Reconstructing Missing Information of Remote Sensing Data Contaminated by Large and Thick Clouds Based on an Improved Multitemporal Dictionary Learning Method

Mu Xia and Kun Jia

***Abstract*— The presence of clouds and cloud shadows has limited the applications of optical remote sensing data. Currently, most cloud removal methods are focused on reconstructing remote sensing data contaminated by small or thin clouds. This study proposes an improved method based on multitemporal dictionary learning to reconstruct missing information of remote sensing data contaminated by large and thick clouds. First, the contaminated target image is initialized using all available adjacent cloud-free reference images. Second, reconstructed images from each of the reference images are produced using dictionary learning and sparse representation methods. Then, weights are determined for the abovementioned reconstructed images based on their reconstruction errors over uncontaminated regions and are used to generate the preliminary reconstruction result. Finally, an error correction step for the contaminated regions is applied to the preliminary result, which is then combined with the original uncontaminated pixels to produce the final reconstruction result. The proposed method was evaluated on simulated clouds/cloud shadows based on remote sensing data with various sizes and land cover types. Visual and quantitative analyses of the reconstruction results show that the proposed method outperformed the generally used geostatistical neighborhood similar pixel interpolator (GNSPI) and nonnegative matrix factorization and error correction (S-NMF-EC) methods. Therefore, the results indicated that the proposed method was capable of accurately and effectively reconstructing data contaminated by large and thick clouds.**

***Index Terms*— Missing information, multitemporal dictionary learning (MDL), reconstruction, sparse representation.**

I. Introduction

With the development of remote sensing technology, software computational efficiency, and hardware capacity, remote sensing data have been widely used in a variety of applications, including precision agriculture [1], biodiversity assessment [2], hydrological modeling [3], geologic mapping [4], and hazard assessment [5]–[7]. Among all the sensors, optical remote sensing systems have the longest history as well as the largest data archives. Recently, it has been convenient to acquire medium- to high spatial resolution optical remote sensing data all over the world at short time intervals. However, clouds form at different altitudes throughout the atmosphere, which cast great influence on optical remote sensing data. Thin clouds do not block the spectral information of the earth's surface entirely, and part of the spectral signals can still pass through and reach the satellite sensors. However, thick clouds obscure all spectral information of the earth's surface at visible and near-infrared wavelengths [8]. Although the remote sensing of clouds is important for atmospheric studies [9], the presence of clouds and cloud shadows has long been one of the greatest obstacles in other optical remote sensing data applications, because they cover approximately 50% of the earth's surface at any time [10].

Many algorithms for removing cloud/cloud shadow effects and recovering cloud-contaminated information from optical remote sensing data have been developed in recent decades. They can be classified into four categories based on the auxiliary information needed, including spatial-based, spectral-based, temporal-based, and hybrid methods [11].

Spatial-based methods utilize information from the cloud-uncontaminated regions to reconstruct the contaminated regions from the same image. Typical algorithms include geostatistical interpolation [12], [13], propagated diffusion [14], [15], and variation-based [16] and exemplar-based [17], [18] methods. These methods assume that the spectral characteristics over both contaminated and uncontaminated regions are similar. Therefore, they only work well for images with small cloud contamination in homogeneous regions. However, clouds and cloud shadows seldom appear so idealized, which limits their actual applications.

Spectral-based methods are developed based on the existing high correlation among neighboring spectral bands and the different abilities of electromagnetic radiation to penetrate clouds at different wavelengths. The main applications of these methods are retrieving a corrupted band using its neighboring

Manuscript received December 30, 2020; revised March 14, 2021 and April 16, 2021; accepted July 1, 2021. Date of publication July 16, 2021; date of current version January 17, 2022. This work was supported in part by the Second Tibetan Plateau Scientific Expedition and Research Program under Grant 2019QZKK0405, in part by the National Natural Science Foundation of China under Grant 41671332, and in part by the National Key Research and Development Program of China under Grant 2016YFB0501404 and Grant 2016YFA0600103. The work of K. Jia was supported by the Tang Scholar of Beijing Normal University. *(Corresponding author: Kun Jia.)*

The authors are with the State Key Laboratory of Remote Sensing Science and the Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China (e-mail: jiakun@bnu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2021.3095067

intact bands [19] and recovering land surface information covered by thin clouds [20]–[22]. However, when thick clouds obstruct the earth surface information over all bands from reaching to the sensors, these methods are not applicable.

Temporal-based methods utilize uncontaminated reference information over the same region with different acquisition times. This category of methods assumes that the land cover types remain unchanged over the acquisition times of all images. Some of the temporal-based methods predict the contaminated data using similar patches were determined by reference data, e.g., Lin *et al.* [23] proposed information cloning by selecting the most similar patches over a sequence of images and formulated the reconstruction process as a global optimization of solving a Poisson equation. Other algorithms predict cloud-contaminated pixels using established relationships between uncontaminated data over multitemporal images [24], [25]. Methods such as the Savitzky–Golay (SG) filter [26], asymmetric Gaussian (AG) model [27], double logistic (DL) technique [28], and harmonic analysis of time series (HANTS) [29] are often used for eliminating noise over time-series data, typically for vegetation indices [30]–[33]. Since vegetation growth follows a certain temporal pattern, these methods can filter out outliers and provide a proper prediction of the missing data according to the pattern curves. However, not all pixels follow temporal patterns like vegetation. In addition, the abovementioned methods are developed mainly based on mathematical derivation processes, which is restricted by the existing knowledge of mathematical axioms. Generally, temporal-based methods are capable of producing better reconstruction results and are more practical than the spatial- and spectral-based methods.

Hybrid methods integrate information from spatial, spectral, and temporal domains. Some hybrid methods estimate contaminated data using their "similar pixels" selected from intact reference images with adjacent acquisition dates. Among them, various methods have been introduced to determine the "similar pixels." The neighborhood similar pixel interpolator (NSPI) [34] and the geostatistical NSPI (GNSPI) [35], [36] methods select similar pixels for each target pixel based on the spectral similarity from time-series data and utilize a geostatistical approach to predict missing data. Other methods select "similar pixels" using methods including spatial–temporal Markov random fields (MRFs) [37], spectral-angle-mapper-based spatial–temporal similarity (SAMSTS) [38], etc. Apart from generating estimations using "similar pixels," many other hybrid methods are proposed in recovering contaminated data. Melgani [39] recovered the contaminated data using an unsupervised contextual prediction process based on established spectral–temporal relationships. Deep learning methods are also adopted to develop missing information reconstruction algorithms. For example, a unified spatial–temporal–spectral deep convolutional neural network (STS-CNN) method was recently proposed to recover corrupted data in remote sensing images [40]. This method can deal with multiple tasks using a unified framework, including reconstructing cloud/cloud shadow contaminated data, Landsat ETM+ SLC-off data, and deadlines in Aqua MODIS band 6. The CNN is trained using samples selected among the time series Landsat Thematic Mapper (TM) image. However, the trained network may not be applicable for other time periods, geographical locations, or data from other sensors because the network relies largely on the selected samples. The Hidden Markov random field (HMRF) framework involving spatial–temporal–spectral and environmental information was also proposed to improve cloud-contaminated MODIS snow products [41]. The HMRF-based framework successfully reduced the cloud cover-related data gap below 1% and improved the current snow mapping accuracies using energy functions in spectral, spatial–temporal, and environment-associated information. Moreover, in recent years, an increasing number of studies have employed dictionary learning and sparse representation methods for cloud removal. This category of methods is capable of recovering sparse signals with high accuracy from a small set of random measurements by solving a linear program [42]–[44]. Lorenzi *et al.* [15] introduced three dictionary learning-based methods utilizing basic pursuit (BP), orthogonal matching pursuit [45], and genetic algorithms (GAs), and achieved satisfactory performances in recovering missing information over different land cover types. Multitemporal dictionary learning (MDL) [46], group sparse representation (GSR) [47], [48], and nonnegative matrix factorization and error correction method (S-NMF-EC) [49] are representative methods of this kind. However, most dictionary learning-based methods change pixel values within the uncontaminated areas along with the reconstruction process. In other words, the reconstruction results would suffer from obvious reconstruction edges between contaminated and uncontaminated regions if the original uncontaminated pixel values are retained. In general, hybrid methods perform better and are more promising in reconstructing contaminated data by integrating information from multiple domains.

Most of the abovementioned methods perform well over the study areas with small amount of clouds. However, in the real situation, clouds often appear in large chunks. Therefore, fusion methods based on auxiliary data from multiple sensors are often used for heavily contaminated images. However, the fusion methods always require time and effort consuming data preprocessing and will also introduce uncertainties from other sensors due to differences in the spectral design and spatial scales. These factors increase the difficulties for the fusion methods to be further applied in real time-series data production. As discussed above, dictionary learning-based methods can greatly improve the computational efficiency by representing features and their intrinsic relationships using representative dictionary atoms and sparse coefficients. Therefore, an improved method to reconstruct missing information contaminated by large and thick clouds based on MDL with the following advantages is proposed. First, it is capable of reconstructing remote sensing data contaminated by large and thick clouds/cloud shadows with high efficiency. Second, it improves previous MDL-based methods in reconstruction accuracy by introducing weighting and error correction steps. Third, our method avoids the application of multisensor data and retains true values within cloud/cloud shadow uncontaminated areas.

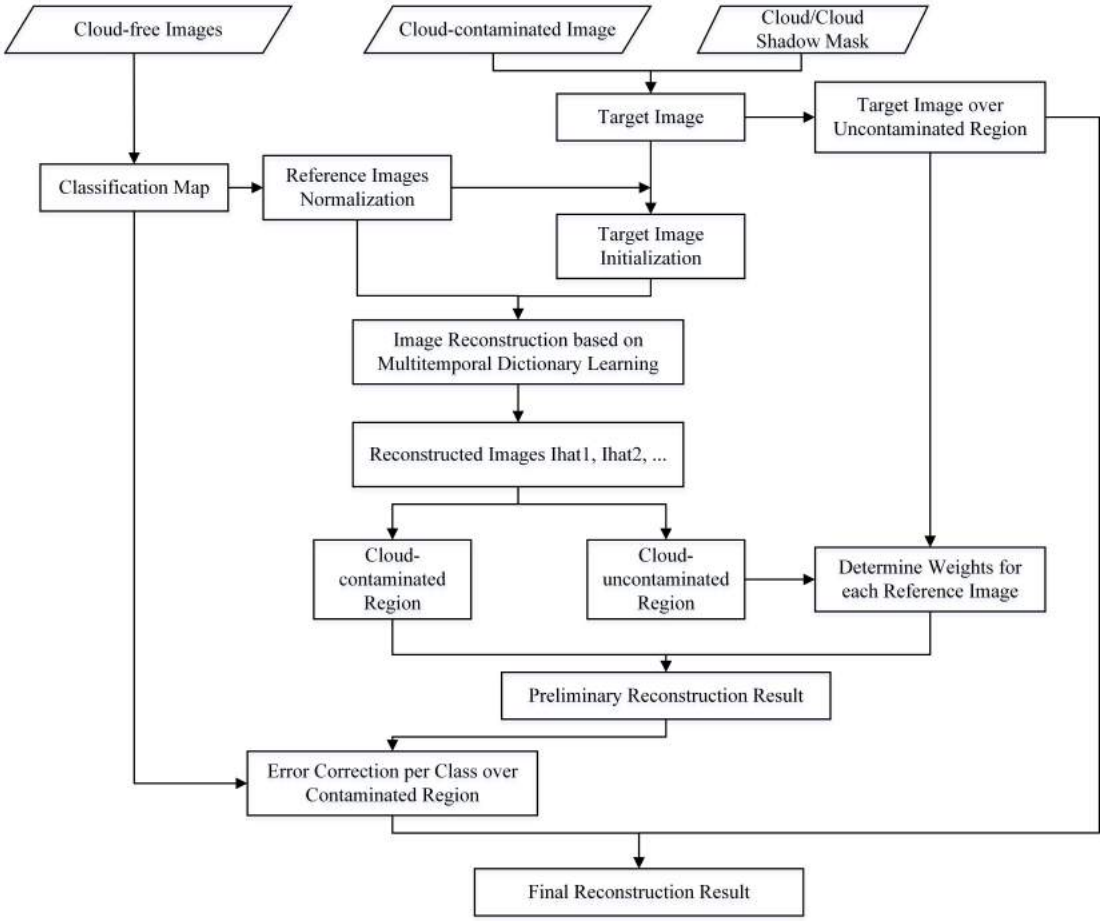

Fig. 1. Flowchart of our method.

II. Methodology

The algorithm flowchart of our method is shown in Fig. 1. First, the cloud/cloud shadow contaminated image, referred to as the target image, is initialized using selected multitemporal cloud-free reference images. K-means classification algorithm is adopted to obtain the classification map based on cloud-free images. Then, based on each reference image, the initialized target image is reconstructed using dictionary learning and sparse representation methods. Next, a weighting step is introduced to determine the weights of all the reconstructed images from each reference. The preliminary reconstructed image is produced by the weighted sum from the corresponding reconstructed images. An error correction step is then applied to the preliminary reconstructed image to reduce the errors between contaminated and uncontaminated regions. Finally, the original uncontaminated pixels are combined with the reconstructed contaminated pixels to produce the final reconstruction result.

To illustrate the procedure, the target image is referred to as $I^t \in R^{M\times N\times Q}$, while cloud-free reference images covering the same area with different acquisition times are referred to as $I^{r(1)}, I^{r(2)}, \ldots I^{r(n)} \in R^{M\times N\times Q}$, where n is the number of reference images and M, N, and Q are the number of rows, columns, and spectral bands of each image, respectively. Acquisition dates t_2–t_n are aligned from the closest to the farthest to t_1. As is shown in Fig. 2, the cloud/cloud shadow contaminated area of the target image is referred to as Ω_+ and the uncontaminated area is referred to as Ω. The corresponding areas in the reference images are also referred to as Ω_+ and Ω. A pixel belonging to the target image is denoted as x_t, while that of a reference image is denoted as x_r.

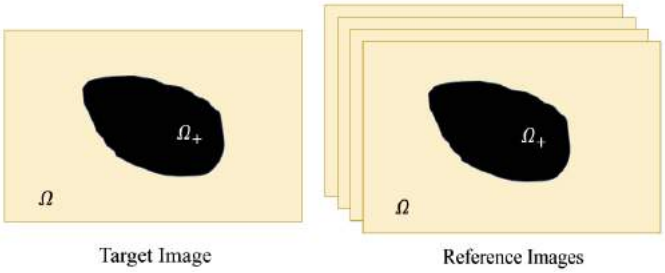

Fig. 2. Cloud/cloud shadow contaminated area Ω_+ and uncontaminated area Ω on target and reference images.

In our method, all the pixels contaminated by clouds/cloud shadows are masked out and substituted with a value of 0. Due to the complexity of cloud types and properties, cloud/cloud shadow detection is not part of this study, and all cloud/cloud shadow masks are either manually simulated or extracted from the Landsat QA bands by the function of mask (FMASK) algorithm [50].

A. Normalization of Reference Images and the Initialization of Target Images

For our method, it is assumed that the land cover types on the target and reference images remain unchanged. Due to different atmospheric conditions, sensor statuses, and vegetation-growing stages during image acquisition times, even temporally close remote sensing data experience reflectance distinctions over the same land cover types. To alleviate these differences, normalization was first applied to all the reference images based on the target image. Since different land cover types show different temporal changes of spectral, a simple k-means cluster algorithm was applied to classify the reference image acquired at t_2 into C classes. The least squares fitting is then applied to each class to establish relationships between the reference image and the target image

$$x(i, j, q) = a \times x^r(i, j, q) + b \tag{1}$$

where $x^r(i, j, q)$ is the value at the ith row, jth column, and qth band in a reference image; $x(i, j, q)$ is the normalized value from x^r; and a and b are the least squares fitting coefficients calculated within each class by

$$\begin{aligned} a = &\Big(MN \sum \big(I^r(i, j, q) \times I^t(i, j, q)\big) \\ &- \sum I^r(i, j, q) \\ &\times \sum I^t(i, j, q)\Big) \Big/ \Big(MN \sum \big(I^r(i, j, q)\big)^2 \\ &- \Big(\sum I^r(i, j, q)\Big)^2\Big) \end{aligned} \tag{2}$$

$$b = \frac{\sum I^t(i, j, q) - MN \sum I^r(i, j, q)}{MN}. \tag{3}$$

It should be noted that the least square coefficients are calculated based only on the pixels within Ω. After all the reference images are normalized, a simple arithmetic average is applied to all the pixels over region Ω_+ in the normalized reference images. The results obtained are used to substitute values in the target image over region Ω_+ as initial values for the missing information. Therefore, the initialized target image X^t is obtained.

B. Image Reconstruction Based on the Dictionary Learning Method

Before implementing dictionary learning on multispectral remote sensing data, a patch extraction process should be carried out first to obtain the patch matrix (Fig. 3). The patch matrix is a combination of patches extracted from the original image $I \in R^{M \times N \times Q}$ by moving an $r \times r$ sliding window with a step length of l. Each column in the patch matrix denotes to a patch vector, representing features from the original data. We assume that the land cover types remain unchanged, and the pixel spectral characteristics are similar among target and reference images. Therefore, the patches were selected using a fixed sliding window with a step length of 1. In this way, each column in the patch matrix refers to the reflectance of a pixel over all spectral bands. The patch extraction process can be expressed as

$$x^{2D}(q, (j-1) \times M + i) = P(x(i, j, q)) \tag{4}$$

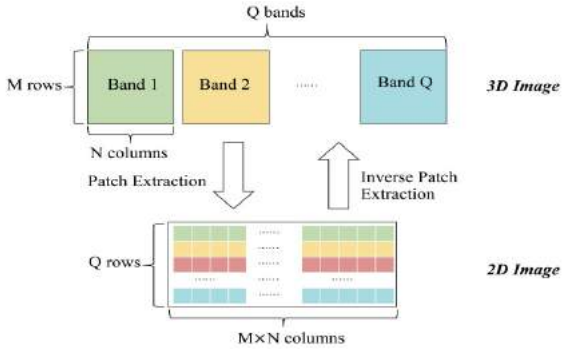

Fig. 3. Patch extraction process of a remote sensing image.

where $P(\cdot)$ is the patch extraction operator; $x(i, j, q)$ is a matrix element in the ith row, jth column, and qth spectral band of an image $I \in R^{M \times N \times Q}$; and x^{2D} is the same element in the new transformed patch matrix. Based on the patch matrices extracted from the initialized target image and reference images P_t and P_r, dictionary learning and sparse representation are then carried out.

Image reconstruction based on dictionary learning requires two components, the dictionary matrix D and the sparse matrix α; both are learned from the patch matrix. A patch matrix $P \in R^{Q \times (\mathrm{MN})}$ can be expressed by the product of $D \in R^{Q \times k}$ and $\alpha \in R^{k \times (\mathrm{MN})}$ as

$$P = D\alpha. \tag{5}$$

In the dictionary matrix, the columns are referred to as dictionary atoms, which show the representative features of columns in P. Each dictionary atom holds Q elements. The number of dictionary atoms K should be determined by the user. Typically, $K > Q$. The sparse matrix α, in which zero elements greatly outnumber nonzero elements, can be used to determine which dictionary atoms are related to each pixel in P and their corresponding significances. All the calculations regarding to dictionary learning and sparse representation in this study was carried out using the SPArse Modeling Software (SPAMS) toolbox [51].

Suppose $P = [p_1, p_2, \ldots, p_{\mathrm{MN}}]$, $D = [d_1, d_2, \ldots, d_k]$, and $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_{\mathrm{MN}}]$, then D and α should satisfy

$$\min_{D,\alpha} \sum_{i=1}^{t} \frac{1}{2} \| p_i - D\alpha_i \|_2^2, \text{ subject to } D \geq 0, \quad \forall \alpha_i \geq 0. \tag{6}$$

This is a joint minimization problem in terms of D and α. The problem can be transformed to a convex optimization to solve each value of D and α when the other is fixed. When D is given, α can be acquired by the sparse coding methods. When α is given, D can be calculated using dictionary learning methods. The online dictionary learning (ODL) algorithm, based on stochastic approximations, is commonly applied for this process. ODL performs faster and trains better dictionaries for large datasets than other classical methods [52]. In this study, NMF, a variation of the ODL algorithm, is adopted. The NMF adds constraints to ODL, which guarantee all components in the dictionary and sparse matrices to be nonnegative [53]. Thus, it is meaningful to reconstruct the surface

reflectance data. The detailed steps are described as follows. First, the initial dictionary D is acquired by randomly selecting K columns from the original dataset, and the sparse coding step is followed to find the sparse matrix α based on the least angle regression (LARS) algorithm [54]. Then, the dictionary matrix is updated to find the solution to (5) when fixing α. These steps are carried out iteratively to update D and α until a convergence is met. More details about the algorithm can be found in [51].

Based on the patch matrices of the initialized target image and reference images P_t and P_r, dictionary and sparse matrices for each image D_t, α_t, D_r, and α_r could be learned separately using the abovementioned method by

$$P_t = D_t \alpha_t \tag{7}$$

$$P_r = D_r \alpha_r \tag{8}$$

where D_t and α_t are the dictionary and sparse matrices for the target patch, and D_r and α_r are those for a reference image.

As is assumed that land cover types remain unchanged among the target and reference images, the relationships between each pixel and dictionary atoms, represented by the sparse matrix, are also considered unchanged. Therefore, the reconstructed image patch from image Y^P can be expressed as

$$Y^P = D_t \alpha_r. \tag{9}$$

The reconstructed image from a reference image y can be obtained by the inversion of patch extraction operator $P^{-1}(\cdot)$ using Y^P

$$y = P^{-1}\left(Y^P\right). \tag{10}$$

In this step, multiple reconstructed images are generated based on multitemporal reference images. The integration of these reconstruction results is discussed in Section II-C. It should be noted that this step reconstructs the whole image, regardless of whether the pixel belongs to Ω or Ω_+. Thus, all the pixel values have been changed over the whole image.

C. Weight Calculations for the Reference Images

Even though image normalization is applied to all the reference images in Section II-B, temporal differences are not completely eliminated. Therefore, a weighting step is introduced to determine the weights of multiple reconstruction results based on their reconstruction accuracies over region Ω. Thus, the errors between the reconstruction results and their corresponding original data within Ω are calculated. Less error indicates a greater weight of the reconstructed image.

The mean absolute error (MAE) of the reconstruction results for uncontaminated area $I^t(\Omega)$ is selected as the accuracy indicator. The weight of the sth reference image $w(s)$ is calculated using

$$w(s) = \frac{\sum_1^n \mathrm{MAE}_\Omega - \mathrm{MAE}_\Omega(s)}{(n-1)\sum_1^n \mathrm{MAE}_\Omega} \tag{11}$$

$$\sum_1^n w = 1 \tag{12}$$

where n is the number of reference images, and the sum of weights for n reference images is 1, and $\mathrm{MAE}_\Omega(s)$ is the MAE within region Ω between the sth reconstructed image $y^s(\Omega)$ and the target image $I^t(\Omega)$

$$\mathrm{MAE}_\Omega(s) = \frac{1}{m}\sum_1^m \left| I^t(\Omega) - y^s(\Omega) \right| \tag{13}$$

where m is the number of uncontaminated pixels in the target or reference images.

Therefore, the preliminary reconstructed image y^{new} is obtained by the weighted average of all the reconstructed images from each reference image

$$y^{\mathrm{new}} = \sum_1^n w(s) \times y^s. \tag{14}$$

D. Error Correction per Class

After the previous steps, some errors may still exist on the preliminary reconstructed image compared with the target image. This might result from the differences between the initialized target image I^t and the actual target image. As stated in Section II-A, $I^t(\Omega_+)$ is obtained from the arithmetic average of normalized reference images over Ω_+. The normalization process could not completely eliminate the reflectance distinctions between images acquired at different times. Therefore, the preliminary reconstructed image might suffer from considerable errors over some regions. Moreover, it is generally believed that the residuals remain stable within the same class and vary among classes [47], since pixels belonging to the same land cover class with similar spectral characteristics react similarly during the reconstruction process. To further improve the reconstruction result by reducing the remaining differences, an error correction process for each class is introduced within Ω_+ in this study. Based on C classes obtained from k-means classification, mean residuals are calculated for each class within Ω

$$r_c = \frac{\sum\left(I_c^t(\Omega) - y_c^{\mathrm{new}}(\Omega)\right)}{n'} \tag{15}$$

where n' is the number of pixels in Ω belonging to the cth class. The reconstruction result after the error correction process X^{new} is obtained by removing the residuals per class from the preliminary reconstructed image y^{new}

$$X_c^{\mathrm{new}} = y_c^{\mathrm{new}} - r_c. \tag{16}$$

To keep pixel values over Ω the same as the target image in the final result, the final reconstructed image X^{final} is obtained by $X^{\mathrm{final}} = x^t(\Omega) \cup X^{\mathrm{new}}(\Omega_+)$.

III. Case Study to Validate the Proposed Method

A. Study Area and Test Data

In this study, Landsat 8 OLI surface reflectance data acquired from two study areas were used to evaluate the performances of our method. The first study area covers the majority of Beijing, northern Tianjin, and part of Hebei Province, China. This area has a warm temperate semihumid continental monsoon climate. Large and thick clouds often

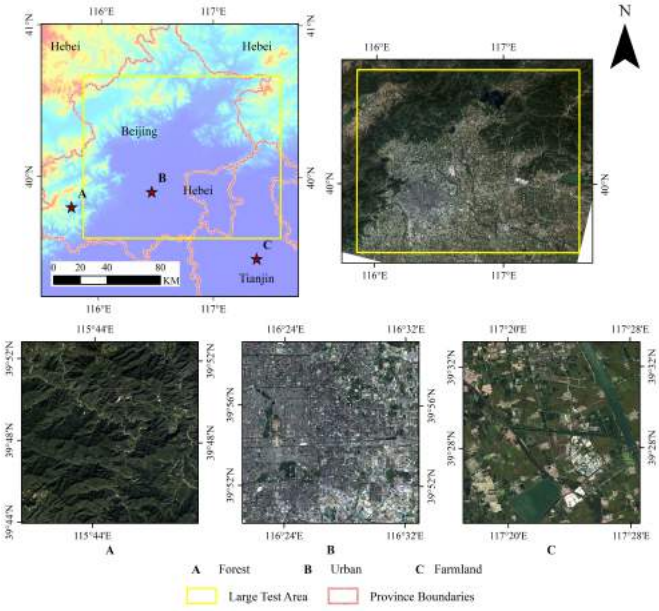

Fig. 4. Study area around Beijing and the distribution of subsets A, B, and C, and the large test area on a Landsat 8 OLI image.

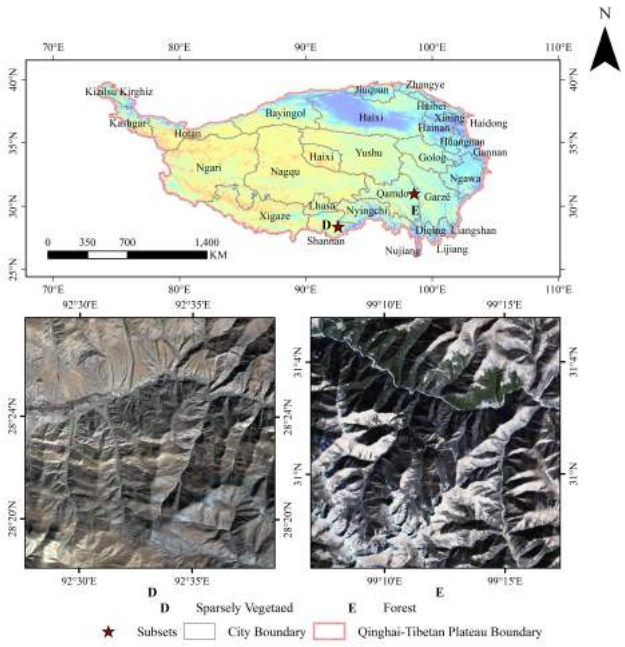

Fig. 5. Study area located in the QTP and the distributions of subsets D and E on Landsat 8 OLI images.

appear during vegetation growing seasons, causing difficulties for land surface observations using remote sensing techniques. The second study area selected is the Qinghai–Tibetan Plateau (QTP). The QTP, with altitudes ranging from 3000 to 8000 m, is dominated by plateau mountains and subtropical monsoon climates. The QTP is drawing increasing attention in environmental studies, since it is highly vulnerable to environmental changes [55]. These two test areas were selected because they vary greatly in climate types, natural conditions, and the intensity of human activities, leading to distinctive land cover types. Thus, the performance and stability of the proposed method can be tested over multiple land cover types under different environmental conditions.

Experiments were conducted separately on five small local subsets with sizes of 500 × 500 pixels and a large subset with a size of 3950 × 4925 pixels, and their distributions are shown in Figs. 4 and 5. Among the five local subsets, three were located around Beijing (subsets A, B, and C) and the other two (subsets D and E) were located on the QTP. Simulated cloud/cloud shadow masks were created manually. Subsets A, B, and C were extracted from the same OLI image (path: 123, row: 32) acquired on September 28, 2017. Two reference images were also applied, which were acquired on September 12 and October 30, 2017. Subset D is located in Shannan, extracted from an OLI image (path: 136, row: 40) acquired on February 11, 2017, and the reference images were acquired on January 26, 2017 and February 27, 2017. Subset E is located in Qamdo, extracted from an OLI image (path: 133, row: 38) acquired on January 21, 2017, and the two reference images were acquired on December 20, 2016 and February 6, 2017.

To demonstrate the reconstruction performance on different land cover types, subsets A, B, and C were covered by forest, urban area, and croplands, while subsets D and E were covered by sparsely vegetated land and evergreen forest in mountain areas. Subset A is a homogeneous mountain area located in the southwestern Beijing. This area is mainly covered with forest. The simulated cloud/cloud shadow [Fig. 6A(b)] accounted for 48.88% of the total area (122 196 missing pixels). Subset B is mostly covered by the urban area of Beijing city. Built-up areas are less likely to change over time compared to vegetated areas. The simulated mask [Fig. 6B(b)] represents missing information in up to 42.75% of the total area (106 875 missing pixels). Subset C shows a rural area located southeast of the study area, north of Tianjin. This area mainly consists of cropland with a small pond and river running through it and sparsely distributed farmhouses. The simulated mask [Fig. 6C(b)] accounted for 41.42% of the total subset (103 553 missing pixels). Subset D shows a mountain area with high altitude in Shannan city and its land surface is covered with sparse vegetation and bare ground. A narrow river runs across the subset. The simulated mask [Fig. 6D(b)] accounted for 52.72% of the total area (131 799 missing pixels). Subset E also shows a mountain area in Qamdo city, with evergreen forests covering mountain tops. The tree lines can be identified clearly. The simulated mask [Fig. 6E(b)] accounted for 52.72% of the total area (131 799 missing pixels).

A large subset was selected to test how our method performs on a large heterogeneous region covered by vast clouds/cloud shadows. This subset was extracted from an OLI image acquired on May 23, 2017, and acquisition dates for the reference images were March 4 and May 7, 2017. The target image was generated by overlaying the real cloud/cloud shadow mask on January 15, 2017, simulating a real cloud distribution. The simulated cloud/cloud shadow contaminated area accounted for approximately 68.63% of the total area. All the remote sensing data used in this study were obtained from EarthExplorer of the United States Geological Survey (USGS) (https://earthexplorer.usgs.gov/).

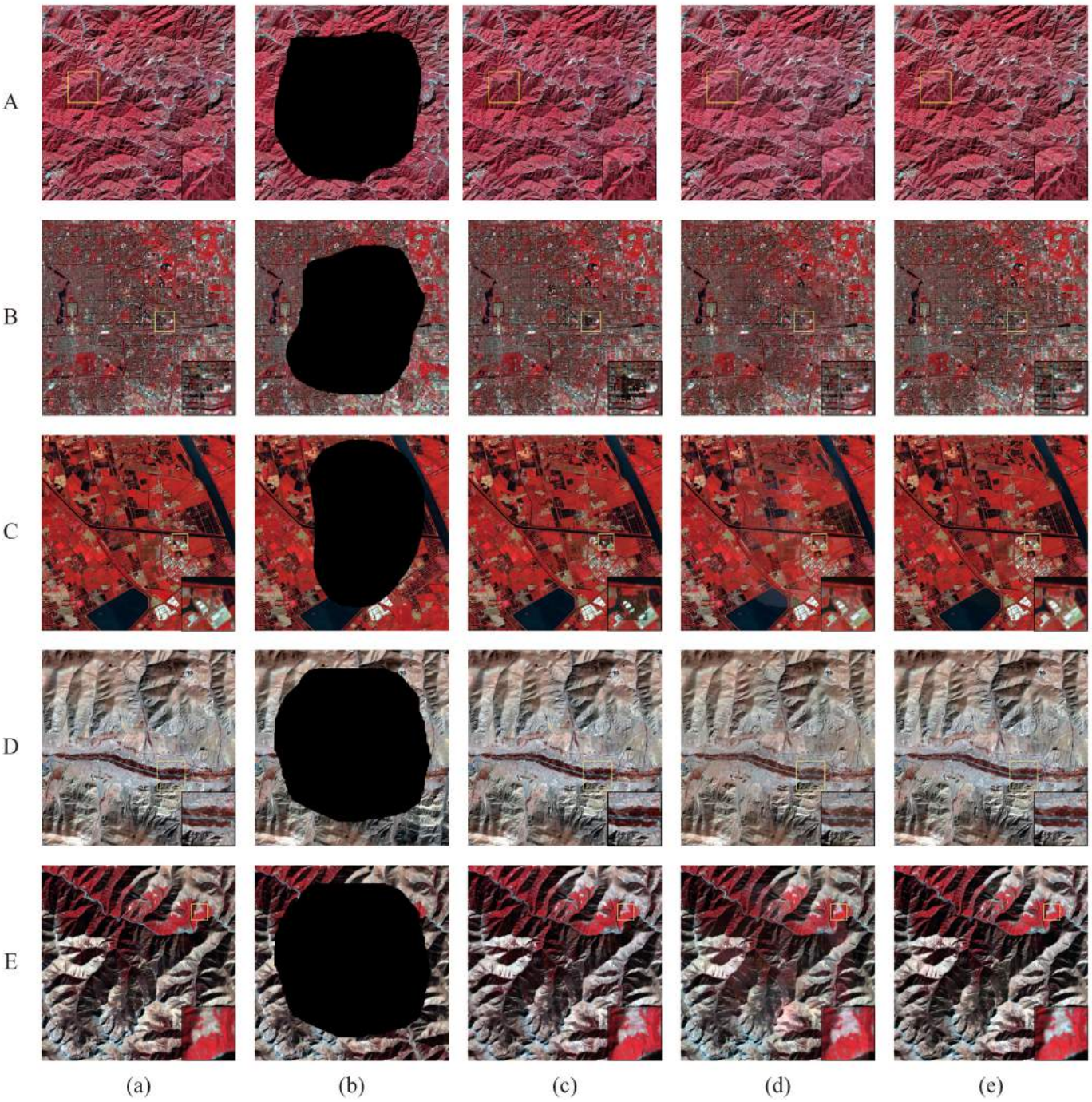

Fig. 6. Results of missing information reconstruction of the simulated missing areas over subsets A, B, C, D, and E (R = band 5, G = band 4, and B = band 3). (a) Actual images. (b) Simulated target images (actual image overlaid with simulated cloud mask). (c) Reconstruction results from GNSPI. (d) Reconstruction results from S-NMF-EC. (e) Reconstruction results from our method.

B. Results

The reconstructed results and the actual original data were compared, and corresponding density scatter plots and accuracy statistics over cloud/cloud shadow contaminated areas were presented to illustrate the performance of our method. To draw comparisons, the GNSPI and the S-NMF-EC methods were also applied over the same test data to reconstruct missing information because all three methods utilized spatio-temporal–spectral information. The GNSPI method is a generally used effective method based on geostatistical approaches, and S-NMF-EC is a state-of-the-art dictionary learning based method that generates the reference image by fusing data from multiple sensors. MODIS daily surface reflectance product (MOD09GA) was used for the S-NMF-EC method to conduct the fusion process. Two reference images were selected in the experiments.

The MAE, root-mean-square error (RMSE), peak signal-to-noise ratio (PSNR), and structure similarity (SSIM) indicators were calculated for bands 2, 3, 4, and 5 in terms of quantifying the accuracy of the reconstructed data. PSNR

TABLE I

QUANTITATIVE ANALYSIS OF RECONSTRUCTION RESULTS FOR SUBSETS A, B, C, D, AND E FROM GNSPI, S-NMF-EC, AND OUR METHOD FOR BANDS 2–5

Subsets	Bands	Methods	MAE ($\times 10^{-4}$)	PSNR	SSIM	R^2	RMSE ($\times 10^{-4}$)
A	Band 2	GNSPI	34.13	48.31	0.93	0.82	54.92
		S-NMF-EC	56.49	46.41	0.92	0.86	68.38
		Ours	**32.69**	**49.60**	**0.94**	**0.88**	**47.35**
	Band 3	GNSPI	36.33	47.95	0.96	0.92	57.26
		S-NMF-EC	53.73	46.59	0.96	0.92	66.95
		Ours	**32.87**	**49.57**	**0.97**	**0.95**	**47.54**
	Band 4	GNSPI	44.35	46.50	0.95	0.91	67.64
		S-NMF-EC	56.53	45.93	0.95	0.93	72.24
		Ours	**40.97**	**47.37**	**0.96**	**0.93**	**61.06**
	Band 5	GNSPI	169.16	35.36	0.96	0.91	243.93
		S-NMF-EC	169.17	36.17	**0.97**	0.94	222.42
		Ours	**139.26**	**37.87**	**0.97**	**0.95**	**182.68**
B	Band 2	GNSPI	72.72	43.16	**0.95**	0.87	106.32
		S-NMF-EC	92.38	41.63	**0.95**	0.85	126.65
		Ours	**60.89**	**44.86**	0.93	**0.91**	**87.36**
	Band 3	GNSPI	76.45	42.37	0.96	0.88	116.48
		S-NMF-EC	77.80	42.81	0.96	0.91	110.68
		Ours	**62.12**	**44.55**	**0.97**	**0.93**	**90.55**
	Band 4	GNSPI	89.62	40.97	0.96	0.88	136.86
		S-NMF-EC	94.50	40.85	0.96	0.91	138.70
		Ours	**74.54**	**42.95**	**0.97**	**0.93**	**108.84**
	Band 5	GNSPI	130.63	37.74	0.96	0.88	198.35
		S-NMF-EC	161.77	36.56	0.94	0.85	227.25
		Ours	**94.72**	**41.10**	**0.98**	**0.94**	**134.69**
C	Band 2	GNSPI	57.62	44.29	0.86	0.88	94.80
		S-NMF-EC	184.25	36.91	0.79	0.70	221.88
		Ours	**48.79**	**45.80**	**0.87**	**0.91**	**79.70**
	Band 3	GNSPI	67.24	43.29	0.87	0.87	106.36
		S-NMF-EC	116.09	39.72	0.81	0.72	160.38
		Ours	**64.36**	**44.15**	**0.88**	**0.89**	**96.32**
	Band 4	GNSPI	83.58	41.67	**0.88**	0.89	128.25
		S-NMF-EC	138.08	38.58	0.83	0.81	182.92
		Ours	**79.99**	**42.53**	**0.88**	**0.91**	**116.15**
	Band 5	GNSPI	228.91	33.05	0.89	0.86	345.98
		S-NMF-EC	271.77	32.68	0.89	0.87	361.06
		Ours	**173.06**	**35.94**	**0.91**	**0.92**	**247.84**
D	Band 2	GNSPI	35.43	44.07	0.90	0.93	62.57
		S-NMF-EC	66.56	35.98	**0.92**	0.56	218.78
		Ours	**34.66**	**44.62**	0.88	**0.94**	**58.75**
	Band 3	GNSPI	50.80	41.50	0.91	0.95	84.10
		S-NMF-EC	70.80	33.05	**0.94**	0.58	306.62
		Ours	**39.77**	**43.16**	0.91	**0.96**	**69.47**
	Band 4	GNSPI	61.34	39.95	0.91	0.95	100.57
		S-NMF-EC	85.58	31.38	**0.95**	0.57	371.62
		Ours	**43.80**	**42.02**	0.92	**0.97**	**79.24**
	Band 5	GNSPI	70.85	38.49	0.91	0.94	119.02
		S-NMF-EC	116.34	29.49	**0.94**	0.50	461.82
		Ours	**51.43**	**40.36**	0.92	**0.96**	**95.92**
E	Band 2	GNSPI	40.18	44.06	0.70	0.94	62.68
		S-NMF-EC	86.87	40.93	**0.83**	0.85	123.7
		Ours	**33.03**	**46.04**	0.75	**0.96**	**49.92**
	Band 3	GNSPI	48.58	42.03	0.76	0.95	79.18
		S-NMF-EC	82.36	41.34	**0.89**	0.93	118.06
		Ours	**39.84**	**44.24**	0.82	**0.97**	**61.41**
	Band 4	GNSPI	58.05	40.25	0.79	0.97	97.11
		S-NMF-EC	99.50	40.07	0.**89**	0.96	136.59
		Ours	**50.89**	**42.11**	0.82	**0.98**	**78.40**
	Band 5	GNSPI	117.60	35.28	0.81	0.96	172.28
		S-NMF-EC	257.82	32.08	**0.89**	0.92	342.81
		Ours	**92.82**	**37.55**	0.85	**0.98**	**132.52**

refers to the ratio between the maximum possible value of an image and the noise, which measures the quality of the reconstructed image. SSIM evaluates the similarity between the reconstructed and original images. Higher PSNR and SSIM and lower MAE and RMSE values indicate better reconstruction accuracies.

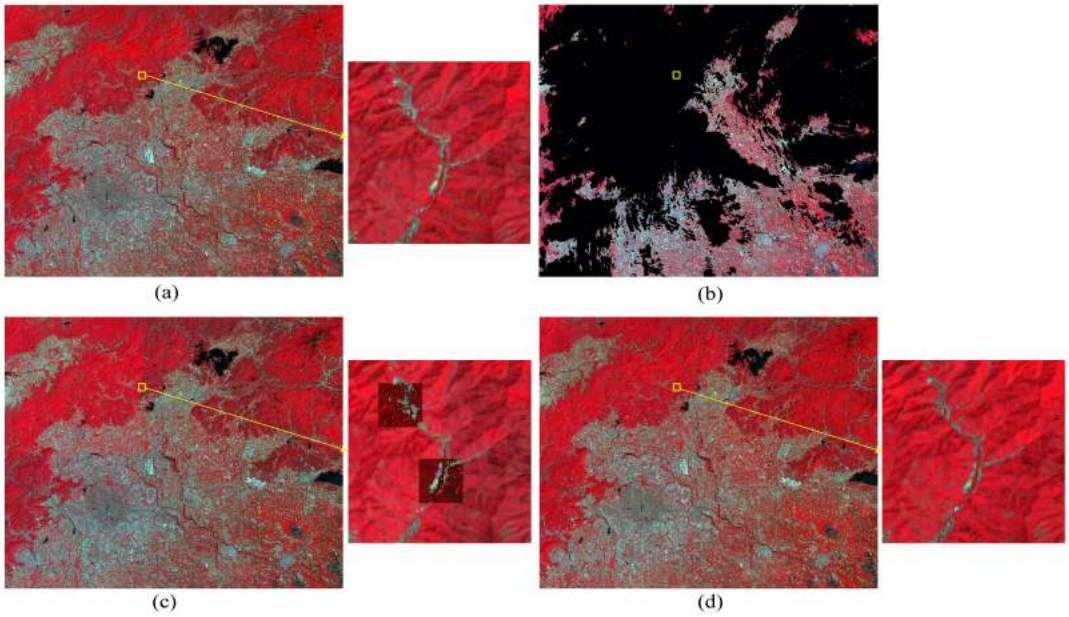

Fig. 7. Results of missing information reconstruction of the simulated missing area over the large subset (R = band 5, G = band 4, and B = band 3). (a) Actual image on May 23, 2017. (b) Simulated target image (actual image overlaid with simulated cloud mask). (c) Reconstruction result from GNSPI. (d) Reconstruction result from our method.

1) Reconstruction Results for Simulated Cloud Cover on Local Subsets: As discussed in Section II, the number of classes (C) and dictionary atoms (K) are two main parameters in our method. Optimal C and K values were chosen for the subsets to achieve the balance between reconstruction accuracy and computational cost for each subset. The influences of C and K on reconstruction accuracies will be discussed in Section IV-A. For subsets A, B, C, D, and E, the number of classes (C) are 10, 10, 30, 4, and 18, while the number of dictionary atoms (K) are 40, 40, 150, 100, and 100, respectively. Reconstruction results for all five subsets are shown in Fig. 6. Visually, all three methods produced satisfactory results without obvious errors. However, the GNSPI method produced unexpected dark or bright patches over some places, as is shown in the enlarged windows in Fig. 6(c). The results generated by the S-NMF-EC method seemed more blurring than the other two methods, and also brighter than the original images, especially in Fig. 6A(d), D(d), and E(d). As the S-NMF-EC method obtains reference data by fusing MODIS and Landsat data, the reconstruction accuracies rely largely on the quality and availability of the MODIS data. Our method reconstructed the contaminated images with better spatial consistency and clearer details. In terms of quantitative analysis (Table I), the reconstruction results from our method showed better performances compared with those of GNSPI and S-NMF-EC in each indicator from band 2 to 5. Overall, our method reconstructed cloud/cloud shadow contaminated regions with the best accuracy under distinctive environmental conditions with various land cover types.

2) Reconstruction Results for Simulated Cloud Cover Over a Large Area: Under common circumstances, large clouds/cloud shadows often appear on remote sensing data over large heterogeneous regions. Because the S-NMF-EC method produced lower accuracies than the other two methods for the five small subsets, we only conducted the comparison between GNSPI and our method for the large subset (Fig. 7). Clear edges can be spotted in the reconstruction results from GNSPI [Fig. 7(c)] since this method operates by a unit of a square block with a certain size (block size = 2000 for this test). Moreover, dark spots have appeared throughout the reconstructed area. The reconstructed image from the proposed method appeared to be more consistent spatially and less likely to generate underestimations. Moreover, Table II indicates that even though the reconstruction results from our method produced slightly higher MAE and RMSE in band 2, it has a much lower MAE and RMSE in the other bands and a higher R^2 over all bands. In addition, our method required approximately 1 h to complete the whole constructing process, while GNSPI took a much longer time of approximately 8 h.

TABLE II
QUANTITATIVE ANALYSIS LARGE AREA RECONSTRUCTION RESULTS FROM GNSPI AND OUR METHOD FOR BANDS 2–5

Bands	Method	MAE ($\times 10^{-4}$)	PSNR	SSIM	R^2	RMSE ($\times 10^{-4}$)
Band 2	GNSPI	**55.57**	**42.08**	**0.88**	0.91	**95.05**
	Ours	63.21	42.04	0.85	**0.92**	95.43
Band 3	GNSPI	68.59	40.89	0.88	0.89	108.95
	Ours	**63.21**	**42.11**	**0.89**	**0.92**	**94.65**
Band 4	GNSPI	87.16	38.77	0.89	0.91	139.14
	Ours	**84.57**	**39.41**	**0.87**	**0.92**	**129.21**
Band 5	GNSPI	271.99	30.07	0.84	0.80	378.51
	Ours	**222.96**	**32.06**	**0.86**	**0.88**	**301.16**

IV. DISCUSSION

This study demonstrated the capability of the proposed method in reconstructing missing information contaminated by

large and thick clouds/cloud shadows. The proposed method efficiently achieved satisfactory reconstruction results on small subsets covered with various land cover types under distinctive environmental conditions, as well as over large images with heterogeneous land cover. In this section, some factors related to the proposed method will be discussed, including the number of dictionary atoms K, the number of classes C, the number of reference images, the improvement of the proposed method in eliminating reconstruction edges, and the limitations of reference images.

A. Number of Dictionary Atoms K and Number of Classes C

The reconstruction results from different values of dictionary atoms (K) and land cover classes (C) were evaluated in the first study area to discuss how K or C influences the reconstruction accuracy. Fig. 8 shows the MAE values under different values of K and C. The black dotted lines in the graphs denote the MAE of the reconstruction result from the GNSPI method for each subset. The blue, green, and cyan lines show the changes in MAE values with C increasing from 4 to 30 when K is 40, 100, and 150, respectively. In Fig. 8(a), most MAE values vary within the range of 60–65 ($\times 10^{-4}$). An overall tendency of decreasing MAE with increasing C can be found. Also, with increasing K, the reconstruction accuracy and stability can be slightly improved. For subset B, good reconstruction accuracy can be achieved with small values for K and C. For subset C, where the spectral characteristics have more complexity compared with the other two subsets, greater values for K and C are efficient in reducing the reconstruction errors. Since land cover of subset C is more heterogeneous, greater values of C and K are needed to capture the spectral characteristics. For all three subsets, only a few sets of parameters (low values of C and K for subset C) produce higher MAE values than that of GNSPI, which provides more evidence that the proposed method outperforms GNSPI. Generally, with increasing K or C, there is a greater chance to achieve better reconstruction accuracy. However, greater K or C also increases computation complexity, which means more time is needed to fulfill the reconstruction process. Thus, the balance between reconstruction accuracy and computational complexity needs to be achieved. Moreover, unsupervised classification is also an important influencing factor for the final reconstruction accuracy. Better classification results can lead to higher reconstruction accuracy and improve the stability of our method.

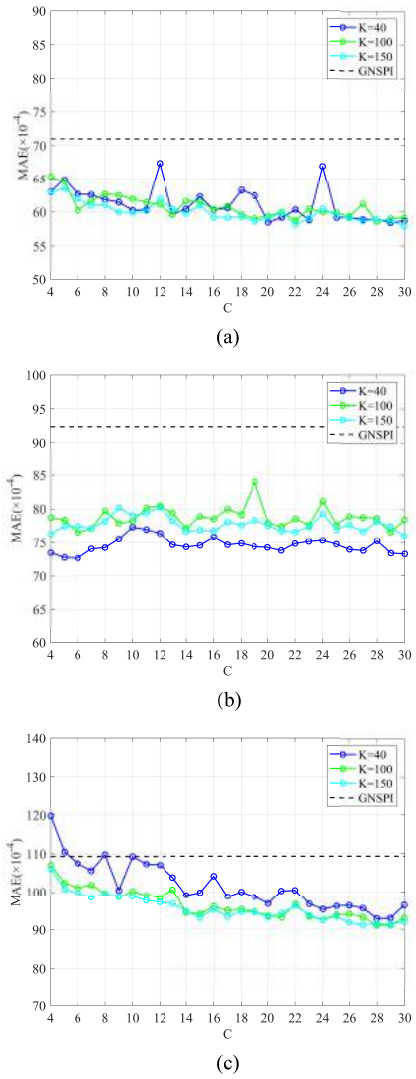

Fig. 8. MAEs of reconstruction results for different C and K values over (a) subset A, (b) subset B, and (c) subset C.

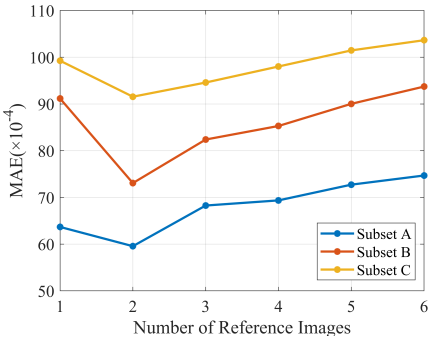

Fig. 9. Relationship between the number of reference images used and the MAE of the reconstruction results for subsets A, B, and C.

B. Number of Reference Images

Our method is designed to utilize multiple reference images. Therefore, based on the previous three subsets in the first study area on September 28, 2017, a total of six images were selected as reference images to test how the number of reference images would affect reconstruction accuracies (Fig. 9). The six reference images were acquired on September 12, October 30, November 15, July 10, May 23, and May 7, 2017, from the closest date to the farthest. It can be inferred from Fig. 9 that for all three subsets, the application of two reference images could achieve the best reconstruction accuracy. Since more reference images introduce longer time intervals between the acquisition dates among the images, potential land cover change might occur. In addition, since the missing information is recovered using the compressed information from reference images, the reconstruction accuracies mainly rely on the similarity between the reference and target images. Therefore, the accuracy of the reconstruction results does not

Fig. 10. Results of missing information reconstruction of the simulated missing areas over the LGC subset (R = band 4, G = band 3, and B = band 2). (a) Actual image. (b) Simulated target image (actual image overlaid with simulated cloud mask). (c) Reconstruction results from WLR. (d) Reconstruction results from GNSPI. (e) Reconstruction results from S-NMF-EC. (f) Reconstruction results from our method.

necessarily increase with the increasing number of reference images. When choosing reference images, the time interval between the reference and target images should also be taken into consideration.

C. Comparisons With Other Methods

According to Section III, our method achieved better performances than the GNSPI and S-NMF-EC methods in the abovementioned cases. To reinforce the statement, we also carried out comparisons on datasets used for the development of S-NMF-EC method [49]. The data were produced by Emelyanova *et al.* [56], which provided matching pairs of Landsat and MODIS imageries for comparing the performances of fusion methods. The data over identical subset with Li *et al.* [49] at the size of 300 × 300 pixels and three spectral bands (R = band 2, G = band 3, and B = band 4 for the Landsat TM images) were generated in the study site of Lower Gwydir Catchment (LGC), located in northern New South Wales, Australia. The target and reference images were obtained from Landsat TM images on October 25, 2004 and August 22, 2004, respectively. A MODIS image on August 22, 2004, was adopted for the S-NMF-EC method. The contaminated area was manually created and accounted for 41.28% of the total image (37 149 pixels). The original datasets of Landsat–MODIS pairs can be downloaded from https://data.csiro.au/collections/collection/CIcsiro:5847v003. Apart from the above three methods, we also carried out the weighted linear regression (WLR) [25], a popular multitemporal missing data reconstruction method, and the reconstruction results from the four methods were compared.

Fig. 10 and Table III show the reconstruction results and their accuracies for WLR, GNSPI, S-NMF-EC, and our method. Among the four methods, our method achieved the best reconstruction accuracy. WLR, GNSPI, and S-NMF-EC methods produced red patches marked by the yellow circles shown in Fig. 10. Our method reconstructed the contaminated regions with better spatial consistency and clearer details. The overall reconstruction accuracies of our method also outperformed the other methods. The reconstruction accuracies may still be improved if multiple reference images are adopted.

These experiments indicated that our method was stable under different environmental conditions and land cover types. Both WLR and GNSPI methods were developed based on predicting missing information using selected similar pixels with the reference of multitemporal images. WLR performs well in reconstructing missing remote sensing data under many circumstances [11], [25], [49]. WLR determines the similar pixels and their corresponding weights based on the spectral and spatial differences to the target pixel. However, when the contaminated region is large, the number of similar pixels may be limited, which will reduce its reconstruction accuracy. The GNSPI method was also capable of recovering cloud/cloud shadow contaminated areas with good accuracy, especially when dealing with simple ground features [11]. However, GNSPI may generate unexpected dark or bright patches, i.e., GNSPI may over or underestimate a group of pixels in some cases [Fig. 6(c)]. The GNSPI estimates a target pixel value using its similar pixels which are selected based on several conditions. First, they belong to the same land cover class as the target pixel. Second, the RMSE between the target and each similar pixel is lesser than the "similar threshold." Third, they are spatially adjacent to the target pixel. If the three requirements cannot be met at the same time, the algorithm selects the adjacent pixels belonging to the same class with the target pixel as similar pixels. For the pixels belonging to the abovementioned unexpected patches, there is no available pixel that satisfies all three requirements. They only have one or two adjacent pixels with the same land cover class. Therefore, there is a higher chance of producing abnormal values from the limited number of similar pixels. However, our method reconstructs missing information based on extracted dictionary atoms and sparse matrices of reference images, which depends more on spectral similarity between pixels

TABLE III
QUANTITATIVE ANALYSIS LARGE AREA RECONSTRUCTION RESULTS FROM WLR, GNSPI, S-NMF-EC, AND OUR METHOD FOR BAND 2, BAND 3, AND BAND 4

Bands	Method	MAE ($\times 10^{-4}$)	PSNR	SSIM	R^2	RMSE ($\times 10^{-4}$)
Band 2	WLR	70.58	43.22	0.70	0.34	107.41
	GNSPI	51.91	46.33	0.76	0.53	75.06
	S-NMF-EC	65.87	43.84	**0.83**	0.59	100.04
	Ours	**51.08**	**47.27**	0.79	**0.66**	**67.38**
Band 3	WLR	97.09	40.25	0.74	0.52	151.24
	GNSPI	76.65	42.29	0.79	0.62	119.58
	S-NMF-EC	156.11	37.91	**0.84**	0.58	197.96
	Ours	**65.34**	**44.93**	0.83	**0.80**	**88.21**
Band 4	WLR	117.50	38.50	0.80	0.67	184.93
	GNSPI	92.69	41.08	0.86	0.76	137.53
	S-NMF-EC	167.17	37.63	**0.85**	**0.82**	204.41
	Ours	**89.56**	**42.42**	**0.85**	**0.82**	**117.85**

and their relationships, rather than spatial adjacency. This may explain why our method achieved better performances and the GNSPI produced the unexpected patches.

The S-NMF-EC method is a newly developed method based on dictionary learning. The main differences between S-NMF-EC and our method are as follows. First, our method does not require auxiliary data from other sensors. However, the S-NMF-EC method requires both MODIS and Landsat data to produce the reference image using the spatial and temporal nonlocal filter-based fusion model (STNLFFM). Thus, the quality of the reference image will largely rely on the availability and quality of MODIS data. Also, the employment of MODIS data will introduce uncertainties resulted from the scale effects and time intervals between MODIS and Landsat data. Second, in the previous dictionary learning-based methods, the data over the uncontaminated region in the result are changed in the algorithms [49]. Both S-NMF-EC and our method employed an error correction step to deal with this problem. However, the error correction steps in these two methods are different. S-NMF-EC calculates errors between the target and reconstructed data over uncontaminated regions and then applies the least square fitting to establish relationships between the reconstructed data and the errors, providing estimations of the errors for each pixel over the entire image. The estimated errors are then added to the corresponding reconstructed data to obtain the final reconstruction result. Our method, as previously discussed, determines the error for each class and carries out correction only over the uncontaminated regions. Therefore, compared to S-NMF-EC, our method has significantly improved the reconstruction accuracies over cloud contaminated regions.

D. Limitations of the Reference Images

In reality, finding a total cloud/cloud shadow-free reference image is a strict limitation. As stated by Xu *et al.* [46], dictionary atoms are fundamental components of an image, which reveal the representative spectral characteristics of an image. Since our method reconstructs missing data based on extracted dictionary atoms from multitemporal reference images, the reconstruction results depend largely on the quality of the extracted dictionary atoms. Therefore, when reference images are partly covered with clouds, dictionary atoms extracted based on uncontaminated pixels from multitemporal reference images can be integrated. The condition should be met so that the clouds/cloud shadows in each multitemporal reference image do not cover the same area. In other words, each pixel should have at least one available value among all the multitemporal reference images to ensure that the relationship between a pixel and the other pixels can be established based on complete dictionary atoms and sparse matrix. Under this condition, the dictionary atoms extracted from cloud-free parts of all the reference images are capable of compensating for each other in providing a complete estimation of area Ω_+ in the target image. Thus, the requirement of reference images became less strict and more practical in reconstructing missing data in reality.

As it is believed that remote sensing data acquired at closer times would have greater similarity, reference images with acquisition dates closer to that of the target image can be selected. The reconstruction results could be improved by useful information acquired at a closer time. Therefore, our method does not always require a certain number of cloud-free reference images to fulfill the reconstruction process as long as clouds/cloud shadows cover different areas in different reference images, and can make the most use of available information with acquisition dates closer to that of the target image to improve the reconstruction results.

V. Conclusion

In this article, an improved method based on MDL to reconstruct missing information of remote sensing data contaminated by clouds/cloud shadows was proposed. The method is capable of accurately reconstructing remote sensing data contaminated by large and thick clouds while greatly improving computation efficiency. Moreover, by applying the weighted average to the reconstruction results from multitemporal reference images and error correction to the preliminary reconstructed results, our method can achieve satisfactory reconstruction accuracy without the application of multisensor remote sensing data. In addition, data within the cloud/cloud shadow-free areas remain unchanged, and no obvious reconstruction boundaries can be spotted. Although only Landsat 8 OLI data were used in the case study, data acquired by other sensors are also applicable. Our method has great potential in reconstructing time-series surface reflectance data at high spatial resolutions, and has a great potential for further generating long-sequence spatially complete land cover classification maps and land surface parameters including fractional vegetation cover (FVC), leaf area index (LAI), etc. In addition, by integrating proper fusion methods, the temporal resolution of the spatially complete land surface parameters can also be improved. This can increase the capability of the current land surface parameters in capturing rapid land surface changes, especially for crops. Moreover, the employment of multisource data for missing data reconstruction and the production of land surface parameters over large regions will also be considered for further studies when consecutive time series of data are unavailable from a single sensor.

References

[1] D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," *Biosyst. Eng.*, vol. 114, no. 4, pp. 358–371, Apr. 2013.

[2] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger, "Remote sensing for biodiversity science and conservation," *Trends Ecol. Evol.*, vol. 18, no. 6, pp. 306–314, Jun. 2003.

[3] T. J. Schmugge, W. P. Kustas, J. C. Ritchie, T. J. Jackson, and A. Rango, "Remote sensing in hydrology," *Adv. Water Resour.*, vol. 25, nos. 8–12, pp. 1367–1385, Aug.-Dec. 2002.

[4] F. Van der Meer, "Remote-sensing image analysis and geostatistics," *Int. J. Remote Sens.*, vol. 33, no. 18, pp. 5644–5676, Mar. 2012.

[5] E. Chuvieco and R. G. Congalton, "Application of remote sensing and geographic information systems to forest fire hazard mapping," *Remote Sens. Environ.*, vol. 29, no. 2, pp. 147–159, Aug. 1989.

[6] C. Huggel, A. Kääb, W. Haeberli, P. Teysseire, and F. Paul, "Remote sensing based assessment of hazards from glacier lake outbursts: A case study in the Swiss Alps," *Can. Geotech. J.*, vol. 39, no. 2, pp. 316–330, Apr. 2002.

[7] R. Nagarajan, A. Mukherjee, A. Roy, and M. V. Khire, "Technical note temporal remote sensing data and GIS application in landslide hazard zonation of part of Western Ghat, India," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 573–585, Mar. 1998.

[8] S. Sahoo, A. Dhar, and A. Kar, "Environmental vulnerability assessment using grey analytic hierarchy Process based model," *Environ. Impact Assessment Rev.*, vol. 56, pp. 145–154, Jan. 2016.

[9] M. D. King, Y. J. Kaufman, W. P. Menzel, and D. Tanre, "Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS)," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 1, pp. 2–27, Jan. 1992.

[10] G. W. Paltridge and C. M. R. Platt, *Radiative Processes in Meteorology and Climatology*. New York, NY, USA: Elsevier, 1976.

[11] H. Shen *et al.*, "Missing information reconstruction of remote sensing data: A technical review," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 61–85, Sep. 2015.

[12] N. Cressie and G. Johannesson, "Fixed rank kriging for very large spatial data sets," *J. Roy. Stat. Soc., B (Stat. Methodol.)*, vol. 70, no. 1, pp. 209–226, Jan. 2008.

[13] R. E. Rossi, J. L. Dungan, and L. R. Beck, "Kriging in the shadows: Geostatistical interpolation for remote sensing," *Remote Sens. Environ.*, vol. 49, no. 1, pp. 32–40, Jul. 1994.

[14] C. A. Z. Barcelos and M. A. Batista, "Image inpainting and denoising by nonlinear partial differential equations," in *Proc. 16th Brazilian Symp. Comput. Graph. Image Process. (SIBGRAPI)*, Marbella, Spain, 2001, pp. 261–266.

[15] L. Lorenzi, F. Melgani, and G. Mercier, "Missing-area reconstruction in multispectral images under a compressive sensing perspective," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 3998–4008, Jul. 2013.

[16] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 175–187, Jan. 2014.

[17] W.-H. Cheng, C.-W. Hsieh, S.-K. Lin, C.-W. Wang, and J.-L. Wu, "Robust algorithm for exemplar-based image inpainting," in *Proc. Int. Conf. Comput. Graph., Imag. Vis. (CGIV)*, 2005, pp. 64–69.

[18] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.

[19] L. Wang, J. J. Qu, X. Xiong, X. Hao, Y. Xie, and N. Che, "A new method for retrieving band 6 of Aqua MODIS," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 2, pp. 267–270, Apr. 2006.

[20] Q. Liu, X. Gao, L. He, and W. Lu, "Haze removal for a single visible remote sensing image," *Signal Process.*, vol. 137, pp. 33–43, Aug. 2017.

[21] M. Xu, X. Jia, and M. Pickering, "Automatic cloud removal for landsat 8 OLI images using cirrus band," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 2511–2514.

[22] M. Xu, X. Jia, M. Pickering, and S. Jia, "Thin cloud removal from optical remote sensing images using the noise-adjusted principal components transform," *ISPRS J. Photogramm. Remote Sens.*, vol. 149, pp. 215–225, Mar. 2019.

[23] C.-H. Lin, P.-H. Tsai, K.-H. Lai, and J.-Y. Chen, "Cloud removal from multitemporal satellite images using information cloning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 232–241, Jan. 2013.

[24] J. Storey, P. Scaramuzza, and J. Barsi, "Landsat 7 scan line corrector-off gap-filled product development," in *Proc. Pecora Conf.*, Sioux Falls, SD, USA, 2005, pp. 23–27.

[25] C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sens. Environ.*, vol. 131, pp. 182–194, Apr. 2013.

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

[27] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1824–1832, Aug. 2002.

[28] P. S. A. Beck, C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI," *Remote Sens. Environ.*, vol. 100, no. 3, pp. 321–334, Feb. 2006.

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

[30] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter," *Remote Sens. Environ.*, vol. 91, nos. 3–4, pp. 332–344, Jun. 2004.

[31] L. Eklundh and P.-G. Jonsson, "Extracting information about vegetation seasons in Africa from Pathfinder AVHRR NDVI imagery using temporal filtering and least-squares fits to asymmetric Gaussian functions," *Proc. SPIE*, vol. 4885, pp. 215–225, Mar. 2003.

[32] G. J. Roerink, M. Menenti, and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *Int. J. Remote Sens.*, vol. 21, no. 9, pp. 1911–1917, Jan. 2000.

[33] X. Zhang *et al.*, "Monitoring vegetation phenology using MODIS," *Remote Sens. Environ.*, vol. 84, no. 3, pp. 471–475, Mar. 2003.

[34] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin, "A simple and effective method for filling gaps in landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1053–1064, Apr. 2011.

[35] X. Zhu, D. Liu, and J. Chen, "A new geostatistical approach for filling gaps in landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 124, pp. 49–60, Sep. 2012.

[36] X. Zhu, F. Gao, D. Liu, and J. Chen, "A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 3, pp. 521–525, May 2012.

[37] Q. Cheng, H. Shen, L. Zhang, Q. Yuan, and C. Zeng, "Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model," *ISPRS J. Photogramm. Remote Sens.*, vol. 92, pp. 54–68, Jun. 2014.

[38] L. Yan and D. Roy, "Large-area gap filling of landsat reflectance time series by spectral-angle-mapper based spatio-temporal similarity (SAMSTS)," *Remote Sens.*, vol. 10, no. 4, p. 609, Apr. 2018.

[39] F. Melgani, "Contextual reconstruction of cloud-contaminated multitemporal multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 442–455, Feb. 2006.

[40] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4274–4288, Aug. 2018.

[41] Y. Huang *et al.*, "Improving MODIS snow products with a HMRF-based spatio-temporal modeling technique in the upper rio grande basin," *Remote Sens. Environ.*, vol. 204, pp. 568–582, Jan. 2018.

[42] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[43] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[44] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[45] T. Yao *et al.*, "Different glacier status with atmospheric circulations in tibetan plateau and surroundings," *Nature Climate Change*, vol. 2, no. 9, pp. 663–667, Jul. 2012.

[46] M. Xu, X. Jia, M. Pickering, and A. J. Plaza, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2998–3006, May 2016.

[47] X. Li, H. Shen, H. Li, and L. Zhang, "Patch matching-based multitemporal group sparse representation for the missing information reconstruction of remote-sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 3629–3641, Aug. 2016.

[48] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.

[49] X. Li, L. Wang, Q. Cheng, P. Wu, W. Gan, and L. Fang, "Cloud removal in remote sensing images using nonnegative matrix factorization and error correction," *ISPRS J. Photogramm. Remote Sens.*, vol. 148, pp. 103–113, Feb. 2019.

[50] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.

[51] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, pp. 19–60, Mar. 2010.

[52] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 689–696.

[53] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[54] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.
[55] D. Li, S. Wu, L. Liu, Y. Zhang, and S. Li, "Vulnerability of the global terrestrial ecosystems to climate change," *Global Change Biol.*, vol. 24, no. 9, pp. 4095–4106, Sep. 2018.
[56] I. V. Emelyanova, T. R. McVicar, T. G. Van Niel, L. T. Li, and A. I. Van Dijk, "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.*, vol. 133, pp. 193–209, Jun. 2013.

Mu Xia received the B.S. degree in geographic information system from the China University of Geosciences, Beijing, China, in 2013, and the M.A. degree in environmental remote sensing and geographic information system from Boston University, Boston, MA, USA, in 2015. She is pursuing the Ph.D. degree in cartography and geography information system with the State Key Laboratory of Remote Sensing Science and the Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing.

Her main research interests are focused on environmental remote sensing and ecological evaluation using remote sensing data.

Kun Jia received the B.S. degree in surveying and mapping engineering from Central South University, Changsha, China, in 2006, and the Ph.D. degree in cartography and GIS from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2011.

He is an Associate Professor with the State Key Laboratory of Remote Sensing Science and also the Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing. His main research interests include estimation of vegetation parameters, land cover classification, and agriculture monitoring using remote sensing data.