ; and $\varepsilon(C, t_i)$ is the residuals. The time-series average reflectance of each land cover class can be calculated by solving Equation (8) using the ordinary least squares technique.

After obtaining the average reflectance of each land cover class, the average reflectance of the subclass is obtained by building the time-series relationship between the land cover class and subclass. If we assume that $f(S, t_i, B)$ is the ratio of the average reflectance of subclass S to the average reflectance of land cover class C (Equation (9)), $f(S, t_i, B)$ (Equation (9)) describes how the subclass

average reflectance changes in terms of land cover class average reflectance variables. The ratio $f(S, t_i, B)$ is greater than 1 when the variation rate of the subclass is greater than that of the land cover class; $f(S, t_i, B)$ is less than 1 when the variation rate of the subclass is less than that of the land cover class.

$$f(S, t_i, B) = \bar{R}_M(S, t_i, B) / \bar{R}_M(C, t_i, B) \quad (9)$$

Because of the spectral temporal variability, $f(S, t_i, B)$ changes over time. Fortunately, the assumption that $f(S, t_i, B)$ changes linearly during short temporal intervals has been shown to be mathematically reasonable [52]. If date t_p is between adjacent dates t_i and t_j , the relationship among $f(S, t_p, B)$, $f(S, t_i, B)$ and $f(S, t_j, B)$ can be described as Equation (10)—that is, Equation (10) can be used to calculate the ratio of average reflectance between subclass and class at any date among t_i and t_j . If there are multiphase images, Equation (10) can be used to build the time-series relationship between the subclass and class.

$$f(S, t_p, B) = f(S, t_i, B) + \frac{(f(S, t_i, B) - f(S, t_j, B)) * (t_p - t_i)}{t_i - t_j} \quad (10)$$

2.1.4. Pixel Reflectance Calculation

Because the pixels within each subclass have similar reflectance variability, the reflectance of each fine-resolution pixel at each time can be calculated by the surface reflectance calculation model (SRCM) proposed by Wu et al. [35]. This assumes that the temporal increment of each fine-resolution pixel in the same class (subclass in this paper) is constant and equal to the average reflectance increment of the class, described as follows:

$$R_L(k, t_p, B) - R_L(k, t_0, B) = \bar{R}_L(S, t_p, B) - \bar{R}_L(S, t_0, B) \quad (11)$$

where $R_L(k, t_0, B)$ and $R_L(k, t_p, B)$ are the reflectance of pixel k of fine resolution at dates t_0 and t_p , respectively, and $\bar{R}_L(k, t_0, B)$ and $\bar{R}_L(k, t_p, B)$ are the average reflectance of subclass S of fine resolution at times t_0 and t_p , respectively.

Combining Equations (6) and (11) yields Equation (12):

$$R_L(k, t_p, B) = a \times (\bar{R}_M(S, t_p, B) - \bar{R}_M(S, t_0, B)) + R_L(k, t_0, B) \quad (12)$$

Combining Equations (9) and (12) yields Equation (13):

$$R_L(k, t_p, B) = R_L(k, t_0, B) + a \times (f(S, t_p, B) \times \bar{R}_M(C, t_p, B) - f(S, t_0, B) \times \bar{R}_M(C, t_0, B)) \quad (13)$$

Since $f(S, t_p, B)$ can be calculated by Equation (9) and Equation (10), $\bar{R}_M(k, t_0, B)$ and $\bar{R}_M(k, t_p, B)$ can be calculated by Equation (8), and a can be calculated by Equation (6), we can obtain $R_L(k, t_p, B)$ with Equation (13)—that is, the time-series reflectance of each fine-resolution pixel can be calculated by Equation (13).

2.2. Comparison with Other Fusion Methods

2.2.1. STDFA

The STDFA was proposed by Wu et al. [35] to generate images with MODIS coverage frequency and Landsat spatial resolution using four steps: (1) clustering the high-resolution pixels to define endmembers based on two pairs of images; (2) calculating the abundances of each endmember within each coarse-resolution pixel; (3) unmixing the coarse-resolution pixel based on linear unmixing theory; and (4) calculating the reflectance of each high-resolution pixel.

The STDFA clusters the high-resolution pixels by dividing the increment of two pairs of images (generally at the beginning and end of the observation period) into several equal parts. After that, a sliding window the size of MODIS pixels is applied to the clustered result to record the endmember

abundance matrix. Then, the unmixing of coarse-resolution pixels is performed by the linear mixing model introduced in Section 2.1.3 (Equation (8)). Then, the reflectance of each high-resolution pixel is calculated by the theory of SRCM (Equation (11)), which assumes that the pixels belonging to the same class have the same temporal changes.

2.2.2. ESTARFM

The STARFM and the ESTARFM are based on the premise that the changes in reflectance between fine- and coarse-resolution images are consistent. After data processing as described in Section 3.2, the ESTARFM applies the following steps to generate high spatial and temporal resolution images. First, a sliding window is applied to the Landsat image to identify similar neighboring pixels. Considering that the reflectance may vary significantly over time, the ESTARFM uses two fine-resolution images to select similar neighbor pixels and then extracts the intersection of the two dates to obtain a more accurate similar-neighbor pixel. Second, the weight W is assigned to each similar neighbor based on (i) the spatial distance between the central pixel and neighboring pixels and (ii) the spectral similarity between a pair base date image. Third, the conversion coefficient V_i is calculated by linear regression analysis. The final step is the calculation of central pixel reflectance, which can be characterized by Equations (14–15). For a more theoretical detail of the ESTARFM algorithm, we refer to [28].

$$T_z = \frac{1/|\sum_{j=1}^w \sum_{l=1}^w C(x_j, y_l, t_z, B) - \sum_{i=1}^w \sum_{l=1}^w C(x_i, y_l, t_p, B)|}{\sum_{z=m,n} (1/|\sum_{j=1}^w \sum_{l=1}^w C(x_j, y_l, t_z, B) - \sum_{i=1}^w \sum_{l=1}^w C(x_i, y_l, t_p, B)|)} \quad (14)$$

where T_z ($z=m, n$) is a proportion that is used to assign the weight that time t_m and t_n contribute to predict time t_p .

$$I(x_{w/2}, y_{w/2}, t_p, B) = T_m \times I_m(x_{w/2}, y_{w/2}, t_p, B) + T_n \times I_n(x_{w/2}, y_{w/2}, t_p, B) \quad (15)$$

2.3. Evaluation Metrics

We selected four different statistical criteria to evaluate the performance of the FCMSTRFM, the ESTARFM and the STDFA: correlation coefficient (R , Equation (16)), root mean square error (RMSE, Equation (17)), the Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS, Equation (18)), and the mean absolute difference (MAD, Equation (19)). R indicates the linear correlativity between the observed and predicted reflectance. The RMSE is calculated as the square root of the average value of the differences between the predicted and observed reflectance. The lower the RMSE is, the more reliable the performance of the model. ERGAS demonstrates the quality of the predicted reflectance. MAD reflects the average differences between the predicted and observed reflectance. The metrics are calculated as follows:

$$R = \frac{\sum_{i=1}^{np} (k_{ai} - \bar{k}_a)(k_{pi} - \bar{k}_p)}{\sqrt{\sum_{i=1}^{np} (k_{ai} - \bar{k}_a)^2 \sum_{i=1}^{np} (k_{pi} - \bar{k}_p)^2}} \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{np} (k_{ai} - k_{pi})^2}{np}} \quad (17)$$

$$ERGAS = 100 \frac{H}{L} \sqrt{\frac{1}{N_b} \sum_{j=1}^{N_b} (RMSE_j / M_j)^2} \quad (18)$$

$$MAD = \frac{1}{np} \sum_{i=1}^{np} |k_{ai} - k_{pi}| \quad (19)$$

where k_{ai} and k_{pi} are the observed and predicted reflectance of pixel i , np is the pixel number of the image, $\overline{K_a}$ and $\overline{K_p}$ are the average reflectance of the observed and generated images, H and L are the pixel sizes of the fine- and coarse-resolution images, N_b is the band number, and M_j is the average reflectance of band j .

3. Experimental Data and Data Processing

3.1. Study Area

The FCMSTRFM was tested and validated in HPIA (Figure 2); the longitude of this region ranges from 127°18'40.28" to 127°45'12.22", and the latitude ranges from 46°51'7.83" to 47°4'8.33". The northeastern area of HPIA is a high mountain, while the southwestern area of HPIA is a plain. The main land cover type of the plain is farmland, where the crop growing period is short and phenology changes rapidly. The high mountains are generally covered by broadleaved deciduous forests. The farmland and high mountains are covered by green vegetation from April to November and covered by snow at other times. HPIA is a complex heterogeneous region with diverse land use types, including grassland, road and water. The heterogeneous region is suitable for testing the FCMSTRFM.

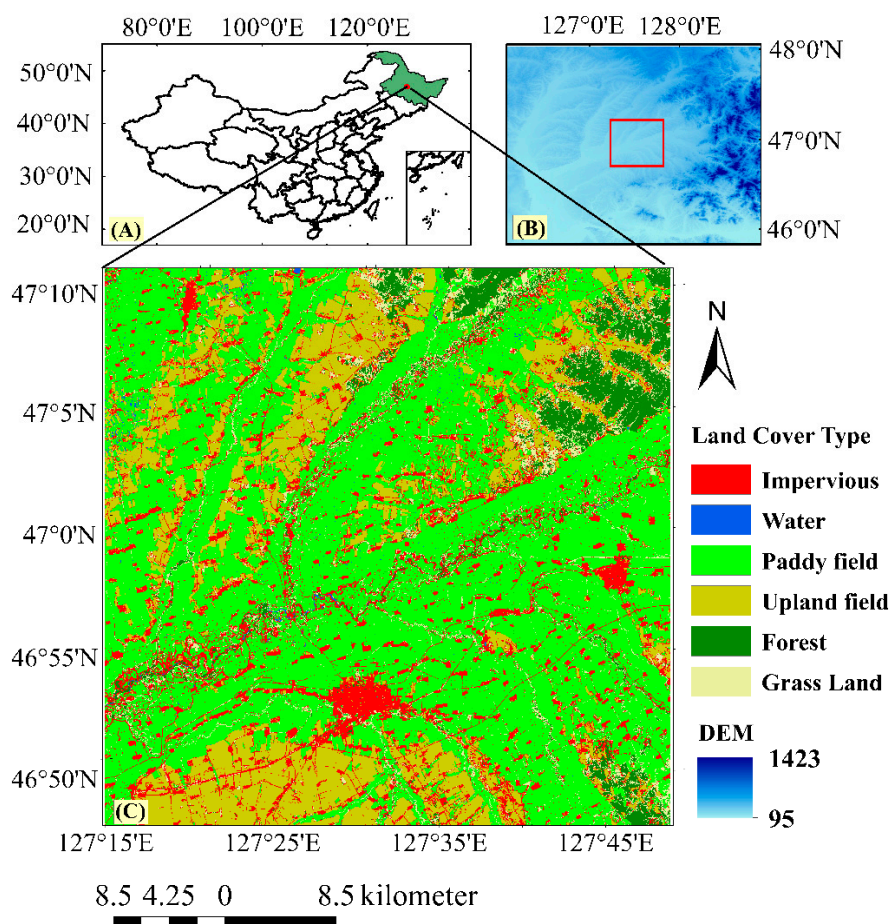


Figure 2. The location of study area (A), the digital elevation model (DEM) at 90 m spatial resolution (B) and land cover mapping (C) of study area.

3.2. Data Preprocessing

Ten Landsat 8 Operational Land Imager (OLI) and MODIS datasets were used in this study (Table 1). Table 2 shows the attributes of the Landsat 8 OLI bands and their matched MODIS bands. All images were acquired in clear sky conditions and provided by the United States Geological Survey (USGS) (<https://glovis.usgs.gov/>). The Landsat images were atmospherically corrected using the FLAASH Atmospheric Correction Model in ENVI 5.3 software (The Environment for Visualizing Images, provided by Harris Geospatial Solutions, Inc in Florida, USA). Landsat 8 OLI has been geometrically corrected, but to obtain higher accuracy, the Landsat images were georeferenced using a topographic map of 1:10,000 scale based on the nearest neighbor resampling method with a position error within 0.5 pixels. MODIS level L2G product MOD09GA and MODIS level L3 product MOD13Q1 were reprojected to the UTM-WGS84 52N projection using the MODIS Reprojection Tool (MRT), clipped to the extent of the Landsat images, and resampled to Landsat resolution using a nearest neighbor approach.

Table 1. Landsat and MODIS data used in this study.

Image	Path/row	Acquisition date/DOY
Landsat	117/027	11/04/2017,17/08/2017, 02/09/2017, 04/10/2017
		25/02/2018, 29/03/2018, 30/04/2018,01/06/2018, 07/10/2018, 20/12/2018
MOD09GA	h25/v04	11/04/2017,17/08/2017, 02/09/2017, 04/10/2017
		25/02/2018, 29/03/2018, 30/04/2018,01/06/2018, 07/10/2018, 20/12/2018
MOD13Q1	h25/v04	11/04/2017,17/08/2017, 02/09/2017, 04/10/2017

Table 2. Landsat8 OLI bands and its matched MODIS bands.

Landsat8 OLI			MODIS		
Band	Bandwidth (μm)	Spatial Resolution (m)	Band	Bandwidth (μm)	Spatial Resolution (m)
1	0.433–0.453	30	9	0.438–0.448	1000
2	0.450–0.515	30	3	0.459–0.479	500
3	0.525–0.600	30	4	0.545–0.565	500
4	0.630–0.680	30	1	0.620–0.670	250
5	0.845–0.885	30	2	0.841–0.876	250
6	1.560–1.660	30	6	1.628–1.652	500
7	2.100–2.300	30	7	2.105–2.155	500
8	0.500–0.680	15	-	-	-
9	1.360–1.390	30	26	1.360–1.390	1000
10	10.60–11.19	100	31	10.780–11.280	1000
11	11.50–12.51	100	32	11.770–12.270	1000

In addition to Landsat and MODIS image data, there is also a need for land cover mapping. The accuracy of land cover mapping has an important influence on the accuracy of the generated images. Although there are many Landsat-scale land cover mapping datasets available, we did not use the downloaded dataset of the land cover data and selected the Landsat 8 OLI image data to classify the area into water, grassland, forest, paddy field, upland field and impervious (Figure 2) based on a support vector machine (SVM). Verification by ground samples indicated that the overall accuracy of the land cover data was 87.33%, and the kappa coefficient was 86.07%.

4. Results

4.1. Evaluation of the FCMSTRFM

The accuracy of the FCMSTRFM was evaluated by comparing the generated images against the observed cloudless images obtained from the Landsat 8 OLI acquired on 01 June 2018, and 30 April 2018. Table 3 shows the accuracy of the generated image of three bands (red, green and NIR bands), indicating that the FCMSTRFM can generate images very similar to the observed images (the FCMSTRFM also can be used in cirrus and thermal bands). Most generated images have a high accuracy with R higher than 0.70 and RMSE less than 0.04. To assess whether the FCMSTRFM is suitable for regions where the land cover type changes, the three reflectance bands generated by the FCMSTRFM were compared with those observed on time series based on the availability of clear scene images (Figure 3). Although the land cover type has suffered complex processes (for example, green vegetation-covered land changed to bare land in December, farmland and grassland was covered by snow in February), it is clear that the images generated by the FCMSTRFM are very similar to those observed.

Table 3. Precision index of six bands of three phases by the spatial and temporal reflectance fusion model based on the fuzzy C-clustering model (FCMSTRFM).

Time	Band	R	RMSE	MAD	ERGAS
30/4/2018	Green	0.7857	0.018	0.0129	1.7107
	Red	0.7914	0.0271	0.0214	2.0703
	NIR	0.8432	0.0382	0.0282	1.7697
1/6/2018	Green	0.8331	0.0163	0.0112	1.6398
	Red	0.8483	0.0227	0.0159	2.2883
	NIR	0.9459	0.0385	0.0273	1.4883

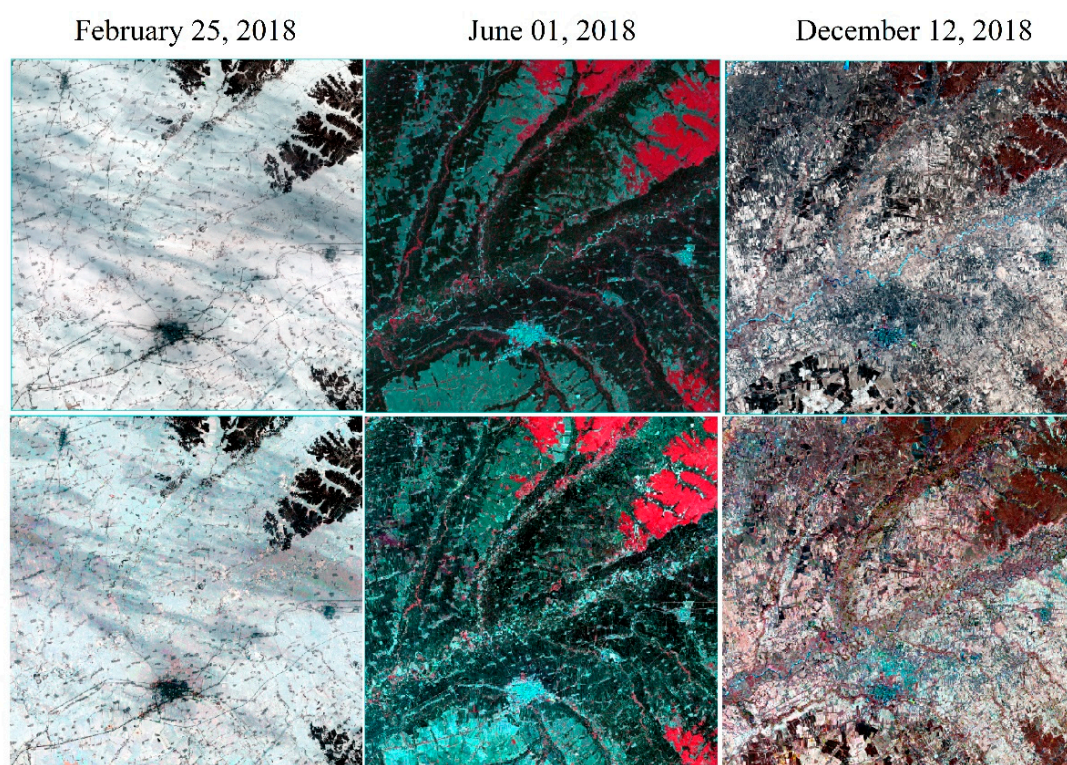
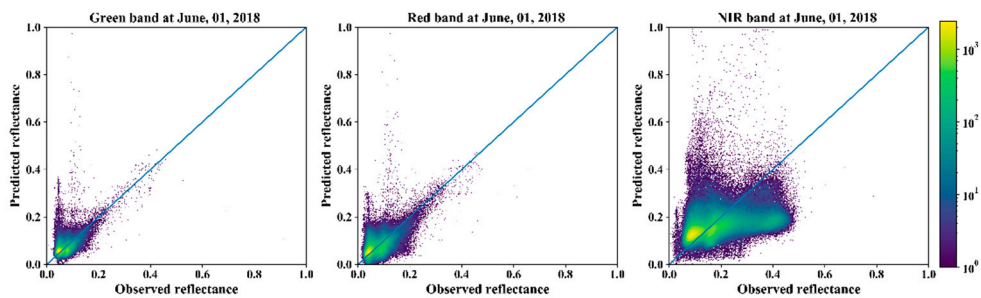


Figure 3. False color display (NIR-red-green combination) of observed and predicted time series reflectance (The order from left to right is covered by snow, green vegetation and bare land). The order from top to bottom is observed reflectance, predicted reflectance.

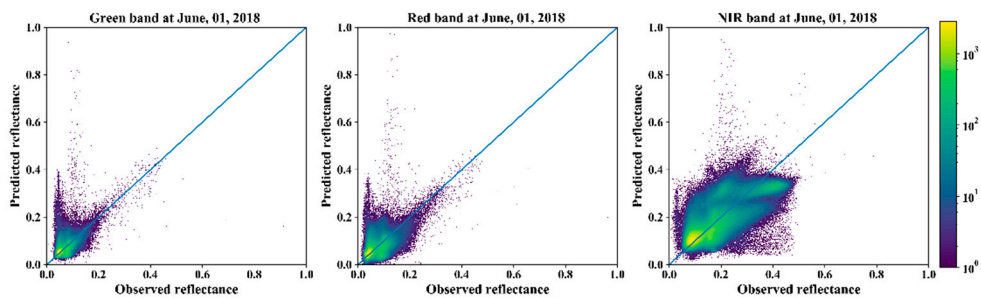
4.2. Comparison with Other Fusion Methods.

Figure 4 shows the relationship between the observed and predicted reflectance on June 01, 2018, for the green, red and NIR bands. For the green and red bands, most data points in the scatterplots are close to the 1:1 line, indicating that the three models can generate reflectance very similar to the actual reflectance in these two bands, but only the FCMSTRFM can generate reflectance similar to the actual reflectance in the NIR band. Figure 5 shows absolute value of different between observed and predicted reflectance on June 01, 2018, for the three models. It is clear that the FCMSTRFM can generate reflectance closest to the actual reflectance. Almost all images, especially those of the ESTARFM, in Figure 5 are visually similar to land cover mapping, indicating that the predicted reflectance is strongly influenced by the land use type. For three models, there is a wavelength dependency in the improvements (Figures 4 and 5), particularly in the NIR region. This is caused by the different of reflectance variation range, green and red band change from 0 to 0.2, but NIR from 0 to 0.5.

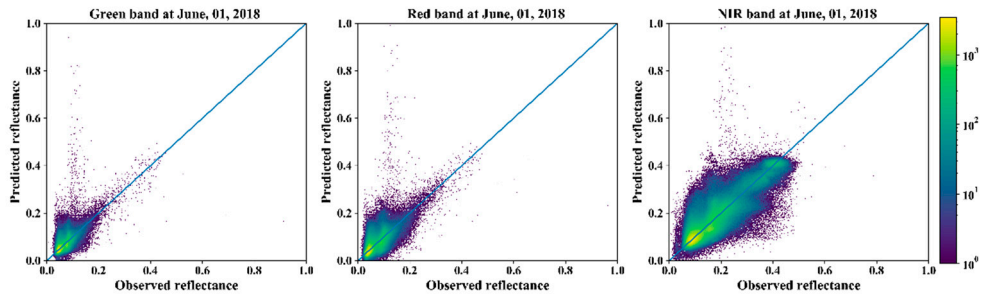
Figure 6 shows the density curve of the difference between the generated image and observed image for the three methods. The curve of the FCMSTRFM is much more concentrated, the bias was close to zero and the difference was constrained to ± 0.01 , except for a few outliers. Based on the statistical comparison (Table 4) between the three methods, the correlation coefficient increased by 0.02 and RMSE decreased by 0.01 on average. Almost all statistical results of the FCMSTRFM are greater than those of the ESTARFM and the STDFA, except for a few marked in black (Table 4), which means that the reflectance predicted by the FCMSTRFM was more accurate than that from the other two methods. Figure 7 shows the Landsat-like images generated by the FCMSTRFM, the STDFA and the ESTARFM. It is clear that the image generated by the ESTARFM (Figure 7d) is more similar to the observed image (Figure 7(c)), and the image generated by the STDFA (Figure 7e) contains some reflectance errors caused by cluster errors of similarity pixels. Somewhat “blurry” areas can be seen in the images generated by the FCMSTRFM and the STDFA. The image generated by the ESTARFM (Figure 7f) contains more spatial details, but some reflectance change information has been stretched or compressed.



(A) ESTARFM



(B) STDFA

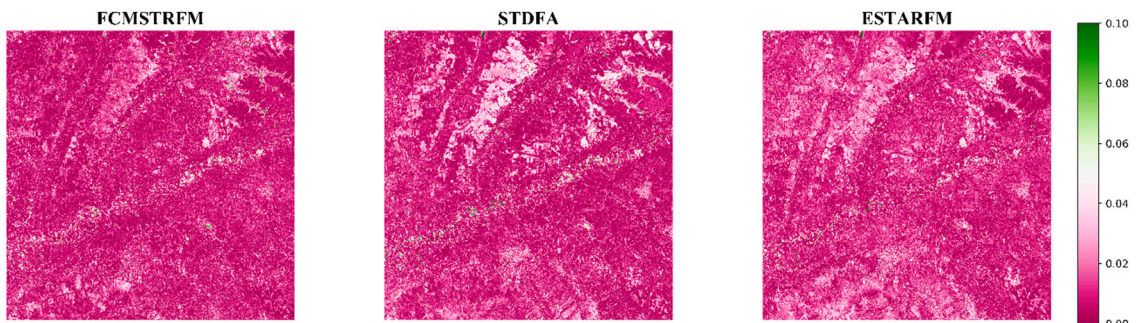


(C) FCMSTRFM

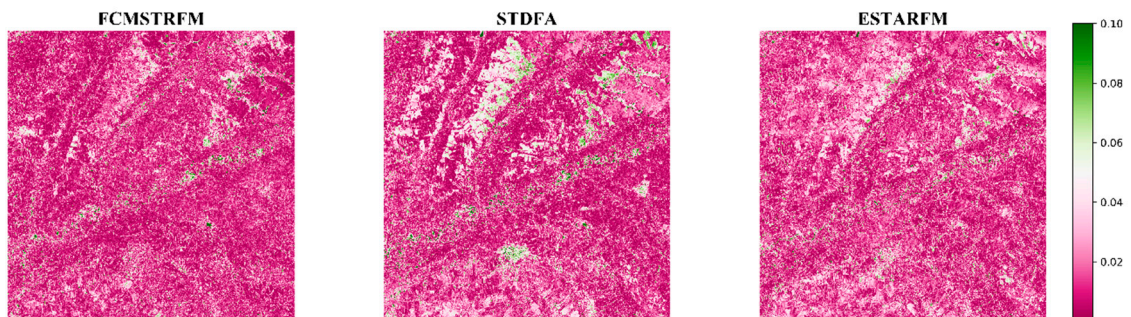
Figure 4. Scatter plots of observed actual reflectance and predicted reflectance by the enhanced spatial and temporal adaptive reflectance fusion model (ESTARFM) (A), the spatial-temporal data fusion approach (STDFA) (B) and the FCMSTRFM (C) for green, red and NIR band.

Table 4. Precision index of six bands of three phases by the FCMSTRFM.

Time	Band	STDFA				ESTARFM			
		R	RMSE	MAD	ERGAS	R	RMSE	MAD	ERGAS
30/4/2018	Green	0.7357	0.0197	0.0138	2.873	0.8288	0.0241	0.0191	2.2861
	Red	0.748	0.0261	0.0191	2.994	0.7966	0.0378	0.028	2.8837
	NIR	0.7993	0.0407	0.031	2.884	0.807	0.0617	0.0424	2.8592
1/6/2018	Green	0.7158	0.0194	0.0129	1.9585	0.6778	0.0184	0.0138	1.8528
	Red	0.7094	0.0275	0.0193	2.7683	0.7513	0.025	0.0189	2.5191
	NIR	0.8913	0.0512	0.0377	1.9769	0.6919	0.0738	0.0524	2.8495



(A)



(B)

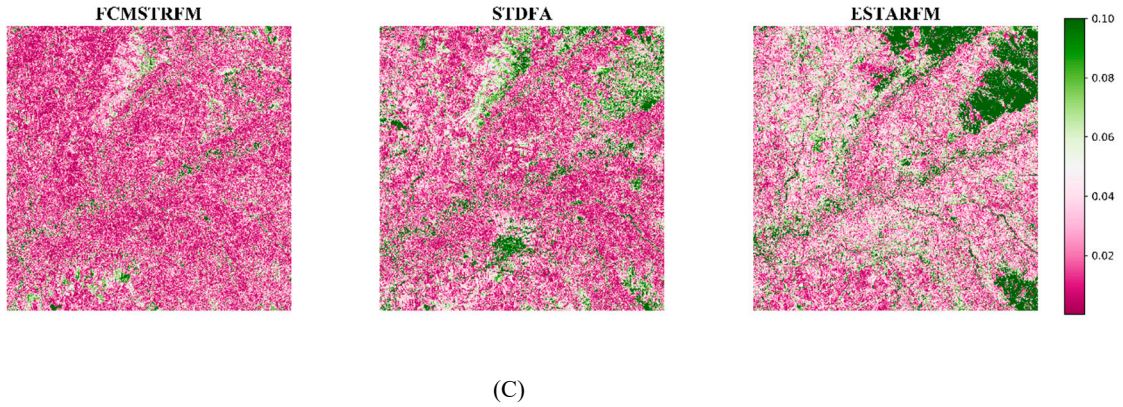


Figure 5. Absolute value of different between observed and predicted reflectance for three models in green band (A), red band (B) and NIR band (C).

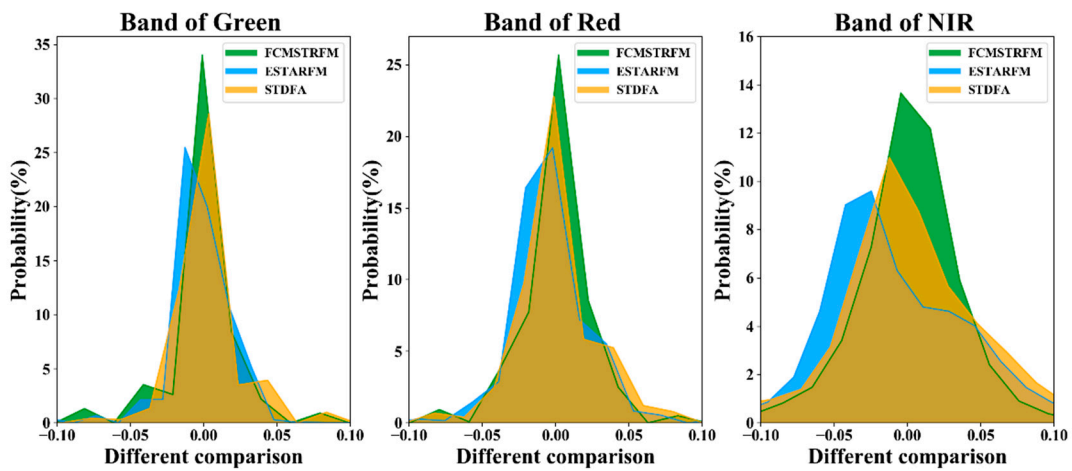


Figure 6. Density curve of different between observed and predicted reflectance in green, red and NIR band for three models.

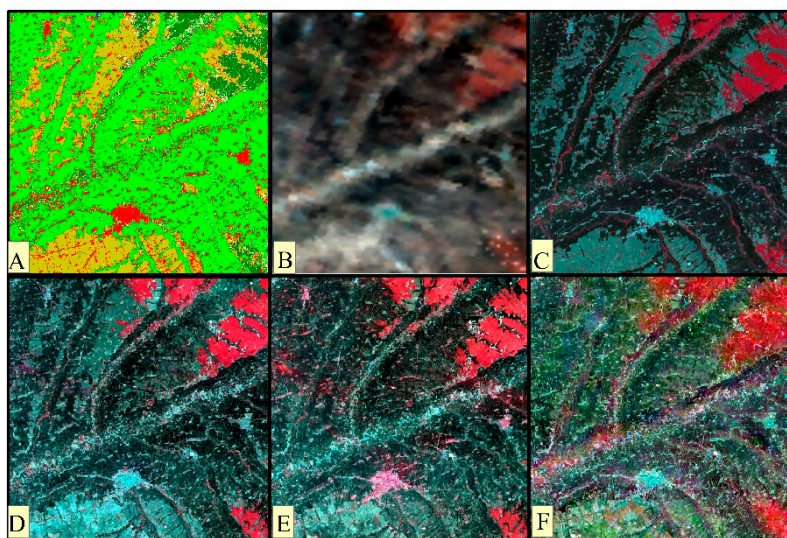


Figure 7. The land cover mapping (A) and false color display (NIR-red-green combination) of reflectance observed by MODIS (B) and Landsat 8 OLI (C) and predicted by the FCMSTRFM (D), the STDFA (E) and the ESTARFM (F).

5. Discussion

5.1. Uncertainties of the FCMSTRFM

5.1.1. Influence of Image Registration

Due to geometric correction and coordinate conversion errors, there is deviation in the image registration. To evaluate the influence of image registration, we moved the Landsat image in one direction (Figure 8) to obtain the RMSE change with the movement of the Landsat image. Figure 9 shows the relationship between the number of moving pixels and the RMSE of the generated reflectance. We found that when the number of moving Landsat pixels exceeded six, the accuracy of the generated images began to decrease substantially, indicating that the accuracy of registration strongly influences the performance of the FCMSTRFM. Therefore, the error of image registration should be small within six Landsat pixels.

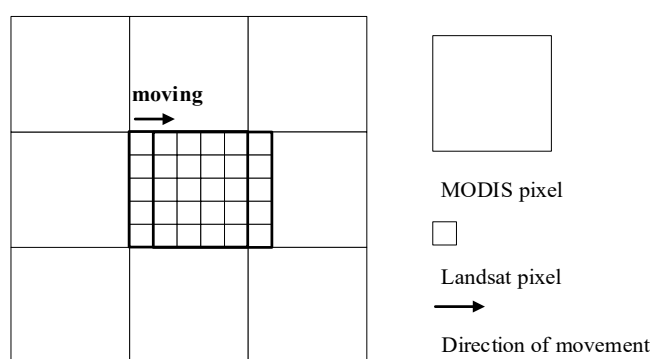


Figure 8. The schematic of Landsat image movement.

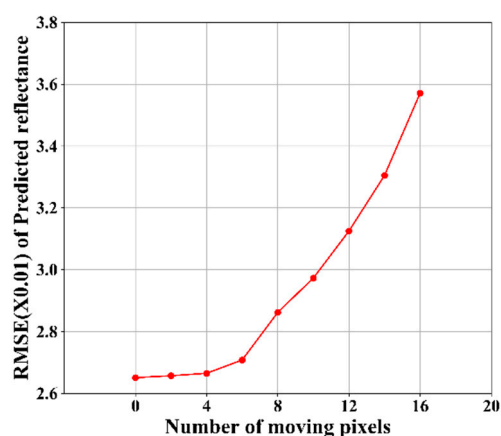


Figure 9. The relationship between number of moving pixels and RMSE.

5.1.2. Influence of the Accuracy of the Land Cover Map

In the FCMSTRFM, one notices that the pixels in the same land cover class have a similarity spectrum within a vegetation growth period, and in PHC, the fine-resolution images are categorized first by land cover mapping. The accuracy of land cover mapping directly determines the accuracy of the first clustering. If the accuracy of the land cover mapping is poor, the similarity of the land cover classes will be low. Therefore, it is clear that the accuracy of land cover mapping strongly influences the performance of the FCMSTRFM. We assess the influence of errors of land cover type classification by forcing randomly convert certain proportion pixels (5%, 10%, 15%, 20% and 25%) to be the wrong type 10 times, such as turning farmland into an impermeable layer or forestland, is used to obtain the RMSE of the generated image change with the conversion of the land cover type. Figure

10 shows the relationship between the converted proportion of land cover type and the average value of RMSEs for all the images generated by the FCMSTRFM. It shows that RMSEs had a positive correlation with the converted proportion of land cover type: the higher the converted proportion of land cover type was, the higher the RMSEs. That means that the lower the accuracy of the land cover map is, the lower the accuracy of the generated images.

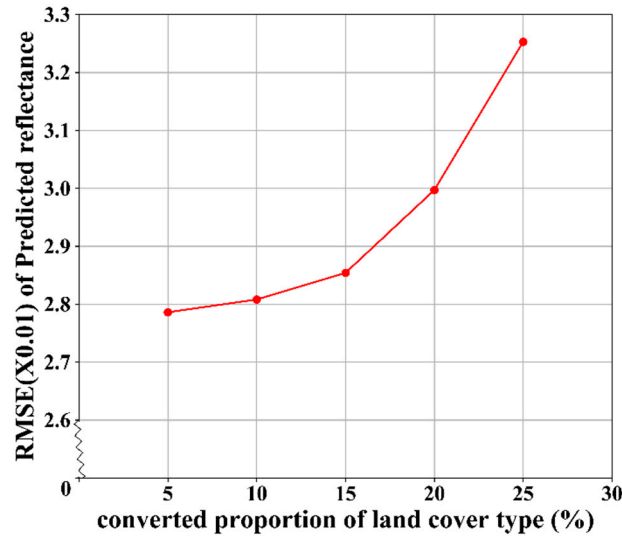


Figure 10. The relationship between the RMSE of the generated image and the converted proportion of land cover type.

5.1.3. Influence of Temporal and Spatial Heterogeneity

To investigate the sensitivity of the three models to temporal changes in reflectance, we chose the variance of difference between two fine-resolution images (σ in Equation (20)) as a measure of the reflectance change intensity, and RMSE as a measure of the accuracy of generated images. The three models have a positive correlation with the RMSE of the generated images (Figure 11). For almost all models, the stronger the temporal changes in reflectance are, the lower the accuracy of the generated images. Moreover, the slope of the FCMSTRFM is less than that of the ESTARFM and the STDFA. When the reflectance drastically changes over time, the accuracy of the FCMSTRFM is minimally reduced. This means that the FCMSTRFM has an advantage in regions with drastic temporal changes in reflectance.

$$\sigma^2_{ij} = \frac{\sum_{g=1}^{np} [(k_{ig} - k_{jg}) - (u_i - u_j)]^2}{np - 1}, \quad (20)$$

where np is the number of pixels in the whole fine-resolution image; k_{ig} and k_{jg} are the reflectance of pixel g at times t_i and t_j , respectively; and u_i and u_j are the average reflectance of the whole image at times t_i and t_j , respectively.

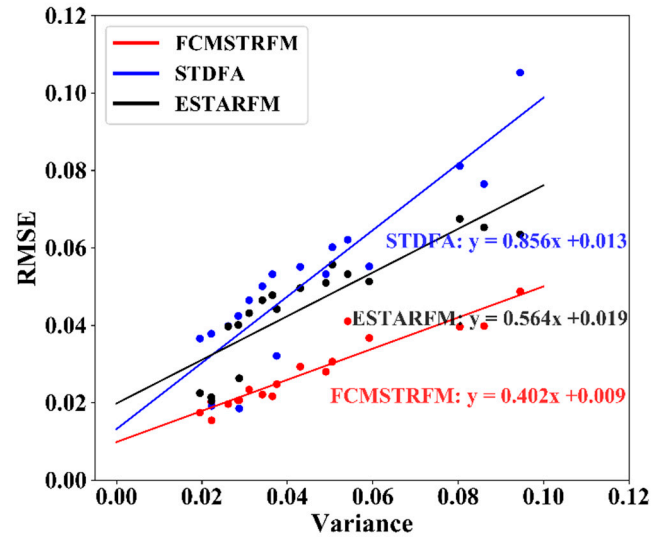


Figure 11. The relationship between the RMSE and the variance for increment of all pixels between two fine-resolution images.

To investigate the sensitivity of the three models to the spatial heterogeneity of reflectance, we chose the variance for all pixels in the actual image as the measure of pixel spatial heterogeneity and RMSE as the measure of accuracy of the generated images. Figure 12 shows that for the three models, the variance for all the pixels in the actual image had a positive correlation with RMSE: the variance increased with the RMSE, indicating that the accuracy of the images generated by STFMs decreased with reflectance spatial heterogeneity. Moreover, the slopes of the FCMSTRFM and the ESTARFM were less than that of the STDFA. When reflectance spatial heterogeneity is complex, the accuracies of the FCMSTRFM and the ESTARFM are minimally reduced.

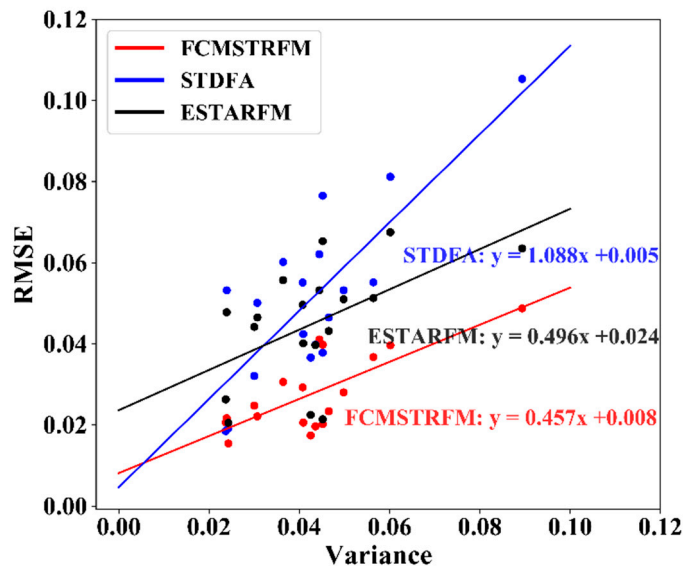


Figure 12. The relationship between the RMSE and the variance of the actual image.

5.2. FCMSTRFM Improvements to Existing Models

Compared with other STFMs, the FCMSTRFM shows the following three distinct advantages. First, the FCMSTRFM develops a new method called PHC that combines land cover map and time-series data cluster methods to group high-resolution pixels. PHC has the capacity, which is not available in other UMX-based methods such as the STDFA, to easily group together pixels that have

similar temporal reflectance changes. The land cover map is a simple and effective division of surface information [53–55] and is unchanged in a short time period. Therefore, land cover maps can be used to coarsely classify time-series images [56–58]. Time-series images contain more information than a single image, and time-series data cluster methods can maximize the similarity of pixels in the same class while minimizing the similarity of different classes [4,59]. Therefore, time-series data cluster methods can be used to subclassify land cover classes.

The similarity of pixels that are grouped together directly determines the accuracy of UMX-based models [39,52]. To evaluate PHC adopted by the FCMSTRFM, we compared three clustering methods: directly using the FCM, the clustering method applied in the STDFA, and PHC applied in the FCMSTRFM. We applied these three methods to cluster high-resolution data into 50 categories (same to the number of subclass), and the similarity of categories was evaluated by the average variance of each category. Table 5 shows the average variance of 50 categories for each method. It is clear that PHC performs better than the other two methods, which simultaneously demonstrates that the use of land cover maps can improve similarity pixel clusters. PHC can provide a reference for obtaining the endmember fraction matrix.

Table 5. The average variance of subclass of each band for the three models.

Band	STDFA	FCM	FCMSTRFM
B3	0.0137	0.0067	0.0059
B4	0.0111	0.0078	0.0064
B5	0.0084	0.0035	0.0022

Second, the FCMSTRFM is suitable for regions where land cover types change over time. Many STFM are based on the assumption that land cover types do not change over time. If the land cover types were to change over time, the accuracy of those STFM would be substantially reduced [22,25,28,33]. The FCMSTRFM uses multiphase images and land cover maps to capture information about changes in land cover type and to group similar pixels, and it is easy to ensure that the pixels grouped together have the same land cover type within the whole period. Influenced by accuracy of land cover mapping, there is a possible that land cover mapping not can ensure the pixels grouped together have the same land cover type within the whole period. However, we subdivided each land cover class and defined the subdivisions as subclasses; compared with land cover class, with subclass, it is easy to ensure that the pixels grouped together have the same land cover type within the whole period. Third, the FCMSTRFM has higher computational efficiency than that of STARFM-based methods. The main steps of the FCMSTRFM are pixel clustering and class-average reflectance calculation, but STARFM-based methods mainly use similar pixel selection and pixel-by-pixel conversion coefficient calculations [20,28,32]. Compared with the FCMSTRFM, STARFM-based methods have higher time and space algorithm complexity. As a result, the FCMSTRFM would be useful for applications in large study areas.

5.3. The Application of FCMSTRFM

Although the FCMSTRFM was designed for spatiotemporal fusion of reflectance data, it also has the potential to blend other remote sensing data, such as vegetation index (VI) products. To test the applicability of the FCMSTRFM to other products, we assessed its performance in fusing the MODIS L3 product MOD13Q1 (providing products EVI and NDVI) with Landsat VI in the HePing Irrigation area. Then, the blended time-series VI products were used to map rice by the method of spectral correlation similarity (SCS). The precision of VI (Table 6) suggested that the FCMSTRFM can be used to generate high-accuracy VI products. The map result is evaluated by a sample from Google Earth images and field investigations, and the accuracy suggests that VI blended by the FCMSTRFM can be used to map crops. For more detailed information about the method and results of rice mapping, we refer to [60]. Furthermore, Sentinel 2 has a five-day revisit time and higher spatial resolution than

Landsat (10–20 m), and has more cloud-free images than Landsat. The FCMSTRFM increases the similarity of pixels grouped together by input multi-temporal images, and use time-series images to record the pixels temporal change. For the FCMSTRFM, the accuracy of the generated images increases with the number of input images. Thus, we consider the FCMSTRFM more suitable for Sentinel 2 data fusion than other models.

Table 6. The precision index of VI generated by the FCMSTRFM.

VI-DOY	R	RMSE	ERGAS	variance
NDVI-229	0.9305	0.0607	1.7132	0.0037
NDVI-245	0.9028	0.0721	1.9655	0.0052
EVI-229	0.9154	0.0622	1.9547	0.0038
EVI-245	0.8744	0.0748	2.2849	0.0055

We also used the ESTARFM to map rice, but it was difficult to map rice from the VI data generated by the ESTARFM. It can be known from Section 4.2 that to get the benefit from the selection of similar neighboring pixels and the calculation of each similar neighboring pixel weight pixel by pixel, the ESTARFM can keep spatial detail information perfectly (Figure 12) when pixels have strong spatial heterogeneity [20,28]. However, because the ESTARFM only captures the temporal changes in pixels through the calculation of VI by two-phase images, there is a large possibility that STARFM-based models cannot capture the pixel time change process perfectly (Figure 11) when the pixel time changes sharply [20,25,52]. The FCMSTRFM captures each subclass time-series change through multiphase images, so it more easily captures the subclass time-series change process and obtains a class (class that is grouped by generated time-series Landsat-like EVI data) series curve that is more similar to reality. We extracted rice area by SCS between the standard series EVI curve and class series curve, so it more accurately mapped rice.

6. Conclusion

This paper developed a new model, the FCMSTRFM, to generate high-spatial-resolution time-series images using fine- and coarse-resolution images. Compared with the STDFM and the ESTARFM, the FCMSTRFM can generate images more accurately, especially for regions with substantial temporal changes in pixels. The main contributions of this study are as follows. First, we developed a new method to group high-resolution pixels to define endmember terms as PHCs. To the best of our knowledge, this is the first time that land cover maps and time-series data cluster methods are combined to group high-resolution pixels. Second, a new strategy is used to calculate the average reflectance of the subclass, which aims to break the ill-posed spectral unmixing limit to the precision of generating images. The results show that the FCMSTRFM is capable of increasing the similarity of pixels grouped together and the precision of the average reflectance and improving the computational efficiency without loss of precision.

Similar to many UMX-based methods, the precision of the FCMSTRFM is currently mainly limited by grouping similar pixels, calculating the average reflectance and quality of images. We believe that the only way to capture the features of temporal changes in pixels is to input as many images to STDFMs as possible; Sentinel 2 has higher temporal resolution than Landsat and can get more cloud-free images, so we think it more suitable for data fusion. Almost all STDFMs based on clear scene images—that is, images partly contaminated by clouds, are unavailable. In fact, cloud-free pixels in images partly contaminated by clouds can provide valuable information for data fusion. If the denoising time-series data clustering algorithm can be used to group pixels, these images can be used to capture the features of temporal and spatial changes in pixels. Aerosols and BRDF effects impact the quality of images, if atmospheric corrections take into account aerosols, and fusing images can be normalized to nadir BRDF adjusted reflectance in STDFMs, the fusing result will be improved.

Author Contributions: J.Y. and Y.Y. prepared the manuscript. Y.W., Y.Z., K.J. and X.Z. revised the manuscript. K.S., X.B. and X.G. participated the discussion. All authors have read and agreed to the published version of the manuscript.

Funding: This work is partially supported by the National Key Research and Development Program of China (No.2016YFC0400101 and No.2016YFA0600102), the Natural Science Fund of China (No. 41671331 and No.41701483) and BNU Interdisciplinary Research Foundation for the First-Year Doctoral Candidates (Grant No. BNUXKJC1907).

Acknowledgments: We would like to thank USGS for providing Landsat and MODIS data. MOD09GA, MOD13Q1 and Landsat8 OLI products were obtained online (<https://earthdata.nasa.gov/>).

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

References

1. Croft, H.; Anderson, K.; Kuhn, N.J. Evaluating the influence of surface soil moisture and soil surface roughness on optical directional reflectance factors. *Eur. J. Soil Sci.* **2014**, *65*, 605–612.
2. Cuppo, F.L.S.; Garcia-Valenzuela, A.; Olivares, J.A. Influence of surface roughness on the diffuse to near-normal viewing reflectance factor of coatings and its consequences on color measurements. *Color Res. Appl.* **2013**, *38*, 177–187.
3. Sun, Z.Q.; Wu, D.; Lv, Y.F.; Zhao, Y.S. Bidirectional Polarized Reflectance Factors of Vegetation Covers: Influence on the BRDF Models Results. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5687–5701.
4. Li, C.X.; Sun, Z.; Jiang, J.Y.; Liu, R.; Chen, W.L.; Xu, K.X. Typical ground object recognition based on principle component analysis and fuzzy clustering with near-infrared diffuse reflectance spectroscopy. *Spectrosc. Spect. Anal.* **2017**, *37*, 3386–3390.
5. Zhong, L.H.; Hu, L.; Zhou, H.; Tao, X. Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US. *Remote Sens. Environ.* **2019**, *233*, doi:10.1016/j.rse.2019.111411.
6. Lee, T.Y.; Kaufman, Y.J. Non-lambertian effects on remote-sensing of surface reflectance and vegetation index. *IEEE Trans. Geosci. Remote Sens.* **1986**, *24*, 699–708, doi:10.1109/Tgrs.1986.289617.
7. Wang, H.; Yang, L.K.; Zhao, M.R.; Du, W.B.; Liu, P.; Sun, X.B. The normalized difference vegetation index and angular variation of surface spectral polarized reflectance relationships: Improvements on aerosol remote sensing over land. *Earth Space Sci.* **2019**, *6*, 982–989.
8. Spanner, M.A.; Pierce, L.L.; Peterson, D.L.; Running, S.W. Remote-sensing of temperate coniferous forest leaf-area index—The influence of canopy closure, understory vegetation and background reflectance. *Int. J. Remote Sens.* **1990**, *11*, 95–111, doi:10.1080/01431169008955002.
9. Zhai, H.; Huang, F.; Qi, H. Generating High Resolution LAI Based on a modified FSDAF model. *Remote Sens.* **2020**, *12*, 150.
10. Ma, R.; Zhang, L.; Tian, X.J.; Zhang, J.C.; Yuan, W.P.; Zheng, Y.; Zhao, X.; Kato, T. Assimilation of remotely-sensed leaf area index into a dynamic vegetation model for gross primary productivity estimation. *Remote Sens.* **2017**, *9*, doi:10.3390/rs9030188.
11. Zheng, Y.; Zhang, L.; Xiao, J.F.; Yuan, W.P.; Yan, M.; Li, T.; Zhang, Z.Q. Sources of uncertainty in gross primary productivity simulated by light use efficiency models: Model structure, parameters, input data, and spatial resolution. *Agr. Forest Meteorol.* **2018**, *263*, 242–257, doi:10.1016/j.agrformet.2018.08.003.
12. Merlin, O.; Chirouze, J.; Olioso, A.; Jarlan, L.; Chehbouni, G.; Boulet, G. An image-based four-source surface energy balance model to estimate crop evapotranspiration from solar reflectance/thermal emission data (SEB-4S). *Agr. Forest Meteorol.* **2014**, *184*, 188–203.
13. Xu, J.; Yao, Y.J.; Liang, S.L.; Liu, S.M.; Fisher, J.B.; Jia, K.; Zhang, X.T.; Lin, Y.; Zhang, L.L.; Chen, X.W. Merging the MODIS and landsat terrestrial latent heat flux products using the multiresolution tree method. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2811–2823.
14. Yao, Y.J.; Qin, Q.M.; Ghulam, A.; Liu, S.M.; Zhao, S.H.; Xu, Z.W.; Dong, H. Simple method to determine the Priestley-Taylor parameter for evapotranspiration estimation using Albedo-VI triangular space from MODIS data. *J. Appl. Remote Sens.* **2011**, *5*, doi:10.1117/1.3557817.
15. Goswami, S.; Gamon, J.A.; Tweedie, C.E. Surface hydrology of an arctic ecosystem: Multiscale analysis of a flooding and draining experiment using spectral reflectance. *J. Geophys. Res. Biogeosci.* **2011**, *116*, doi:10.1029/2010JG001346.
16. Jin, J.; Wang, Q.; Wang, J.L.; Otieno, D. Tracing water and energy fluxes and reflectance in an arid ecosystem using the integrated model SCOPE. *J. Environ. Manag.* **2019**, *231*, 1082–1090.

17. Yao, Y.J.; Liang, S.L.; Li, X.L.; Chen, J.Q.; Wang, K.C.; Jia, K.; Cheng, J.; Jiang, B.; Fisher, J.B.; Mu, Q.Z.; et al. A satellite-based hybrid algorithm to determine the Priestley-Taylor parameter for global terrestrial latent heat flux estimation across multiple biomes (vol 165, pg 216, 2015). *Remote Sens. Environ.* **2015**, *169*, 454.
18. Yao, Y.J.; Liang, S.L.; Qin, Q.M.; Wang, K.C. Monitoring drought over the conterminous United States using MODIS and NCEP reanalysis-2 data. *J. Appl. Meteorol. Clim.* **2010**, *49*, 1665–1680.
19. Xue, J.; Leung, Y.; Fung, T. An unmixing-based bayesian model for spatio-temporal satellite image fusion in heterogeneous landscapes. *Remote Sens.* **2019**, *11*, 324.
20. Gevaert, C.M.; Garcia-Haro, F.J. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* **2015**, *156*, 34–44.
21. Emelyanova, I.V.; McVicar, T.R.; Van Niel, T.G.; Li, L.T.; van Dijk, A.I.J.M. Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. *Remote Sens. Environ.* **2013**, *133*, 193–209.
22. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.
23. Sun, R.; Chen, S.H.; Su, H.B.; Mi, C.R.; Jin, N. The Effect of NDVI time series density derived from spatiotemporal fusion of multisource remote sensing data on crop classification accuracy. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, doi:10.3390/ijgi8110502.
24. Ehlers, M. Multisensor image fusion techniques in remote-sensing. *ISPRS J. Photogramm. Remote Sens.* **1991**, *46*, 19–30.
25. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.G.; Gao, F.; White, J.C. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627.
26. Wang, T.; Tang, R.L.; Li, Z.L.; Jiang, Y.Z.; Liu, M.; Niu, L. An Improved spatio-temporal adaptive data fusion algorithm for evapotranspiration mapping. *Remote Sens.* **2019**, *11*, 761.
27. Xia, H.P.; Chen, Y.H.; Li, Y.; Quan, J.L. Combining kernel-driven and fusion-based methods to generate daily high-spatial-resolution land surface temperatures. *Remote Sens. Environ.* **2019**, *224*, 259–274.
28. Zhu, X.L.; Chen, J.; Gao, F.; Chen, X.H.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623.
29. Cui, J.T.; Zhang, X.; Luo, M.Y. Combining linear pixel unmixing and STARFM for spatiotemporal fusion of gaofen-1 wide field of view imagery and MODIS imagery. *Remote Sens.* **2018**, *10*, doi:10.3390/rs10071047.
30. Roy, D.P.; Ju, J.; Lewis, P.; Schaaf, C.; Gao, F.; Hansen, M.; Lindquist, E. Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3112–3130.
31. Walker, J.J.; de Beurs, K.M.; Wynne, R.H.; Gao, F. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sens. Environ.* **2012**, *117*, 381–393.
32. Xie, D.F.; Zhang, J.S.; Zhu, X.F.; Pan, Y.Z.; Liu, H.L.; Yuan, Z.M.Q.; Yun, Y. An improved STARFM with help of an unmixing-based method to generate high spatial and temporal resolution remote sensing data in complex heterogeneous regions. *Sensors* **2016**, *16*, doi:10.3390/s16020207.
33. Wu, M.Q.; Wu, C.Y.; Huang, W.J.; Niu, Z.; Wang, C.Y.; Li, W.; Hao, P.Y. An improved high spatial and temporal data fusion approach for combining Landsat and MODIS data to generate daily synthetic Landsat imagery. *Inform. Fusion* **2016**, *31*, 14–25.
34. Liu, M.; Yang, W.; Zhu, X.L.; Chen, J.; Chen, X.H.; Yang, L.Q.; Helmer, E.H. An Improved Flexible Spatiotemporal DATA Fusion (IFSADF) method for producing high spatiotemporal resolution normalized difference vegetation index time series. *Remote Sens. Environ.* **2019**, *227*, 74–89.
35. Wu, M.Q.; Niu, Z.; Wang, C.Y.; Wu, C.Y.; Wang, L. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model. *J. Appl. Remote Sens.* **2012**, *6*, doi:10.1117/1.Jrs.6.063507.
36. Zurita-Milla, R.; Kaiser, G.; Clevers, J.G.P.W.; Schneider, W.; Schaepman, M.E. Downscaling time series of MERIS full resolution data to monitor vegetation seasonal dynamics. *Remote Sens. Environ.* **2009**, *113*, 1874–1885, doi:10.1016/j.rse.2009.04.011.
37. Zhong, D.T.; Zhou, F.Q. Improvement of clustering methods for modelling abrupt land surface changes in satellite image fusions. *Remote Sens.* **2019**, *11*, 1759.
38. Zurita-Milla, R.; Clevers, J.G.P.W.; Schdepman, M.E. Unmixing-based Landsat TM and MERIS FR data fusion. *IEEE Trans. Geosci. Remote Sens.* **2008**, *5*, 453–457.

39. Zhukov, B.; Oertel, D.; Lanzl, F.; Reinhackel, G. Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1212–1226.
40. Maselli, F. Definition of spatially variable spectral endmembers by locally calibrated multivariate regression analyses. *Remote Sens. Environ.* **2001**, *75*, 29–38.
41. Zurita-Milla, R.; Gomez-Chova, L.; Guanter, L.; Clevers, J.G.P.W.; Camps-Valls, G. Multitemporal unmixing of medium-spatial-resolution satellite images: A case study using MERIS images for land-cover mapping. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 4308–4317.
42. Comber, A.J.; Law, A.N.R.; Lishman, J.R. Application of knowledge for automated land cover change monitoring. *Int. J. Remote Sens.* **2004**, *25*, 3177–3192, doi:10.1080/01431160310001657795.
43. Hansen, M.C.; Defries, R.S.; Townshend, J.R.G.; Sohlberg, R. Global land cover classification at 1km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **2000**, *21*, 1331–1364, doi:10.1080/014311600210209.
44. King, R.B. Land cover mapping principles: a return to interpretation fundamentals. *Int. J. Remote Sens.* **2002**, *23*, 3525–3545, doi:10.1080/01431160110109606.
45. Jokinen, J.; Raty, T.; Lintonen, T. Clustering structure analysis in time-series data With density-based clusterability measure. *IEEE/CAA J. Autom. Sin.* **2019**, *6*, 1332–1343, doi:10.1109/Jas.2019.1911744.
46. Wang, Y.; Ru, Y.N.; Chai, J.P. Time series clustering based on sparse subspace clustering algorithm and its application to daily box-office data analysis. *Neural Comput. Appl.* **2019**, *31*, 4809–4818, doi:10.1007/s00521-018-3731-7.
47. Zhang, L.; Weng, Q.H.; Shao, Z.F. An evaluation of monthly impervious surface dynamics by fusing Landsat and MODIS time series in the Pearl River Delta, China, from 2000 to 2015. *Remote Sens. Environ.* **2017**, *201*, 99–114.
48. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM - the Fuzzy C-Means Clustering-Algorithm. *Comput. Geosci. UK* **1984**, *10*, 191–203.
49. Rodriguez, A.; Tomas, M.S.; Rubio-Martinez, J. A benchmark calculation for the fuzzy c-means clustering algorithm: initial memberships. *J. Math. Chem.* **2012**, *50*, 2703–2715.
50. Saxena, A.; Prasad, M.; Gupta, A.; Bharill, N.; Patel, O.P.; Tiwari, A.; Er, M.J.; Ding, W.P.; Lin, C.T. A review of clustering techniques and developments. *Neurocomputing* **2017**, *267*, 664–681.
51. Garcia-Haro, F.J.; Sommer, S.; Kemper, T. A new tool for variable multiple endmember spectral mixture analysis (VMESMA). *Int. J. Remote Sens.* **2005**, *26*, 2135–2162.
52. Yang, J.M.; Wu, Y.; Wei, Y.X.; Wang, B.; Ru, C.; Ma, Y.Y.; Zhang, Y. A model for the fusion of multi-source data to generate high temporal and spatial resolution VI data. *J. Remote Sens.* **2019**, *23*, 935–943, doi:10.11834/jrs.20198204.
53. Jin, X.; Jin, Y.X.; Yuan, D.H.; Mao, X.F. Effects of land-use data resolution on hydrologic modelling, a case study in the upper reach of the Heihe River, Northwest China. *Ecol. Model.* **2019**, *404*, 61–68.
54. Pokonieczny, K.; Moscicka, A. The Influence of the shape and size of the cell on developing military passability maps. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 261.
55. Weigand, M.; Staab, J.; Wurm, M.; Taubenbock, H. Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *88*, doi:10.1016/j.jag.2020.102065.
56. Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.R.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.T.A.; et al. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. *Remote Sens.* **2016**, *8*, doi:10.3390/rs8010070.
57. Lei, G.B.; Li, A.N.; Bian, J.H.; Zhang, Z.J. The roles of criteria, data and classification methods in designing land cover classification systems: Evidence from existing land cover data sets. *Int. J. Remote Sens.* **2020**, *41*, 5062–5082, doi:10.1080/01431161.2020.1724349.
58. Xu, G.; Zhang, H.R.; Chen, B.Z.; Zhang, H.F.; Yan, J.W.; Chen, J.; Che, M.L.; Lin, X.F.; Dou, X.M. A Bayesian based method to generate a synergetic land-cover map from existing land-cover products. *Remote Sens.* **2014**, *6*, 5589–5613, doi:10.3390/rs6065589.

59. Liu, Y.L.; Wang, X.M.; Liu, Q.L.; Chen, Y.Y.; Liu, L.L. An improved density-based time series clustering method based on image resampling: A case study of surface deformation pattern analysis. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, doi:10.3390/ijgi6040118.
60. Wei, Y.X.; Yang, J.M.; Wu, Y.; Wang, B.; Shaban, M.; Hou, J.X. Rice planting area extraction based on multi-source data fusion. *Trans. Chin. Soc. Agric. Machin.* **2018**, *49*, 300–306.



© 2020 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/>).