https://dx.doi.org/10.1016/j.rse.2022.113136

Journal: Remote Sensing of Environment

Title: A new object and class based gap-filling method for PlanetScope satellite image time series

Author List: Jing Wang[1], Calvin KF Lee[1], Xiaolin Zhu[2], Ruyin Cao[3], Yating Gu[1], Shengbiao Wu[1], Jin Wu[1,4,5*]

Author Affiliations:

(1) Research Areas of Ecology and Biodiversity, School for Biological Sciences, The University of Hong Kong, Hong Kong, China

(2) Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

(3) School of Resources and Environment, University of Electronic Science and Technology of China, Sichuan, China

(4) Institute for Climate and Carbon Neutrality, The University of Hong Kong, Hong Kong, China

(5) State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong, China

*** Corresponding Author:** Jin Wu

School for Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong (email: jinwu@hku.hk; phone: +852 2299-0655)

Figures/Tables Record:

Number of figures/tables: 10 figures, 4 tables

Supporting materials: 13 figures, 1 table

Abstract

PlanetScope CubeSats data with a 3-meter resolution, frequent revisits, and global coverage have provided an unprecedented opportunity to advance land surface monitoring over the recent years. Similar to other optical satellites, cloud-induced data missing in PlanetScope satellites substantially hinders its use for broad applications. However, effective gap-filling in PlanetScope image time series remains challenging and is subject to whether it can 1) consistently generate high accuracy results regardless of different gap sizes, especially for heterogeneous landscapes, and 2) effectively recover the missing pixels associated with rapid land cover changes. To address these challenges, we proposed an object and class based gap-filling ('OCBGF') method. Two major novelties of OCBGF include 1) adopting an **object**-based segmentation method in conjunction with an unsupervised **class**ification method to help characterize the landscape heterogeneity and facilitate the search of neighboring valid pixels for gap-filling, improving its applicability regardless of the gap size; 2) employing a scenario-specific gap-filling approach that enables effective gap-filling of areas with rapid land cover change. We tested OCBGF at four sites representative of different land cover types (plantation, cropland, urban, and forest). For each site, we evaluated the performance of OCBGF on both simulated and real cloud-contaminated scenarios, and compared our results with three state-of-the-art methods, namely Neighborhood Similar Pixel Interpolator (NSPI), AutoRegression to Remove Clouds (ARRC), and Spectral-Angle-Mapper Based Spatio-Temporal Similarity (SAMSTS). Our results show that across all four sites, OCBGF consistently obtains the highest accuracy in gap-filling when applied to scenarios with various gap sizes (RMSE=0.0065, 0.0090, 0.0092, and 0.0113 for OCBGF, SAMSTS, ARRC, and NSPI, respectively) and with/without rapid land cover changes (RMSE=0.0082, 0.0112, 0.0119, and 0.0120 for OCBGF, SAMSTS, ARRC, and NSPI, respectively). These results demonstrate the effectiveness of OCBGF for gap-filling PlanetScope image time series, with potential to be extended to other satellites.

Keywords: Gap-filling, CubeSats, image reconstruction, cloud removal, object-based segmentation

1. Introduction

Recent advances in Earth observation satellites with increasing spatial and temporal resolutions have created unprecedented opportunities for monitoring rapid and fine-scale changes on the Earth's surface. One typical example of these advances is the PlanetScope constellation that is made of 180+ micro satellites (CubeSats) (Planet, 2021). These CubeSats altogether provide a daily-to-weekly global coverage at a 3-meter spatial resolution with four spectral bands (i.e. red, green, blue, and near-infrared (NIR)) (Roy et al., 2021; Wang et al., 2020). As a result of these specifications, PlanetScope satellites have been increasingly suggested as a powerful and new means to improve fine-scale Earth's surface monitoring. Similar to other optical satellite observations, PlanetScope data are subject to cloud/cloud shadow contamination. According to the global statistics (Ju and Roy, 2008; Norris et al., 2016), around one-third of the global land surface is covered by clouds within a year, resulting in large quantities of missing data that hinder the wider applications of optical satellite measurements including PlanetScope (Roy et al., 2021). Therefore, accurate and effective gap-filling is important to advance the use of PlanetScope data for monitoring those rapid changes in land surface statuses and processes, such as fine-scale land use development, agricultural expansion, natural disasters and associated impact assessments (Feng et al., 2022; Halls and Magolan, 2019; Wang et al., 2019; Zeng et al., 2018, 2021), forecasting seasonal crop growth (Kimm et al., 2020; Sadeh et al., 2021), quantifying fine-scale plant phenology (Chen et al., 2019; Wang et al., 2020; Wu et al., 2021), and characterizing surface carbon and water fluxes (Dechant et al., 2022; Kong et al., 2022; McCabe et al., 2017).

To gap-fill missing data due to cloud/cloud shadow contamination in satellite images, many methods have been developed. These methods can be grouped into four categories according to the auxiliary information used, including fusion-, spatial-, temporal-, and spatiotemporal- based methods (Cao et al., 2020; Shen et al., 2015). Fusion-based methods have been used for gap-filling based on the integration of multisource images from other optical satellites (Luo et al., 2018; Roy et al., 2008) or synthetic aperture radar (SAR) images (Huang et al., 2015; Li et al., 2020b). However, these methods suffer from inconsistencies in spatial resolution and radiometric characteristics among different data types (Cao et al., 2020). Spatial-based methods rely on the assumption that adjacent pixels tend to be more similar due to spatial autocorrelation, and often use the neighboring clear pixels to fill gaps via various approaches such as spatial interpolation methods (Pringle et al., 2009; Zhang et al., 2007) and inpainting methods (Lorenzi et al., 2011; Maalouf et al., 2009). These methods have been demonstrated to be effective for filling data gaps associated with small cloud regions or gaps due to instrument errors such as Landsat Scan Line Corrector (SLC)-off images, but uncertainty massively increases with the size of data gaps (Shen et al., 2015). Temporal-based methods rely on the assumption that there are no obvious land cover changes over a short period of time, and missing pixels in the cloudy image ('target image') can be substituted or modeled with the pixels from cloud-free images acquired on adjacent dates ('reference images') after a relative normalization process. These methods include temporal filtering, temporal replacement, and machine learning based methods (Li et al., 2020a; Lin et al., 2013; Yan and Roy, 2020; Zeng et al., 2013). Nevertheless, these methods are sensitive to the selection of the reference images and generate large uncertainty in heterogeneous landscapes as they usually assume that spectrally similar pixels come from the same class and have the same temporal changing patterns without accounting for the intra-class temporal variability (Chen et al., 2011; Shen et al., 2015). The fourth type is spatiotemporal-based methods that integrate both spatial and temporal information for gap-filling (Zhang et al., 2018; Zhu et al., 2012a). For example, a neighborhood similar pixel interpolator (NSPI) method that combines both temporal predictions from the cloud-free reference image and spatial predictions from pixels outside the data gaps of the target image has been proposed to gap-fill the SLC-off data missing in individual Landsat images (Zhu et al., 2012a), which was further modified to gap-fill the cloud/cloud shadow-induced data missing in a time series of Landsat images (Zhu et al., 2018). However, it is criticized for the sensitivity to reference images used (Cao et al., 2020). To reduce the dependence on specific reference images, Cao et al. (2020) developed an autoregression method to remove clouds (ARRC) for Landsat data that explicitly uses the autoregression of Landsat image time series of the adjacent clear pixels to gap-fill those missing pixels.

Among the four categories of gap-filling methods described above, spatiotemporal-based methods are more commonly used since they generally obtain satisfactory results (Cao et al., 2020; Shen et al., 2015). For example, recently developed NSPI and ARRC have been tested across multiple sites across the globe with demonstrated high performance in gap-filling (Cao et al., 2020; Zhu et al., 2018). However, most of these methods are tested on medium to coarse spatial resolution satellite data (e.g. Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat, and

Sentinel-2) (Griffiths et al., 2019; Liu et al., 2017; Shen et al., 2015), and have rarely been explored in high spatial resolution satellite data, such as PlanetScope with a 3-m spatial resolution. Therefore, their applicability to the high-resolution PlanetScope remains unknown and may be constrained by the following two factors.

The first one is associated with the gap size concern. For the spatial component in spatiotemporal-based methods, the accuracy and efficiency have been demonstrated to decrease significantly with increasing data gap size (or the number of pixels with missing values). Meanwhile, the temporal-based component in the spatiotemporal-based methods relies on the spectrally similar pixels in the neighborhood of data missing pixels to model the temporal changing pattern between the target and reference images, and the search for the spectrally similar pixels for each data missing pixel is time-consuming, especially for the large data gaps. Moreover, due to a much higher spatial resolution of PlanetScope images than traditional satellite data, two novel issues could also emerge. One is even for the same cloud/shadow size (in terms of m^2), there would be more pixels with data missing in PlanetScope images than others. Second, due to the spatial heterogeneity of land cover in nature, there could be multiple objects within the same class or multiple classes within the same object in high-resolution PlanetScope images (e.g. Fig. S1). In other words, concerns about the spatial heterogeneity issue can be more serious in PlanetScope images than in traditional satellite data, especially for those big-size data gaps. Thus, methodological improvements are needed to address the above issues that are particularly important to high-resolution PlanetScope data. Object-based segmentation methods have been increasingly used to automatically segment spectrally and spatially similar pixels into independent objects in satellite images (Hossain and Chen, 2019; Myint et al., 2011), and have been recently implemented in gap-filling (Case and Vitti, 2021; Maxwell et al., 2007; Wu et al., 2018; Yan and Roy, 2018) and up to spatiotemporal fusion methods (Guan et al., 2017; Huang and Zhang, 2014; Luo et al., 2018; Xi et al., 2019). For example, a Spectral-Angle-Mapper Based Spatio-Temporal Similarity (SAMSTS) method (Yan and Roy, 2018) that segmented image time series into a segmentation map has been demonstrated effective for gap-filling large gaps in Landsat image time series. Such object-based segmentation methods may be able to address the above concern by reducing the sensitivity on gap size while improving the efficiency for filling large gaps in very high-resolution satellite images (Case and Vitti, 2021), but the relevant workflow has not yet been tested on PlanetScope images.

In addition to the gap size concern, the appropriate selection of reference images is vital to achieve successful and accurate gap-filling, but could result in large uncertainties when rapid land cover changes occur. Gap-filling methods often use either single or multiple reference images. The single-reference-image methods require a clear reference image that is acquired at a close date to the target image for gap-filling (Zhu et al., 2012b). Although PlanetScope has a daily-to-weekly temporal coverage, this data requirement remains limited in tropical or subtropical areas due to continuous cloud contamination, especially during high rainfall wet seasons (Roy et. al., 2021; Wang et al., 2020). To address this issue, the multiple-reference-image methods select multiple reference images that are cloud-free and have the highest similarity with the target image for each cloud patch instead of the whole image (Lin et al., 2014). Some studies have demonstrated that multiple-reference-image methods have better performance than single-reference-image methods (Chen et al., 2017; Lin et al., 2013), but it remains difficult to obtain cloud-free reference images for large cloud patches, as the similarities between the reference and target images for all objects even within a cloud patch can vary considerably (Cao et al., 2020), particularly for the heterogeneous landscapes that appear more often in the high-resolution PlanetScope image (as Fig. S1). For this, here we test whether adaptively selecting either single or two reference images on an object-class basis rather than at the cloud patch level can help improve gap-filling performance by reducing the uncertainties associated with both spatial heterogeneity and diverse (with/without rapid) land cover change scenarios.

The goal of this study thus aims to develop an accurate and robust gap-filling method to automatically reconstruct missing data in PlanetScope image time series. Given the challenges in gap-filling PlanetScope satellites, we highlighted two major novelties of this work as follows. First, we employed an object-based segmentation method in conjunction with an unsupervised classification method to effectively reconstruct the missing information regardless of image gap size (corresponding to the number of image pixels contaminated by clouds/cloud shadows). Second, for each object-class (i.e. a group of pixels of the same object and the same class), we adopted an adaptive method to automatically select one or two reference images from the image time series for gap-filling, and assigned the best guess of the temporal change scenarios (with/without rapid land cover change) based on their temporal changing patterns. We named this method the object and class based gap-filling (OCBGF) method. With this new method, we hope to effectively gap-fill missing pixels in PlanetScope image time series and enhance the data availability and continuity for fine-scale land surface monitoring. As a proof-of-concept, we selected the four sites with different land cover types covering a variety of climate zones on the global land surface, and compared OCBGF with three state-of-the-art methods, NSPI, ARRC, and SAMSTS.

2. Study sites and materials

2.1 Study sites

To test the OCBGF method, we selected four sites that are representative of different land cover types and climate zones across the global land surface (Fig. 1). These include 1) a managed forest landscape from a Eucalyptus plantation site in South Brazil (henceforth ‘Euc-plantation’), 2) a cropland landscape from Iowa in the United States (henceforth ‘Iowa-cropland’), 3) a metropolitan urban area from Beijing city in China (henceforth ‘Beijing-urban’), and 4) a moist forest landscape from the Barro Colorado Island in Panama (henceforth ‘BCI-forest’). Details regarding the location, climate, seasons, and spatial extent of these four sites are shown in Table 1.

We selected these study sites for two reasons. First, these four sites represent different land cover types and climate zones with various levels of landscape heterogeneity and span a large range of annual precipitation from 492-2052 mm per year (Table 1). For example, the Beijing-urban site contains a large number of built-ups, such as buildings, highways, airport, as well as vegetated areas in north-eastern suburbs; the Iowa-cropland site is typical heterogeneous cropland with various crops types, such as corn, soybeans, and oats; the BCI-forest site includes a moist forest mixed with deciduous and evergreen broadleaf species, roads, buildings, and bare soils in surrounding areas of the Barro Colorado Island; and the Euc-plantation site has mixed land cover types, including commercial Eucalyptus, forest patches, buildings, roads, and bare soils. In addition, among these sites, the BCI-forest site has a tropical climate with the most serious cloud contamination resulting in continuous data missing in a year, especially during the wet season with monthly rainfall higher than 100 mm.

Second, there is modest to strong temporal variability in surface reflectance at these sites. For most gap-filling methods, large temporal variations in surface reflectance caused by plant phenology and abrupt land cover change could cause uncertainties (Cao et al., 2020). The Beijing-urban site has experienced substantial urbanization in recent years, converting vegetated surfaces into buildings (Cao et al., 2020). The Iowa-cropland site has a significant crop rotation between corn and soybean and vegetation phenology changes, such as seeding, growing, and harvesting (Cao et al., 2020). The BCI-forest site exhibits modest to large seasonal reflectance variability with significant amounts of leaf shedding and leaf exchange during the high-light dry season (Detto et al., 2018; Park et al., 2019). The Euc-plantation site experiences strong seasonal reflectance changes caused by harvesting, and land cover changes caused by deforestation and reforestation (Lopez-Poma et al., 2020; Qin et al., 2019).

For more details about the ecological, hydrological, and topographic characteristics of these four sites, refer to previous studies (Campoe et al., 2012; Leigh, 1999; Liu et al., 2018; Qi et al., 2011).

2.2 Materials

The four-band, 3-m resolution PlanetScope data from Planet Labs PBC. (San Francisco, CA, USA) were used in this study. We accessed the data from https://www.planet.com/ through a research and education license with Planet Labs PBC. We downloaded the PlanetScope data for all four sites that span the whole annual cycle covering a wide range of percent data missing for each site (Fig. 1c). The level 3B surface reflectance product of PlanetScope was used, which has been orthorectified and pre-processed (including geometric, radiometric, and atmospheric corrections) (Planet, 2021). To retain as much good data as possible, we assessed all available data that met the following criteria: 1) "standard" quality level (that refers to an image meeting a variety of quality standards; Planet, 2021), 2) rectification with ground control points, 3) solar zenith angle < 80° and view zenith angle < 5°, 4) snow cover<5%, 5) cloud cover<80%, 6) thin cloud cover<5%. Consequently, across the four sites throughout the entire annual cycle of 2018, a total of 271 days of PlanetScope land surface reflectance images were accessed (Table 1). Each of these images is from the sensor type of PS2 (denoting Dove Classic) and has a spatial extent of 20km×20km (6667×6667 pixels). Notably, since the BCI-forest site has some water surfaces occupied by lakes and rivers, prior to testing our method, we masked out these elements in the corresponding

PlanetScope images using the same masking method as Wang et al. (2021). This was conducted to minimize the gap-filling uncertainty associated with large spectral variability with the water surfaces across both space and time (Qiu et al., 2019; Zhu et al., 2015; Zhu and Woodcock, 2012) and misclassification error encountered in cloud and shadow detection on water surfaces (Zhu and Helmer, 2018).

3. Methods

We applied the OCBGF method for gap-filling PlanetScope time-series images using one target image (I_{t0}, where $t0$ is the acquisition date of the target image) in this time series as an example. The same procedure also applies to any other images in the time series that need to be gap-filled. The time-series images can be fully or partly cloud-free (e.g. percent data missing ranging from 0-100%; Fig. 1c), and OCBGF automatically selects those image pixels in the time series with valid values at or nearby the locations of gaps to gap-fill I_{t0}. Particularly, we assumed that pixels with both high spectral similarity and spatial continuity tend to have similar temporal changing patterns (Chen et al., 2011; Zhu et al., 2012b). Based on this assumption, we first developed a model to capture the temporal changing patterns using spectrally and spatially similar pixels outside the gaps, and then applied this model to gap-fill the missing pixels in I_{t0}. There are four tasks in OCBGF for gap-filling (Fig. 2). First, we conducted pixel-level quality control for each image in the time series to minimize potential cloud and cloud shadow impacts. Second, before gap-filling, we integrated a cloud-free time series (I_{ts}) covering around a one-month time period with the minimum data missing nearby each I_{t0}, then segmented and classified I_{ts} into independent objects of different classes. Third, for each object-class (i.e. a group of pixels of the same object and the same class) with missing pixels in I_{t0}, we performed gap-filling with two scenarios: single-reference-image (no rapid land cover change) and two-reference-images (rapid land cover change) scenarios. Finally, we conducted post-image-processing with a guided filtering approach to further reduce random noises in the gap-filled images while retaining detailed information.

3.1 Pixel-level quality control (Task 1)

Before gap-filling, strict pixel-level quality control is essential (Chen et al., 2004). For this, we adopted a two-step approach. First, we applied a recently developed automatic cloud/cloud shadow detection method, STI-ACSS (Wang et al, 2021), for initial and automatic screening of cloud/cloud shadow across the entire image time series. The STI-ACSS method was used because it has been demonstrated to be more effective in cloud/cloud shadow screening in PlanetScope images compared with the default PlanetScope quality control layers and other state-of-the-art cloud/cloud shadow methods (Wang et al., 2021). Second, since no cloud/cloud shadow detection algorithm is perfect, to minimize the uncertainty associated with the residual effects in STI-ACSS, we applied additional pixel-level quality control to automatically mask out any remaining cloud/cloud shadow pixels with two sub-steps: (1) to minimize the potential effects associated with thin clouds/cloud shadows surrounding the detected clouds/cloud shadows, we used morphological dilation with a structure element on each cloud and cloud shadow mask to label the border areas of clouds/cloud shadows (Soille, 1999; Soille and Pesaresi, 2002). The structure element is a disk-shaped matrix (Zheng, 1995) with a user-specified size (5×5 pixels in this study); and (2) to best filter remaining cloud/cloud shadow pixels that have larger temporal variability in their spectral reflectance than clear-sky pixels (Wang et al., 2021), we examined their band-specific pixel values with a user-specified threshold pair (e.g. 1 and 99 percentiles) across the entire image time series and masked out the pixels with values either greater than the upper bound (e.g. 99 percentile) or smaller than the lower bound (e.g. 1 percentile). More details regarding the three parameters used in this task, i.e., the size of the structure element, and a pair of percentile thresholds for additional data quality control, the way to determine their values, and the values of recommendation are summarized in Table S1.

3.2 Object segmentation and classification (Task 2)

Because the natural landscape is often mixed with multiple objects within the same class or multiple classes belonging to the same object like the example shown in Fig. S1, the conventional approaches that either rely on any given moving window or are based on the object-based image segmentation are not sufficiently accurate to address this issue. Thus, here we developed a new method relying on the concept of object-class, and also compared this new method with the conventional approaches that relied on either object or class alone. The new object-class method integrated an object-based segmentation method with an unsupervised classification method for automatic identifications of pixels, by which it helps to group those spectrally and spatially similar pixels into the same object and the same class.

Specifically, we first used an object-based image analysis (OBIA) method as implemented by Watkins and Van Niekerk (2019a, 2019b) to segment all pixels in a composited cloud-free time series (I_{ts}) into individual objects with irregular sizes based on their spectral and spatial characteristics. This OBIA method was adopted because of its high accuracy in detecting object boundaries as well as its capability to operate without any prior knowledge (Watkins and Van Niekerk, 2019a). This method includes the following four steps, through which we generated a segmentation map (S_{ts}) for I_{ts}:

(Step 1) We generated a composited cloud-free time series I_{ts} for automatic image segmentation. For this, the 10 temporally-adjacent images (i.e. 5 images before/after I_{t0}) were automatically selected from the whole image time series, with which we further derived I_{ts} by identifying the clearest images with minimum missing pixels from each 5-image group (Fig. S2). Each 5-images before and after I_{t0} were selected because this roughly represents a short temporal period of one month in PlanetScope images (Roy et al., 2021). The two selected clearest images (I_{ts}; n=2, with one before I_{t0} and one after I_{t0}) were subsequently used for image automatic segmentation as described in steps 2-4. It is noted that i) when there are more than one clearest images within each 5-image group, the image with the least temporal distance to I_{t0} is selected; and ii) these two clearest images both need to cover the full area of I_{t0}; otherwise, we would use the clearest image across the entire time series. It is also important to note that the approach used here is empirical and usable for PlanetScope satellites of high temporal resolution, but might need some fine-tuning when applying to other satellites of different temporal resolutions.

(Step 2) We employed an image high-pass filtering approach with a 4-neighborhood Laplacian filter (Gonzalez and Woods, 2006; Solomon and Breckon, 2010) for each image in I_{ts}, through which we could include more spatial details to facilitate the subsequent image segmentation.

(Step 3) We applied a commonly-used Canny edge-detection operator (Canny, 1986) on each sharpened image in I_{ts} and aggregated the edge layers with a union operation to generate one composite edge layer for image segmentation.

(Step 4) We employed a widely-used region-based segmentation approach, watershed segmentation (Li et al., 2010), on the above-derived edge layer for deriving S_{ts} (Fig. 2). This approach divides regions of local minima (catchment basins) into individual objects based on edges with high gradient magnitudes (Salman, 2006). This watershed segmentation approach used 8-connected connectivity to specify the directions of adjacent pixels in the neighborhood of a given pixel.

After deriving individual objects using the above image segmentation method, we employed a commonly-used unsupervised classifier, *k*-means (Lloyd, 1982), on I_{ts}. The *k*-means classifier automatically classifies all valid pixels of I_{ts} into *K* classes based on their spectral similarity, and generates a classification map C_{ts} (Fig. 2). The classifier minimized the sum of the squared Euclidean distance of spectral reflectance between each pixel and the class centroid to estimate classification results. There is only one parameter, *K* (the number of classes), in this unsupervised classification. To determine *K*, we specified a range of *K* values (5-10 in our study) and then followed the Calinski-Harabasz clustering evaluation criterion (Calinski and Harabasz, 1974) to automatically determine the optimal *K* within this range. A further sensitivity analysis based on the setting of different *K* ranges (Fig. S3) confirms that the *K* range set for this study is valid. It is worthy to note that the setting of *K* range can vary with the level of landscape heterogeneity, thus we included this recommendation in Table S1 for the readers' reference.

With the S_{ts} and C_{ts} derived above, we finally identified the object and class type for each pixel, and grouped the pixels belonging to the same object and the same class as a representative gap-filling unit of 'object-class'.

3.3 Gap-filling with two scenarios (Task 3)

Since rapid land cover change could also introduce uncertainty into gap-filling, we first differentiated the two land cover change scenarios (with or without rapid land cover change) and then adopted scenario-specific gap-filling procedures. These two steps were conducted on an object-class basis as follows.

3.3.1 Differentiations of two land cover change scenarios

For each given object-class with missing pixels in I_{t0}, we divided them into the following two cases: (1) only part of the pixels were missing, and (2) all pixels were missing.

For case 1, we first searched each target object-class throughout the full image time series, and then identified the reference images that had valid pixels for gap-filling. For each object-class, based on their temporal distances to I_{t0}, we further determined the reference images with the closest temporal distance before I_{t0} (noted as I_{t0-1}) and after I_{t0} (noted as I_{t0+1}) as well as the image with the closest absolute temporal distance to I_{t0} (denoted as I_{tn}, which could be either I_{t0-1} or I_{t0+1}; Fig. 2). For case 2, we turned to the neighboring object-class that had valid pixels in I_{t0}, belonged to the same class, and was spatially closest to the target object-class, and repeated the same process as case 1, through which we identified the corresponding reference images of I_{t0-1}, I_{t0+1}, and I_{tn}. Since the reference image(s) were selected on an object-class basis, our approach can thus optimize the temporal distance between reference image(s) and the target image, consequently reducing the gap-filling uncertainty associated with the landscape heterogeneity and rapid land cover changes over time.

With the identified I_{tn}, we then calculated the correlation coefficient for each spectral band between all the valid pixels of I_{t0} and I_{tn} that either belongs to the target object-class (for case 1) or the nearest adjacent object-class (for case 2). We further calculated the average correlation coefficient across all four spectral bands ($\overline{r_h}$) and assigned the land cover change scenarios by comparing $\overline{r_h}$ with an empirically determined threshold (r_T) as Eq. 1.

$$\begin{cases} \overline{r_h} \geq r_T, \text{no rapid land cover change scenario} \\ \overline{r_h} < r_T, \text{rapid land cover change scenario} \end{cases} \tag{1}$$

To assess how the threshold value (r_T) would affect our results, we varied r_T from 0.65 to 0.95 with an interval of 0.05. Additionally, to assess how our method would perform in these two independent scenarios in Eq. 1, we also separately tested these two scenarios. The sensitivity analysis across four sites (Figs. S4 and S5) demonstrated that using a fixed threshold of 0.80 across all sites obtains the highest averaged accuracies, and is also quite comparable with the site-specific optimized threshold with a very narrow range of 0.75~0.80 across all four sites. Details regarding how to determine r_T as well as the recommended value of r_T for use are summarized in Table S1.

3.3.2 Scenario-specific gap-filling

Scenario 1 – Single-reference-image

For each object (h)-class (k), under the no rapid land cover change scenario, we used only one image I_{tn} for gap-filling. This single-image referencing method relies on a linear regression model between I_{t0} and I_{tn}:

$$I_{t0,b,i} = \alpha_{hk,b} \times I_{tn,b,i} + \beta_{hk,b} + \varepsilon_{hk,b} \tag{2}$$

where $I_{t0,b,i}$ and $I_{tn,b,i}$ are spectral reflectance values of the valid pixel i in band b for I_{t0} and I_{tn}, respectively, and $\alpha_{hk,b}$ and $\beta_{hk,b}$ are coefficients of this linear model for band b, and $\varepsilon_{hk,b}$ is the residual error. The band-specific coefficients can be solved by linearly regressing the spectral reflectance values of all the valid pixels between I_{t0} and I_{tn}.

Before gap-filling, we further assessed the type of any given object-class into the following two categories: 1) if only part of the pixels within the target object-class were missing, we labelled them as category 1, by which we applied Eq. 2 to the remaining valid pixels of the same object-class to first derive the linear regression model, and then applied this model to gap-fill the missing ones using Eq. 3; and 2) if all the pixels within the target object-class were missing, we labelled them as category 2, by which we applied Eq. 2 to those valid pixels of the same class available from the nearest adjacent object to derive the linear model, and then applied the derived model to gap-fill those missing ones using Eq. 4 below.

$$\widehat{I_{t0,b,j}} = \alpha_{hk,b} \times I_{tn,b,j} + \beta_{hk,b} \tag{3}$$

Where $\widehat{I_{t0,b,j}}$ is the predicted reflectance value of missing pixel j in band b for I_{t0}, $I_{tn,b,j}$ is the spectral reflectance value of valid pixel j in band b for I_{tn}, and $\alpha_{hk,b}$ and $\beta_{hk,b}$ are estimated coefficients using Eq. 2 for the category 1.

$$\widehat{I_{t0,b,j}} = \alpha_{ck,b} \times I_{tn,b,j} + \beta_{ck,b} \tag{4}$$

Where $\widehat{I_{t0,b,j}}$ is the predicted reflectance value of missing pixel j in band b for I_{t0}, $I_{tn,b,j}$ is the spectral reflectance value of valid pixel j in band b for I_{tn}, and $\alpha_{ck,b}$ and $\beta_{ck,b}$ are estimated coefficients using Eq. 2 for the category 2, which represents the closest object (c) to the target object (h) with valid pixels in both I_{t0} and I_{tn} for class (k).

Scenario 2 – Two-reference-images

For each object (h)-class(k), under the rapid land cover change scenario, we used the two images (one before I_{t0}, I_{t0-1}, and one after I_{t0}, I_{t0+1}) for gap-filling. This two-image referencing method relies on a linear regression model between I_{t0} and the two reference images (Eqs. 5). The key assumption of this method is that the use of a pair of reference images would be more efficient to capture the rapid land cover change compared with the use of a single reference image for gap filling, especially when the target image is different from either reference image (i.e. $\overline{r_h}<r_T$).

$$I_{t0,b,i} - I_{t0+1,b,i} = \alpha'_{hk,b}(I_{t0-1,b,i} - I_{t0+1,b,i}) + \beta_{hk,b} + \varepsilon_{hk,b} \quad (5)$$

where $I_{t0,b,i}$, $I_{t0-1,b,i}$, and $I_{t0+1,b,i}$ are spectral reflectance values of valid pixel i in band b for I_{t0}, I_{t0-1} and I_{t0+1}, respectively, and $\alpha'_{hk,b}$ and $\beta_{hk,b}$ are coefficients of this linear model in band b, and $\varepsilon_{hk,b}$ is the residual error. The band-specific coefficients can be solved by linearly regressing the spectral reflectance values of all the valid pixels between $I_{t0} - I_{t0+1}$ and $I_{t0-1} - I_{t0+1}$.

Similar to scenario 1, before gap-filling, we further assessed the type of any given object-class into the following two categories: 1) if only part of the pixels within the target object-class were missing, we labelled them as category 3, by which we applied Eq. 5 to the remaining valid pixels of the same object-class to first derive the linear regression model, and then applied this model to gap-fill those missing ones using Eq. 6 below; and 2) if all the pixels within the target object-class were missing, we labelled them as category 4, by which we applied Eq. 5 to those valid pixels of the same class available from the nearest adjacent object to derive the linear model, and then applied the derived model to gap-fill those missing ones using Eq. 7 below.

$$\widehat{I_{t0,b,j}} = \alpha'_{hk,b} \times \left(I_{t0-1,b,j} - I_{t0+1,b,j}\right) + I_{t0+1,b,j} + \beta_{hk,b} \quad (6)$$

Where $\widehat{I_{t0,b,j}}$ is the predicted reflectance value of missing pixel j in band b for I_{t0} , $\left(I_{t0-1,b,j} - I_{t0+1,b,j}\right)$ is the spectral reflectance difference of valid pixel j in band b between the two reference images, and $\alpha_{hk,b}$ and $\beta_{hk,b}$ are estimated coefficients using Eq. 5 for the category 3.

$$\widehat{I_{t0,b,j}} = \alpha'_{ck,b} \times \left(I_{t0-1,b,j} - I_{t0+1,b,j}\right) + I_{t0+1,b,j} + \beta_{ck,b} \quad (7)$$

Where $\widehat{I_{t0,b,j}}$ is the predicted reflectance value of missing pixel j in band b for I_{t0}, $\left(I_{t0-1,b,j} - I_{t0+1,b,j}\right)$ is the spectral reflectance difference of valid pixel j in band b between the two reference images, and $\alpha'_{ck,b}$ and $\beta_{ck,b}$ are estimated coefficients using Eq. 5 for the category 4, which represents the closest object (c) to the target object (h) with valid pixels for class (k) in I_{t0}.

3.4 Post-image-processing (Task 4)

To further reduce random noises (e.g. salt-and-pepper noises; Fig. S6b) while retaining high-resolution details in the gap-filling results derived above, we conducted post-image-processing for the target image using a commonly-used edge-preserving filter, the guided filter (He et al., 2013; Fig. S6c), with a demonstrated improvement in the gap-filling results (Fig. S7). The guided filter computes the filtering output by a linear transformation of a guidance image, but does not suffer from the gradient reversal artefacts (He et al., 2013; Li et al., 2017). Specifically, a cloud-free image or gap-filled image with the maximum number of valid pixels across the full image time series is selected as the guidance image for the subsequent guided filter on the target image.

3.5 Evaluation

To evaluate the performance of our method, we performed three tests. The first two tests were based on simulation experiments, and the third test used real cloud contaminated images. Throughout all these three tests, we also cross-compared our OCBGF model performance with the other three state-of-the-art methods, i.e. NSPI (Zhu et al., 2012a), ARRC (Cao et al., 2020), and SAMSTS (Yan and Roy, 2018).

In test 1, we evaluated the effects of gap size on the gap-filling results by randomly selecting a fully or near-fully clear-sky image from each study site and artificially creating seven gaps ranging from 10^4 to 10^6 pixels (10^5 to 10^7 m^2) on each image (Table 2). Since the percent data missing of the above test remains small (e.g. within 20%), to further evaluate the boundary beyond which level of percent data missing would largely affect the gap-filling results among different methods, we used the Euc-plantation and Iowa-cropland sites as examples, and artificially created six circle gaps sharing the same center (as the image center) but having a radius range from 500 to 3000 pixels with an increment of 500 pixels. These six gaps corresponded to the percent data missing of 2%, 7%, 16%, 28%, 44%, and 64%, respectively, on the selected images.

In test 2, we evaluated the effects of rapid land cover change on the gap-filling results by selecting 24 fully or near-fully clear-sky images (10 from Euc-plantation, 7 from Iowa-cropland, 4 from Beijing-urban, and 3 from BCI-forest sites). These images were selected based on the fact that there were obvious land cover changes between the temporally adjacent images and the images covered different seasons of the year 2018. With these images, we then purposely masked out the areas that experienced land cover changes over time.

In both test 1 and 2, we used the original clear images as benchmark. The model performance was assessed on a spectral band basis, using the following three metrics, i.e., the root mean square error (RMSE), the correlation coefficient (CC), and the structure similarity index (SSIM) (Wang et al., 2004). We used these three metrics in combination, as each of them reflects different aspects of model performance assessments, with RMSE assessing the spectral band differences between the predicted and reference images, CC reflecting the degree of correlation and similarity between the predicted and reference images, and SSIM quantifying the structural similarities between the predicted and reference images.

In test 3, we randomly selected one cloud contaminated image from each study site, and visually assessed the structural continuity and color consistency by comparing the gap-filled areas with those clear areas spatially/temporally adjacent to the gaps. For cross-method comparisons, we used the default parameters of NSPI, ARRC, and SAMSTS, following Zhu et al. (2012a), Cao et al. (2020), and Yan and Roy (2018), respectively.

4. Results

4.1 Cross-method comparison for different simulated gap sizes

To assess the effects of gap size on the gap-filled results for all four methods (NSPI, ARRC, SAMSTS, and OCBGF), we conducted accuracy assessments on seven deliberately created gaps of different size categories using the original clear images as benchmark. Across all four sites, we found that OCBGF yielded the highest accuracy with the smallest variation across seven gap size categories (RMSE=0.0065 (mean) ± 0.0027 (standard deviation), CC=0.95±0.06, and SSIM= 0.94±0.06; Fig. 3 and Table 3). By contrast, the NSPI method generated the lowest model accuracies with the largest variations across different gap size categories (RMSE=0.0113 (mean) ± 0.0058 (standard deviation), CC=0.87±0.15, and SSIM= 0.86±0.16; Fig. 3 and Table 3). Among these four sites, OCBGF obtained much higher accuracies compared to the other three methods, consistently across all gap size categories in Euc-plantation and BCI-forest (Fig. 3); OCBGF achieves a comparable accuracy in Iowa-cropland and Beijing-urban as ARRC, and in Iowa-cropland as SAMSTS, all of which are higher than NSPI (Fig. 3). Moreover, within each site and across all three accuracy metrics, we observed that OCBGF yields more stable model accuracies across all seven gap size categories than the other three models (Fig. 3). Especially in Euc-plantation, Beijing-urban, and BCI-forest, the three accuracy metrics of OCBGF tend to stabilize with gap size while NSPI and SAMSTS display larger accuracy fluctuations across gap sizes, with large near-linear reductions in model accuracies with gaps size in Beijing-urban and BCI-forest. The stability in model accuracies of ARRC with gap size is comparable with OCBGF in Euc-plantation, Iowa-cropland, and Beijing-urban, but shows much larger fluctuations with a near-linear reduction in model accuracies with gaps size in BCI-forest. These results altogether demonstrate that OCBGF is the most accurate method showing the least sensitivity to the gap size among all comparative methods, followed by SAMSTS, ARRC, and NSPI (Table 3).

Among the three accuracy metrics across all four methods, we observed a much lower CC and SSIM at the BCI-forest site relative to the other three sites (Fig. 4). This is likely associated with the low data availability at this site with a high-rainfall environment (Fig. 1c and Table 1). Meanwhile, we observed that all the four methods maintain comparable CC and SSIM in the other three lower-rainfall sites (i.e. Euc-plantation, Iowa-cropland, and Beijing-urban; Table 1), and the accuracy reduction of CC and SSIM from these drier sites to the high-rainfall BCI-forest site is much smaller in OCBGF than the other three methods, demonstrating that OCBGF could be more robust than the other three methods in gap-filling, especially in the high rainfall environment.

To further aid the visual interpretation of these cross-method comparisons, we next presented the gap-filling results (Fig. S8) and associated magnified correspondences (Fig. 5) for each of the four sites. Across all three methods, we observed that the gap-filling results are overall very comparable in the smaller gap size categories from 1 to 5 (Fig. S8), but tend to diverge in the remaining two bigger gap size categories (Figs. S8 and 5), with the OCBGF results being the closest to original images in larger gaps. Also in Fig. 5, NSPI displays some blocking artifacts (Figs. 5a, b, and d) in the interior areas of Euc-plantation, Iowa-cropland, and BCI-forest, while generating some strip artifacts (Figs. 5c and e) in the heterogeneous areas of Beijing-urban and BCI-forest. ARRC exhibits obvious salt-and-pepper noises (Figs. 5f-h) in heterogeneous areas mixed with different land covers across all four sites. SAMSTS also presents some blocking artifacts (Fig. 5i), accompanied by moderate errors in gap-filling land surface areas of high reflectance values (Figs. 5j and k). In contrast, OCBGF achieves a more satisfying performance across all these large gaps, despite small portions of fine details missing as compared to the original images (Fig. 5). Collectively, both quantitative and visual cross-method comparisons show that OCBGF is more accurate and stable across all land cover types and gap size categories than the other three methods.

Finally, to evaluate whether a large percent of data missing (e.g. >30%) would affect the gap-filling results among different methods, we conducted accuracy assessments on the six artificial circle gaps covering the percent data missing ranging from 2%, 7%, 16%, 28%, 44%, and 64% (Fig. 6a). Our results show that OCBGF consistently yields the best and most stable model performance across the full range of percent data missing (Figs. 6b and c), in contrast with the other three models that display much weaker and more variable model accuracies across different percent data missing. For example, when the percent data missing is relatively low (2% and 7%), ARRC and SAMSTS obtain higher accuracies across all three accuracy metrics. With a further increase in percent data missing (16% and 28%), SAMSTS and NSPI obtain higher accuracies in CC and SSIM, while NSPI generates lower accuracy in RMSE. When the percent data missing is much higher (44% and 64%), NSPI generates higher accuracies in CC and SSIM but a lower accuracy in RMSE. Our proposed method of OCBGF performs best consistently and most stable across all the three accuracy metrics, with only a very minor accuracy reduction with an increasing percent data missing.

4.2 Cross-method comparison for rapid land cover changes

To examine the effects of rapid land cover changes on the gap-filled results for all four methods, we turned to the simulation approach, and conducted accuracy assessments on 24 representative images using original clear images as benchmarks. Across all sites and seasons, we found that

OCBGF outperforms the other three methods (Fig. 7), with the highest accuracies and least accuracy variations (RMSE = 0.0082 (mean) ± 0.0023 (standard deviation), CC = 0.95±0.031, and SSIM = 0.94±0.035; Table 4), followed by SAMSTS and ARRC that performs slightly better than NSPI but with much larger variations (Table 4). Among these four sites and seasons, we also observed that OCBGF obtains the consistently highest accuracies and lowest variations across all four sites (Fig. 8). While for ARRC, SAMSTS, and NSPI, we found that these three methods have mixed successes, with ARRC obtaining the lowest accuracies in the BCI-forest site, NSPI obtaining the lowest accuracies in most cases of the Iowa-cropland site, and SAMSTS obtaining the lowest accuracies in Beijing-urban site.

To visually interpret these cross-method comparisons further, we presented the two magnified regions with rapid land cover changes for each site using their original clear images as benchmarks (Fig. 9). Specifically, there are two types of rapid changes being selected: 1) gradual color transformation (e.g. Figs. 9b, h, and i) and 2) abrupt color transformations (e.g. Figs. 9a, c, d, e, f, and g). To put the gap-filling results in the time-series context, we displayed the two images (before and after) temporally adjacent to each target image. Across the cases showing a gradual color transformation (Figs. 9b, h, and i), our results show that the gap-filling results from OCBGF are closer to the benchmarks than the other three methods, in which NSPI yields some blurred colors, ARRC gives some salt-and-pepper noises, while SAMSTS exhibits some obvious errors around the heterogeneous areas with complex land covers. Moreover, we also examined the cases with abrupt color transformations across the image time series (Figs. 9a, c, d, e, and f). Across these cases, we found that OCBGF is consistently more efficient in capturing the rapid land cover changes with well-constructed spatial details compared with the other three methods. However, we also found that there remains a small portion (Fig. 9g) of areas in Beijing-urban being inaccurately reconstructed by all four methods. Compared with OCBGF, NSPI, ARRC, and SAMSTS are found to have even larger portions of areas being inaccurately reconstructed, especially for those heterogeneous areas that often come with obvious errors/noises. These results altogether demonstrate that OCBGF performs the best in gap-filling those areas with rapid land cover changes among all four methods examined here.

4.3 Cross-method comparison in the real-world practices

To assess the performance of these methods in real-world scenarios, we selected one representative image mixed with cloud contaminations for each of the four sites. Our results show that OCBGF achieves the best model performance with the least sensitivity to clouds/cloud shadows in general (Fig. 10). Particularly, the magnified gap-filling results (highlighted in orange squares; Fig. 10) show that OCBGF obtains the highest consistency with the temporally most adjacent images while displaying the best spatial continuity near gap boundaries relative to the other three methods. In contrast, NSPI exhibits blurred colors, especially in large gaps, ARRC exhibits obvious noises and errors in heterogeneous areas, and SAMSTS displays imprecise boundaries, especially at the intersections among different land cover types. These results again demonstrate that OCBGF generates the best gap-filling results, with potential to best recover the signal for those areas with big data gaps and rapid land cover changes.

5. Discussion

Over recent years, PlanetScope data has been increasingly used for monitoring rapid and fine-scale land surface dynamics that can scale up to create impacts on understanding global environmental change and ecosystem responses (Forzieri et al., 2021; Heinrich et al., 2021; Taubert et al., 2018; Zeng et al., 2018). However, a critical challenge remains with the lack of an effective and accurate method for gap-filling missing data caused by cloud and cloud shadow contaminations. For example, as shown in a 50-ha plot at the Panamanian BCI-forest site (Fig. S9), the original PlanetScope time-series images are not able to accurately quantify leaf phenology at the patch scale of a 12 m×12 m area when they are seriously contaminated by clouds/cloud shadows, while the gap-filled time-series images could help recover the considerable within-site fine-scale phenology variability, suggesting the necessity of gap-filling. To address this challenge, we developed an object and class based gap-filling method, OCBGF, and evaluated the effectiveness of this method across four sites spanning a large variety in land cover types (i.e. plantation, cropland, urban, and forest; Fig. 1b), spatial heterogeneity (e.g. homogenous and heterogeneous landscapes; Fig. 1b), annual precipitation (e.g. 492-2052 mm per year; Table 1), and percent data missing (e.g. ranging from 0% to 100%; Fig. 1c). When comparing our OCBGF method with the other three state-of-the-art methods (NSPI, ARRC, and SAMSTS) on both simulation tests and real-world cloud-contaminated cases (Figs. 3-10, Tables 3 and 4), our results consistently show that OCBGF is the most accurate and stable gap-filling method for recovering cloud/cloud shadow-induced data gaps in PlanetScope time-series images.

The effectiveness of OCBGF relies on the following two strengths. First, to minimize the uncertainty associated with large data gaps and landscape heterogeneity, OCBGF adopts an object-based segmentation method in conjunction with a classification method to leverage valid pixels outside the gaps for gap-filling. It is widely known that large data gaps (corresponding to a large pixel number contaminated by clouds/cloud shadows) often result in insufficient valid pixels with spectral similarity and spatial continuity, making the gap-filling difficult with large uncertainty, especially for those interior pixels of large gaps (Shen et al., 2015; Yan and Roy, 2018). Meanwhile, in high spatial resolution PlanetScope images, the landscape heterogeneity can be an increasingly important issue, as there are often mixed classes within an object and mixed objects within a class (e.g. Fig. S1), making the conventional gap-filling methods relying on either moving window or object-based segmentation approaches challenging to resolve this issue. To address these issues, similar to previous approaches of NSPI (Zhu et al., 2018) and ARRC (Cao et al., 2020), we also employed an unsupervised classification method to classify the composited cloud-free time series into different classes based on their spectral similarity. Different from those previous approaches relying on a moving window to identify the valid pixels of the same class to gap-fill each target pixel (Cao et al., 2020; Zhu et al., 2018), we integrated an object-based segmentation method with an unsupervised classification method to identify the object and class type of pixels, by which we searched valid pixels belonging to the same object-class or the nearest object-class to help gap-fill each target object-class with missing pixels. This object-class integrated approach facilitates the search for spectrally and spatially similar pixels, producing results with higher accuracies when gap-filling missing pixels in general, and particularly when gaps are large (Figs. 3, 4, and 6). An additional analysis (Fig. S10), showing that our object-class approach can generate more stable and accurate gap-filling results compared with the other two approaches respectively relying on object and class alone, further suggests that it can be a more accurate way to characterize the landscape heterogeneity in the real world. Although SAMSTS applied a similar integration framework—segment-and-clustering to gap-fill large gaps in Landsat images (Yan and Roy, 2018), it is different from our method in two aspects: 1) SAMSTS used a region-growing image segmentation method based on spectral similarity, while our method adopted an object-based segmentation method based on both spectral characteristics and spatial texture; 2) SAMSTS performed the classification task on derived segmented objects while our classification was performed on a pixel level, which makes our method more accurate as there were often more than one land cover types within a segmented object. Our cross-model comparisons (Figs. 4, 6, and 8) also demonstrate that our method is superior to SAMSTS for gap-filling PlanetScope image time series across all sites, seasons, data gap sizes, and land cover change scenarios. This study thus represents a new attempt at using the object-class integration approach to effectively recover big data gaps in high spatial-resolution satellite images.

Second, to minimize the gap-filling uncertainty associated with rapid land cover changes, OCBGF adopts a scenario-specific gap-filling approach. With a 3-m spatial resolution, PlanetScope is able to detect finer-scale changes, including rapid changes in land cover and land use types, than traditional satellites (Cheng et al., 2020; Rao et al., 2021; Rasanen and Virtanen, 2019). However, these fine-scale rapid changes also present a challenge due to the difficulty of identifying valid pixels in heterogeneous landscapes suitable for gap-filling. To address this challenge, the selection of the closest and clearest reference images has been increasingly recognized as an essential step to gap-fill areas with rapid land cover changes (Cao et al., 2020). However, two issues remain in the existing methods that rely on either single-reference-image or multiple-reference-image. First, existing gap-filling methods often require one or more cloud-free reference images for the whole image or cloud patch acquired close to the target image to aid gap-filling (Chen et al., 2011; Zhu et al., 2012a), resulting in an extended temporal distance to the target image and consequently introducing more uncertainty during gap-filling. Second, some of these methods commonly select the reference images based on their spectral similarities with the target image on the scale of a whole image or cloud patch, without considering the potential diverse land cover types underneath the target cloud patch (Chen et al., 2017; Lin et al., 2013). To address these two issues, we also turned to the reference image(s) but focused on the scale of each object-class, through which the diversity in temporal changing patterns associated with different object-classes, even within the same cloud patch, is accounted for. Meanwhile, our approach also minimized the temporal distance between the reference image(s) and the target image, as the reference image(s) were now selected at the object-class scale, which is much smaller than the scale of a whole image or cloud patch and facilitates the search of the temporally nearest reference image(s). Additionally, to account for rapid land cover changes, we adopted an extra step to first use a fixed threshold (Figs. S4 and S5) to determine the scenarios with/without rapid cover changes for each target object-class, and then adopted the scenario-specific gap-filling procedure, through which we aimed to further minimize the gap-filling uncertainty associated with rapid land cover changes. Such a scenario-specific approach integrating two different scenarios also generates consistently higher accuracies across all four sites than independent scenarios (Fig. S4), demonstrating the effectiveness of including this approach in OCBGF for gap-filling. As landscape heterogeneity is a common feature for high spatial resolution satellite imagery and is becoming more common with increasing anthropogenic activities, our successful implementation of OCBGF with higher accuracies than NSPI, ARRC, and SAMSTS (Figs. 7 and 8) highlights the potential to extend this approach to other high spatial-resolution satellite images in the future.

Our study also identifies three important next steps that need to be considered for future advances. First, we observed that there remained some errors in OCBGF for gap-filling areas experiencing rapid fine-scale changes with varying degrees of inconsistency in temporal changing patterns when compared with valid pixels of the same or the nearest object-class (Fig. 9g). This is likely because OCBGF gap-filled each target object-class with missing pixels by assuming that valid pixels of the same or the nearest object-class tend to have similar temporal changing patterns to the missing pixels, which would not work if the temporal changing patterns are inconsistent (Fig. 9g). A potential way of solving this issue is by using deep learning based approaches, as several recent studies have shown that they are effective at recognizing subtle patterns and characterizing nonlinear relationships between the missing and valid pixels in time-series imagery (Li et al., 2019; Zhang et al., 2020; Zi et al., 2021). Second, although OCBGF has demonstrated a higher accuracy than the other three methods at the high-rainfall BCI-forest site (Figs. 4 and 8), it may encounter challenges with extremely low availability of valid observations, such as in certain tropical/subtropical areas which are subjected to persistent cloud cover at the month-to-year scale (Wang et al., 2020, 2021). To investigate such impacts, we examined the sensitivity of OCBGF on the frequency of low percent valid data (e.g. < 10%) and the time interval of clear-sky observations, respectively. Our results in Fig. S11 demonstrated that the model accuracy is stable when the low percent valid data frequency is not more than 30% and then decreases afterward. Meanwhile, we also observed that OCBGF yields stable model accuracies when the average time interval of clear-sky observations is not more than 19 days and accuracy decreased linearly afterward (Fig. S12). This high sensitivity to the frequency of valid observations is likely because OCBGF relies on time-series images temporally adjacent to the target image for gap-filling, and low valid data availability might lead to insufficient or no valid pixels to build models to capture temporal changing patterns. We thus recommend using as many as available time-series images as input to increase the valid-observation frequency. As for certain tropical/subtropical areas with extremely low valid data availability, we recommend fusing optical PlanetScope with synthetic aperture radar (SAR) data (Huang et al., 2015; Li et al., 2020b; Meraner et al., 2020; Pipia et al., 2019) that are unaffected by cloud interference, which could be an important next step in the future. Third, although all the data used in this study came from the PS2 sensor type, there are more and more sensor types available in PlanetScope constellation, and the inconsistent radiometric calibration across sensor types could also introduce uncertainties into the gap-filling results. To investigate such impacts, we used the data in the year 2020 from the Iowa-cropland site that are mixed with three different sensor types (PS2, PS2.SD, and PSB.SD) as an example, and addressed the cross-sensor radiometric inconsistency issue using a cross-calibration approach developed specifically for the PlanetScope constellation (Wang et al., 2020). Using the gap-filling results from the calibrated PlanetScope as benchmarks, our results (Fig. S13) show that the raw and calibrated gap-filled results are almost identical. These suggest that cross-sensor radiometric inconsistency may only cause very minor differences in the gap-filling results, but a more comprehensive assessment is still needed.

6. Conclusions

Here, we developed a new object and class based gap-filling (OCBGF) method for automatic gap-filling of missing pixels in PlanetScope time-series images. This method integrates object segmentation and classification for automatic identification of those spectrally and spatially similar pixels, and uses a scenario-specific gap-filling procedure to recover missing pixels for each target object-class in the image time series. The accuracy of OCBGF was evaluated at four contrasting study sites spanning large gradients in land cover types (plantation, cropland, urban, and forest; all excluding water bodies), spatial heterogeneity, annual precipitation, and percent data missing. Relative to three other state-of-the-art gap-filling methods, NSPI, ARRC, and SAMSTS, OCBGF obtained the highest accuracy regardless of simulation tests covering various gap sizes and rapid land cover changes, and also achieved better performances for gap-filling cloud/cloud shadow-induced data gaps in real-world practices. With these assessments, our results suggest that OCBGF is an accurate and robust approach for gap-filling PlanetScope time-series images, and has potential to be extended to other high spatial resolution satellites, e.g. Gaofen-1, Skysat, and SPOT-7 (Gaofen-1, 2020; SkySat, 2020; SPOT-7, 2020), in the future attempts.

Acknowledgements: We would like to thank Helene C. Muller-Landau for sharing the drone data collected over Barro Colorado Island, Panama (Fig. S9), which was collected by a team of scientists, including Helene C. Muller-Landau, Jonathan Dandois, Milton Garcia, Stephanie Bohlman, Sam Grubinger, and Raquel Araujo. The work was supported by National Natural Science Foundation of China (#31922090). J. Wu was in part supported by the Innovation and Technology Fund (funding support to State Key Laboratories in Hong Kong of Agorobiotechnology) of the HKSAR, China. J. Wang was in part supported by the Division of Ecology and Biodiversity PDF research award; C. Lee was in part supported by HKU 45th round PDF scheme and HKU Seed Fund for Basic Research (#202011159154). The PlanetScope data used in this study were accessed through the Education and Research Program, contracted between Planet Labs PBC and HKU.

References:

Calinski, T., & Harabasz, J. (1974). A Dendrite Method for Cluster Analysis. *Communications in Statistics - Theory and Methods, 3*, 1-27.

Campoe, O.C., Stape, J.L., Laclau, J.P., Marsden, C., & Nouvellon, Y. (2012). Stand-level patterns of carbon fluxes and partitioning in a Eucalyptus grandis plantation across a gradient of productivity, in Sao Paulo State, Brazil. *Tree Physiology*, *32*, 696-706.

Canny, J. (1986). A Computational Approach to Edge-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*, 679-698.

Cao, R.Y., Chen, Y., Chen, J., Zhu, X.L., & Shen, M.G. (2020). Thick cloud removal in Landsat images based on autoregression of Landsat time-series data. *Remote Sensing of Environment*, *249*, https://doi.org/10.1016/j.rse.2020.112001.

Case, N., & Vitti, A. (2021). Reconstruction of Multi-Temporal Satellite Imagery by Coupling Variational Segmentation and Radiometric Analysis. *ISPRS International Journal of Geo-Information*, *10*, https://doi.org/10.3390/ijgi10010017.

Chen, B., Huang, B., Chen, L.F., & Xu, B. (2017). Spatially and Temporally Weighted Regression: A Novel Method to Produce Continuous Cloud-Free Landsat Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *55*, 27-37.

Chen, B., Jin, Y.F., & Brown, P. (2019). An enhanced bloom index for quantifying floral phenology using multi-scale remote sensing observations. Isprs Journal of Photogrammetry and Remote Sensing, 156, 108-120.

Chen, J., Jonsson, P., Tamura, M., Gu, Z.H., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment*, *91*, 332-344.

Chen, J., Zhu, X.L., Vogelmann, J.E., Gao, F., & Jin, S.M. (2011). A simple and effective method for filling gaps in Landsat ETM plus SLC-off images. *Remote Sensing of Environment*, *115*, 1053-1064.

Cheng, Y., Vrieling, A., Fava, F., Meroni, M., Marshall, M., & Gachoki, S. (2020). Phenology of short vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sensing of Environment*, *248*, https://doi.org/10.1016/j.rse.2020.112004.

Dechant, B., Ryu, Y., Badgley, G., Kohler, P., Rascher, U., Migliavacca, M., Zhang, Y.G., Tagliabue, G., Guan, K.Y., Rossini, M., Goulas, Y., Zeng, Y.L., Frankenberg, C., & Berry, J.A. (2022). NIRvP: A robust structural proxy for sun-induced chlorophyll fluorescence and photosynthesis across scales. Remote Sensing of Environment, 268.

Detto, M., Wright, S.J., Calderon, O., & Muller-Landau, H.C. (2018). Resource acquisition and reproductive strategies of tropical forest in response to the El Nino-Southern Oscillation. *Nature Communications*, *9*, https://doi.org/10.1038/s41467-41018-03306-41469.

Feng, Y., Zeng, Z.Z., Searchinger, T.D., Ziegler, A.D., Wu, J., Wang, D.S., He, X.Y., Elsen, P.R., Ciais, P., Xu, R.R., Guo, Z.L., Peng, L.Q., Tao, Y.H., Spracklen, D.V., Holden, J., Liu, X.P., Zheng, Y., Xu, P., Chen, J., Jiang, X., Song, X.P., Lakshmi, V., Wood, E.F., & Zheng, C.M. (2022). Doubling of annual forest carbon loss over the tropics during the early twenty-first century. Nature Sustainability, 1-8.

Forzieri, G., Girardello, M., Ceccherini, G., Spinoni, J., Feyen, L., Hartmann, H., Beck, P.S.A., Camps-Valls, G., Chirici, G., Mauri, A., & Cescatti, A. (2021). Emergent vulnerability to climate-driven disturbances in European forests. *Nature Communications*, *12*, 1-12.

Gaofen-1 (2020). GF-1 (Gaofen-1) High-resolution Imaging Satellite / CHEOS series of China. In: https://directory.eoportal.org/web/eoportal/satellite-missions/g/gaofen-1 (accessed 03 October 2020).

Gonzalez, R.C., & Woods, R.E. (2006). *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc.

Griffiths, P., Nendel, C., & Hostert, P. (2019). Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment, 220*, 135-151.

Guan, X.D., Liu, G.H., Huang, C., Liu, Q.S., Wu, C.S., Jin, Y., & Li, Y.F. (2017). An Object-Based Linear Weight Assignment Fusion Scheme to Improve Classification Accuracy Using Landsat and MODIS Data at the Decision Level. Ieee Transactions on Geoscience and Remote Sensing, 55, 6989-7002.

Halls, J.N., & Magolan, J.L. (2019). A Methodology to Assess Land Use Development, Flooding, and Wetland Change as Indicators of Coastal Vulnerability. Remote Sensing, 11, https://doi.org/10.3390/rs11192260.

He, K.M., Sun, J., & Tang, X.O. (2013). Guided Image Filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*, 1397-1409.

Heinrich, V.H.A., Dalagnol, R., Cassol, H.L.G., Rosan, T.M., de Almeida, C.T., Silva, C.H.L., Campanharo, W.A., House, J.I., Sitch, S., Hales, T.C., Adami, M., Anderson, L.O., & Aragao, L.E.O.C. (2021). Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nature Communications*, *12*, 1-11.

Hossain, M.D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, *150*, 115-134.

Huang, B., Li, Y., Han, X.Y., Cui, Y.Z., Li, W.B., & Li, R.R. (2015). Cloud Removal From Optical Satellite Imagery With SAR Imagery Using Sparse Representation. *IEEE Geoscience and Remote Sensing Letters*, *12*, 1046-1050.

Huang, B., & Zhang, H.K. (2014). Spatio-temporal reflectance fusion via unmixing: accounting for both phenological and land-cover changes. International Journal of Remote Sensing, 35, 6213-6233.

Ju, J.C., & Roy, D.P. (2008). The availability of cloud-free Landsat ETM plus data over the conterminous United States and globally. *Remote Sensing of Environment*, *112*, 1196-1211.

Kimm, H., Guan, K.Y., Jiang, C.Y., Peng, B., Gentry, L.F., Wilkin, S.C., Wang, S.B., Cai, Y.P., Bernacchi, C.J., Peng, J., & Luo, Y.N. (2020). Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the US Corn Belt using Planet Labs CubeSat and STAIR fusion data. *Remote Sensing of Environment, 239*, https://doi.org/10.1016/j.rse.2019.111615.

Kong, J., Ryu, Y., Liu, J., Dechant, B., Rey-Sanchez, C., Shortt, R., Szutu, D., Verfaillie, J., Houborg, R., & Baldocchi, D.D. (2022). Matching high resolution satellite data and flux tower footprints improves their agreement in photosynthesis estimates. Agricultural and Forest Meteorology, 316, 108878.

Leigh, E.G. (1999). *Tropical forest ecology: a view from Barro Colorado Island*. New York, NY: Oxford University Press.

Li, D.R., Zhang, G.F., Wu, Z.C., & Yi, L.N. (2010). An Edge Embedded Marker-Based Watershed Algorithm for High Spatial Resolution Remote Sensing Image Segmentation. *IEEE Transactions on Image Processing*, *19*, 2781-2787.

Li, L.F., Franklin, M., Girguis, M., Lurmann, F., Wu, J., Pavlovic, N., Breton, C., Gilliland, F., & Habre, R. (2020a). Spatiotemporal imputation of MAIAC AOD using deep learning with downscaling. *Remote Sensing of Environment*, *237*, https://doi.org/10.1016/j.rse.2019.111584.

Li, W.B., Li, Y., & Chan, J.C.W. (2020b). Thick Cloud Removal With Optical and SAR Imagery via Convolutional-Mapping-Deconvolutional Network. *IEEE Transactions on Geoscience and Remote Sensing*, *58*, 2865-2879.

Li, W.B., Li, Y., Chen, D., & Chan, J.C.W. (2019). Thin cloud removal with residual symmetrical concatenation network. *ISPRS Journal of Photogrammetry and Remote Sensing, 153*, 137-150.

Li, Z.W., Shen, H.F., Li, H.F., Xia, G.S., Gamba, P., & Zhang, L.P. (2017). Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery. *Remote Sensing of Environment, 191*, 342-358.

Lin, C.H., Lai, K.H., Chen, Z.B., & Chen, J.Y. (2014). Patch-Based Information Reconstruction of Cloud-Contaminated Multitemporal Images. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 163-174.

Lin, C.H., Tsai, P.H., Lai, K.H., & Chen, J.Y. (2013). Cloud Removal From Multitemporal Satellite Images Using Information Cloning. *IEEE Transactions on Geoscience and Remote Sensing, 51*, 232-241.

Liu, J., Zhenyao, S.Y., Yan, T.Z., & Yang, Y.C. (2018). Source identification and impact of landscape pattern on riverine nitrogen pollution in a typical urbanized watershed, Beijing, China. *Science of the Total Environment, 628-629*, 1296-1307.

Liu, R.G., Shang, R., Liu, Y., & Lu, X.L. (2017). Global evaluation of gap-filling approaches for seasonal NDVI with considering vegetation growth trajectory, protection of key point, noise resistance and curve stability. *Remote Sensing of Environment, 189*, 164-179.

Lloyd, S.P. (1982). Least-Squares Quantization in Pcm. *IEEE Transactions on Information Theory*, *28*, 129-137.

Lopez-Poma, R., Pivello, V.R., de Brito, G.S., & Bautista, S. (2020). Impact of the conversion of Brazilian woodland savanna (cerradao) to pasture and Eucalyptus plantations on soil nitrogen mineralization. *Science of the Total Environment, 704*, https://doi.org/10.1016/j.scitotenv.2019.135397.

Lorenzi, L., Melgani, F., & Mercier, G. (2011). Inpainting Strategies for Reconstruction of Missing Data in VHR Images. *IEEE Geoscience and Remote Sensing Letters*, *8*, 914-918.

Luo, Y.N., Guanb, K.Y., & Pen, J. (2018). STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product. *Remote Sensing of Environment*, *214*, 87-99.

Maalouf, A., Carre, P., Augereau, B., & Fernandez-Maloigne, C. (2009). A Bandelet-Based Inpainting Technique for Clouds Removal From Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 2363-2371.

Maxwell, S.K., Schmidt, G.L., & Storey, J.C. (2007). A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images. *International Journal of Remote Sensing*, *28*, 5339-5356.

McCabe, M.F., Aragon, B., Houborg, R., & Mascaro, J. (2017). CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation. Water Resources Research, 53, 10017-10024.

Meraner, A., Ebel, P., Zhu, X.X., & Schmitt, M. (2020). Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. *ISPRS Journal of Photogrammetry and Remote Sensing, 166*, 333-346.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q.H. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment, 115*, 1145-1161.

Norris, J.R., Allen, R.J., Evan, A.T., Zelinka, M.D., O'Dell, C.W., & Klein, S.A. (2016). Evidence for climate change in the satellite cloud record. *Nature*, *536*, 72–75.

Park, J.Y., Muller-Landau, H.C., Lichstein, J.W., Rifai, S.W., Dandois, J.P., & Bohlman, S.A. (2019). Quantifying Leaf Phenology of Individual Trees and Species in a Tropical Forest Using Unmanned Aerial Vehicle (UAV) Images. *Remote Sensing*, *11*, https://doi.org/10.3390/rs11131534.

Pipia, L., Munoz-Mari, J., Amin, E., Belda, S., Camps-Valls, G., & Verrelst, J. (2019). Fusing optical and SAR time series for LAI gap filling with multioutput Gaussian processes. *Remote Sensing of Environment, 235*, https://doi.org/10.1016/j.rse.2019.111452.

Planet (2021). PLANET SURFACE REFLECTANCE VERSION 2.0 [WWW Document]. In. https://assets.planet.com/marketing/PDF/Planet_Surface_Reflectance_Technical_White_Paper.pdf (accessed 14 April 2021).

Pringle, M.J., Schmidt, M., & Muir, J.S. (2009). Geostatistical interpolation of SLC-off Landsat ETM plus images. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*, 654-664.

Qi, Z.M., Helmers, M.J., & Kaleita, A.L. (2011). Soil water dynamics under various agricultural land covers on a subsurface drained field in north-central Iowa, USA. *Agricultural Water Management, 98*, 665-674.

Qin, Y.W., Xiao, X.M., Dong, J.W., Zhang, Y., Wu, X.C., Shimabukuro, Y., Arai, E., Biradar, C., Wang, J., Zou, Z.H., Liu, F., Shi, Z., Doughty, R., & Moore, B. (2019). Improved estimates of forest cover and loss in the Brazilian Amazon in 2000-2017. *Nature Sustainability, 2*, 764-772.

Qiu, S., Zhu, Z., & He, B.B. (2019). Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery. Remote Sensing of Environment, 231, https://doi.org/10.1016/j.rse.2019.1005.1024.

Rao, P., Zhou, W.Q., Bhattarai, N., Srivastava, A.K., Singh, B., Poonia, S., Lobell, D.B., & Jain, M. (2021). Using Sentinel-1, Sentinel-2, and Planet Imagery to Map Crop Type of Smallholder Farms. *Remote Sensing, 13*, https://doi.org/10.3390/rs13101870.

Rasanen, A., & Virtanen, T. (2019). Data and resolution requirements in mapping vegetation in spatially heterogeneous landscapes. *Remote Sensing of Environment, 230*, https://doi.org/10.1016/j.rse.2019.1005.1026.

Roy, D.P., Huang, H.Y., Houborg, R., & Martins, V.S. (2021). A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery. *Remote Sensing of Environment*, *264*, https://doi.org/10.1016/j.rse.2021.112586.

Roy, D.P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., & Lindquist, E. (2008). Multi-temporal MODIS-Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment, 112*, 3112-3130.

Sadeh, Y., Zhu, X., Dunkerley, D., Walker, J.P., Zhang, Y.X., Rozenstein, O., Manivasagam, V.S., & Chenu, K. (2021). Fusion of Sentinel-2 and PlanetScope time-series data into daily 3 m surface reflectance and wheat LAI monitoring. *International Journal of Applied Earth Observation and Geoinformation, 96*, 102260.

Salman, N. (2006). Image segmentation based on watershed and edge detection techniques. *The International Arab Journal of Information Technology, 3*, 104-110.

Shen, H.F., Li, X.H., Cheng, Q., Zeng, C., Yang, G., Li, H.F., & Zhang, L.P. (2015). Missing Information Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geoscience and Remote Sensing Magazine, 3*, 61-85.

SkySat (2020). SkySat constellation of Terra Bella - formerly SkySat Imaging Program of Skybox Imaging. In. https://directory.eoportal.org/web/eoportal/satellite-missions/s/skysat (accessed 03 October 2020).

Soille, P. (1999). *Morphological image analysis: Principles and applications* Springer-Verlag Berlin Heidelberg.

Soille, P., & Pesaresi, M. (2002). Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 2042-2055.

Solomon, C., & Breckon, T. (2010). *Fundaments of Digital Image Processing, A Practical Approach with example in Matlab*. Wiley-Blackwell, Hoboken.

SPOT-7 (2020). SPOT-6 and SPOT-7 Commercial Imaging Constellation. In. https://directory.eoportal.org/web/eoportal/satellite-missions/s/spot-6-7 (accessed 03 October 2019).

Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Muller, M.S., Rodig, E., Wiegand, T., & Huth, A. (2018). Global patterns of tropical forest fragmentation. *Nature, 554*, 519-522.

Wang, F.R., Fan, X.M., Yunus, A.P., Subramanian, S.S., Alonso-Rodriguez, A., Dai, L.X., Xu, Q., & Huang, R.Q. (2019). Coseismic landslides triggered by the 2018 Hokkaido, Japan (M-w 6.6), earthquake: spatial distribution, controlling factors, and possible failure mechanism. Landslides, 16, 1551-1566.

Wang, J., Yang, D.D., Chen, S.L., Zhu, X.L., Wu, S.B., Bogonovich, M., Guo, Z.F., Zhu, Z., & Wu, J. (2021). Automatic cloud and cloud shadow detection in tropical areas for PlanetScope satellite images. *Remote Sensing of Environment, 264*, https://doi.org/10.1016/j.rse.2021.112604.

Wang, J., Yang, D.D., Detto, M., Nelson, B.W., Chen, M., Guan, K.Y., Wu, S.B., Yan, Z.B., & Wu, J. (2020). Multi-scale integration of satellite remote sensing improves characterization of dry-season green-up in an Amazon tropical evergreen forest. *Remote Sensing of Environment, 246*, https://doi.org/10.1016/j.rse.2020.111865.

Wang, Z., Bovik, A.C., Sheikh, H.R., & Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*, 600-612.

Watkins, B., & Van Niekerk, A. (2019a). Automating field boundary delineation with multi-temporal Sentinel-2 imagery. *Computers and Electronics in Agriculture, 167*, https://doi.org/10.1016/j.compag.2019.105078.

Watkins, B., & Van Niekerk, A. (2019b). A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery. *Computers and Electronics in Agriculture, 158*, 294-302.

Wu, S.B., Wang, J., Yan, Z.B., Song, G.Q., Chen, Y., Ma, Q., Deng, M.F., Wu, Y.T., Zhao, Y.Y., Guo, Z.F., Xu, X.T., Yang, X., Su, Y.J., Liu, L.L., & Wu, J. (2021). Monitoring tree-crown scale autumn leaf phenology in a temperate forest with an integration of PlanetScope and drone remote sensing observations. *ISPRS Journal of Photogrammetry and Remote Sensing, 171*, 36-48.

Wu, W., Ge, L.Q., Luo, J.C., Huan, R.H., & Yang, Y.P. (2018). A Spectral-Temporal Patch-Based Missing Area Reconstruction for Time-Series Images. *Remote Sensing, 10*, https://doi.org/10.3390/rs10101560.

Xi, W.Q., Du, S.H., Wang, Y.C., & Zhang, X.Y. (2019). A spatiotemporal cube model for analyzing satellite image time series: Application to land-cover mapping and change detection. Remote Sensing of Environment, 231.

Yan, L., & Roy, D.P. (2018). Large-Area Gap Filling of Landsat Reflectance Time Series by Spectral-Angle-Mapper Based Spatio-Temporal Similarity (SAMSTS). *Remote Sensing, 10*, https://doi.org/10.3390/rs10040609.

Yan, L., & Roy, D.P. (2020). Spatially and temporally complete Landsat reflectance time series modelling: The fill-and-fit approach. *Remote Sensing of Environment, 241*, https://doi.org/10.1016/j.rse.2020.111718.

Zeng, C., Shen, H.F., & Zhang, L.P. (2013). Recovering missing pixels for Landsat ETM plus SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment, 131*, 182-194.

Zeng, Z.Z., Estes, L., Ziegler, A.D., Chen, A.P., Searchinger, T., Hua, F.Y., Guan, K.Y., Jintrawet, A., & Wood, E.F. (2018). Highland cropland expansion and forest loss in Southeast Asia in the twenty-first century. *Nature Geoscience, 11*, 556-562.

Zeng, Z.Z., Wang, D.S., Yang, L., Wu, J., Ziegler, A.D., Liu, M.F., Ciais, P., Searchinger, T.D., Yang, Z.L., Chen, D.L., Chen, A.P., Li, L.Z.X., Piao, S.L., Taylor, D., Cai, X.T., Pan, M., Peng, L.Q., Lin, P.R., Gower, D., Feng, Y., Zheng, C.M., Guan, K.Y., Lian, X., Wang, T., Wang, L., Jeong, S.J., Wei, Z.W., Sheffield, J., Caylor, K., & Wood, E.F. (2021). Deforestation-induced warming over tropical mountain regions regulated by elevation. Nature Geoscience, 14, 23-29.

Zhang, C., Li, W., & Travis, D. (2007). Gaps-fill of SLC-off Landsat ETM plus satellite image using a geostatistical approach. *International Journal of Remote Sensing, 28*, 5103-5122.

Zhang, Q., Yuan, Q.Q., Li, J., Li, Z.W., Shen, H.F., & Zhang, L.P. (2020). Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing, 162*, 148-160.

Zhang, Q., Yuan, Q.Q., Zeng, C., Li, X.H., & Wei, Y.C. (2018). Missing Data Reconstruction in Remote Sensing Image With a Unified Spatial-Temporal-Spectral Deep Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing, 56*, 4274-4288.

Zheng, X., Gong, P., Strome, M. (1995). Characterizing spatial structure of tree canopy using colour photographs and mathematical morphology. *Canadian Journal of Remote Sensing, 21*, 421-429.

Zhu, X.L., Gao, F., Liu, D.S., & Chen, J. (2012a). A Modified Neighborhood Similar Pixel Interpolator Approach for Removing Thick Clouds in Landsat Images. *IEEE Geoscience and Remote Sensing Letters, 9*, 521-525.

Zhu, X.L., & Helmer, E.H. (2018). An automatic method for screening clouds and cloud shadows in optical satellite image time series in cloudy regions. *Remote Sensing of Environment, 214*, 135-153.

Zhu, X.L., Helmer, E.H., Chen, J., & Liu, D.S. (2018). *An Automatic System for Reconstructing High-Quality Seasonal Landsat Time Series. Chapter 2 (Pages 25-42), In: Qihao Weng, Ed. Remote Sensing: Time Series Image Processing*. Boca Raton: CRC Press.

Zhu, X.L., Liu, D.S., & Chen, J. (2012b). A new geostatistical approach for filling gaps in Landsat ETM plus SLC-off images. *Remote Sensing of Environment, 124*, 49-60.

Zhu, Z., Wang, S.X., & Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. Remote Sensing of Environment, 159, 269-277.

Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sensing of Environment, 118, 83-94.

Zi, Y., Xie, F.Y., Zhang, N., Jiang, Z.G., Zhu, W.T., & Zhang, H.P. (2021). Thin Cloud Removal for Multispectral Remote Sensing Images Using Convolutional Neural Networks Combined With an Imaging Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 3811-3823.

Figure 1. Study sites. (a) locations (the yellow stars), (b) false color composites (RGB=NIR-Red-Green) of PlanetScope images, and (c) percent data missing of PlanetScope time-series images of the four testing sites (i.e. S1-S4; spatial extent: 20km×20km, temporal coverage: January-December, 2018), including Euc-plantation, Iowa-cropland, Beijing-urban, and Barro Colorado Island (BCI)-forest. The map in panel (a) is adapted from National Geographic, ESRI. The percent data missing in panel (c) is derived from STI-ACSS method (Wang et al., 2021).

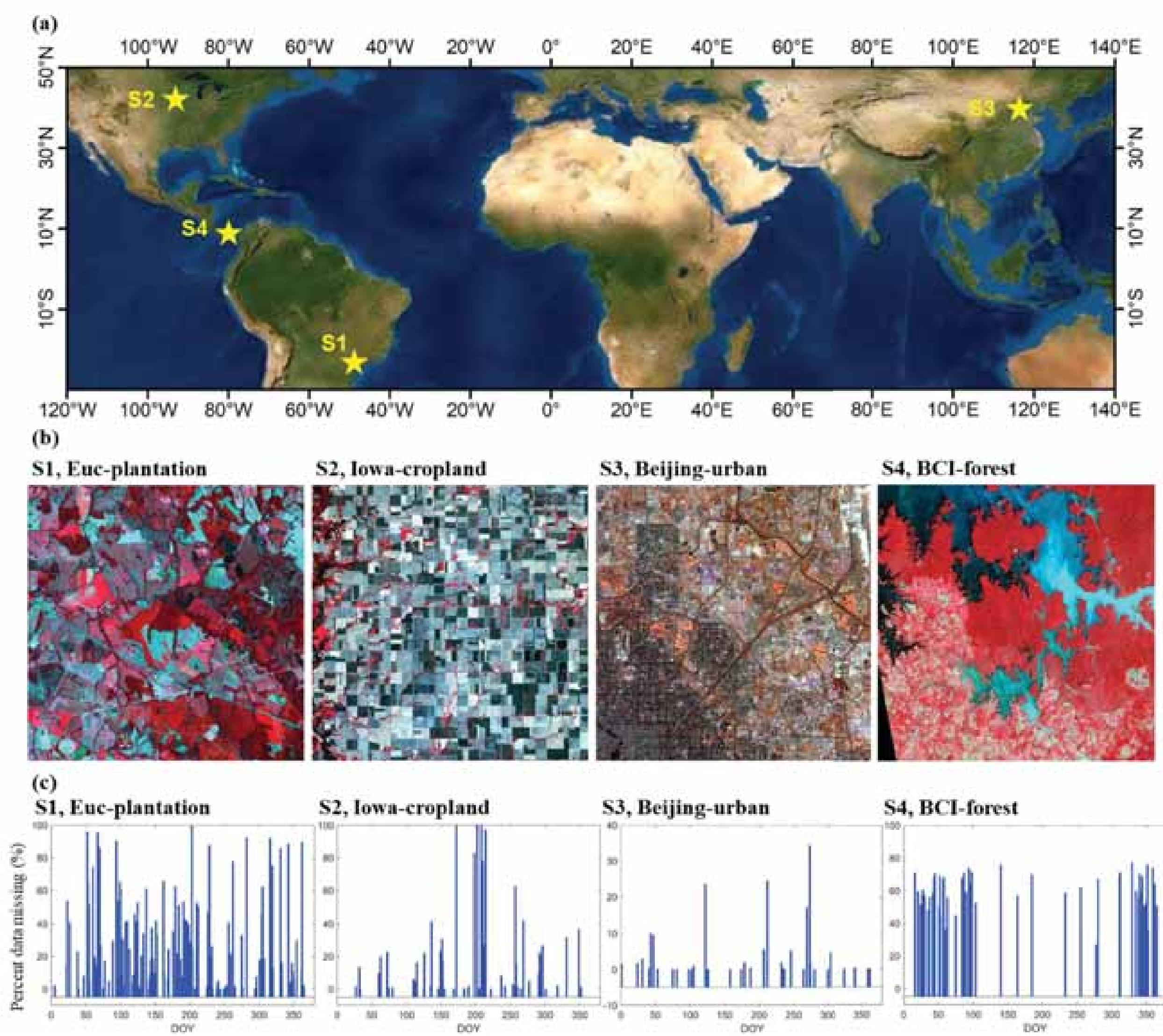

Figure 2. Flowchart of the four key tasks of the object and class based gap-filling (OCBGF) method. Task 1: pixel-level quality control; Task 2: object segmentation and classification; Task 3: gap-filling with two scenarios: single-reference-image and two-reference-images scenarios; Task 4: post-image-processing.

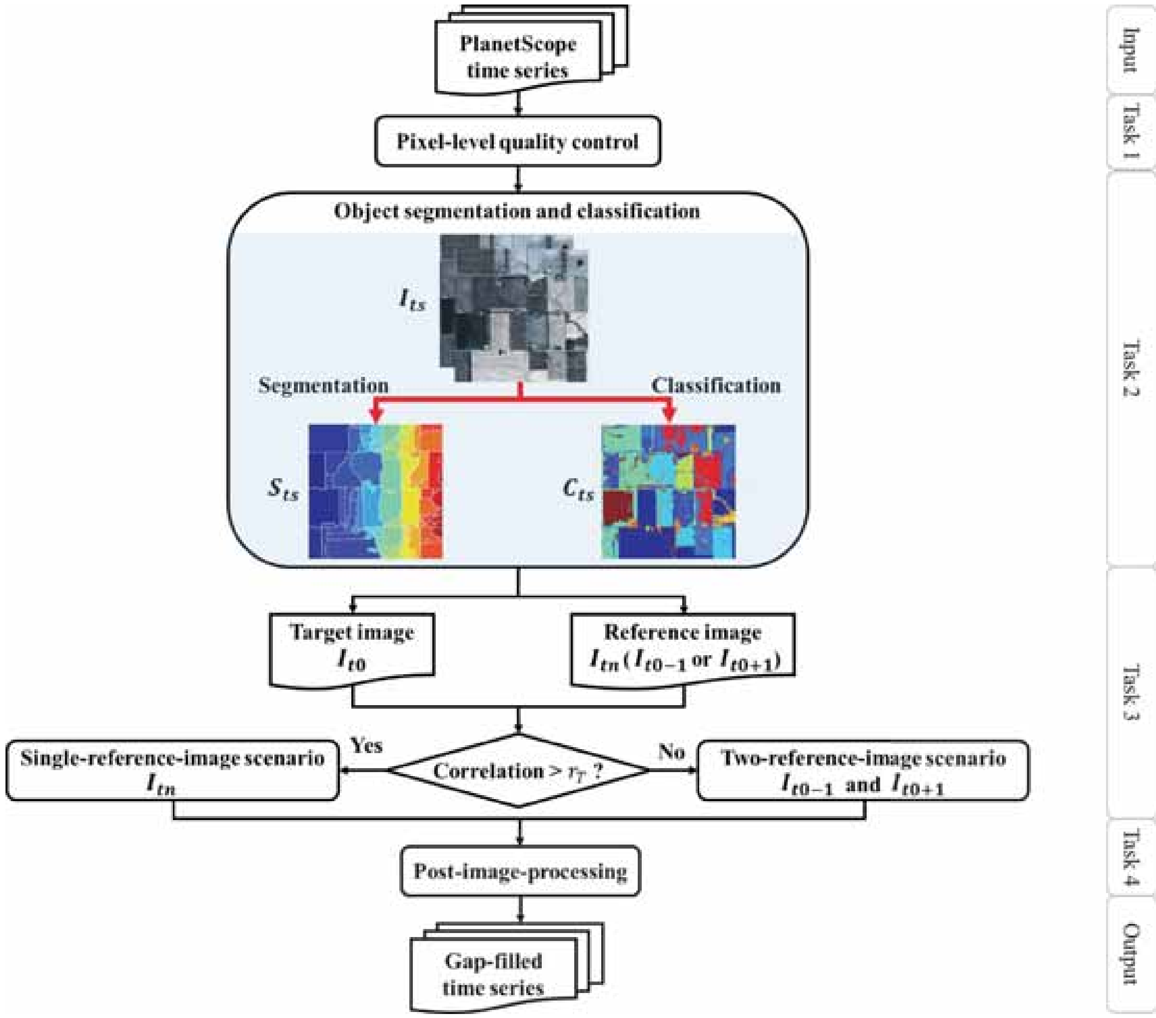

Figure 3. Cross-method comparison of gap-filled results across seven gap size categories (Table 2) at four sites using original images as benchmarks. The results of cross-method comparison include four-band average values of three indices: the root mean square error (RMSE), the correlation coefficient (CC), and the structure similarity index (SSIM) across all four sites. The four methods examined here include Neighborhood Similar Pixel Interpolator (NSPI), AutoRegression to Remove Clouds (ARRC), Spectral-Angle-Mapper Based Spatio-Temporal Similarity (SAMSTS), and OCBGF.

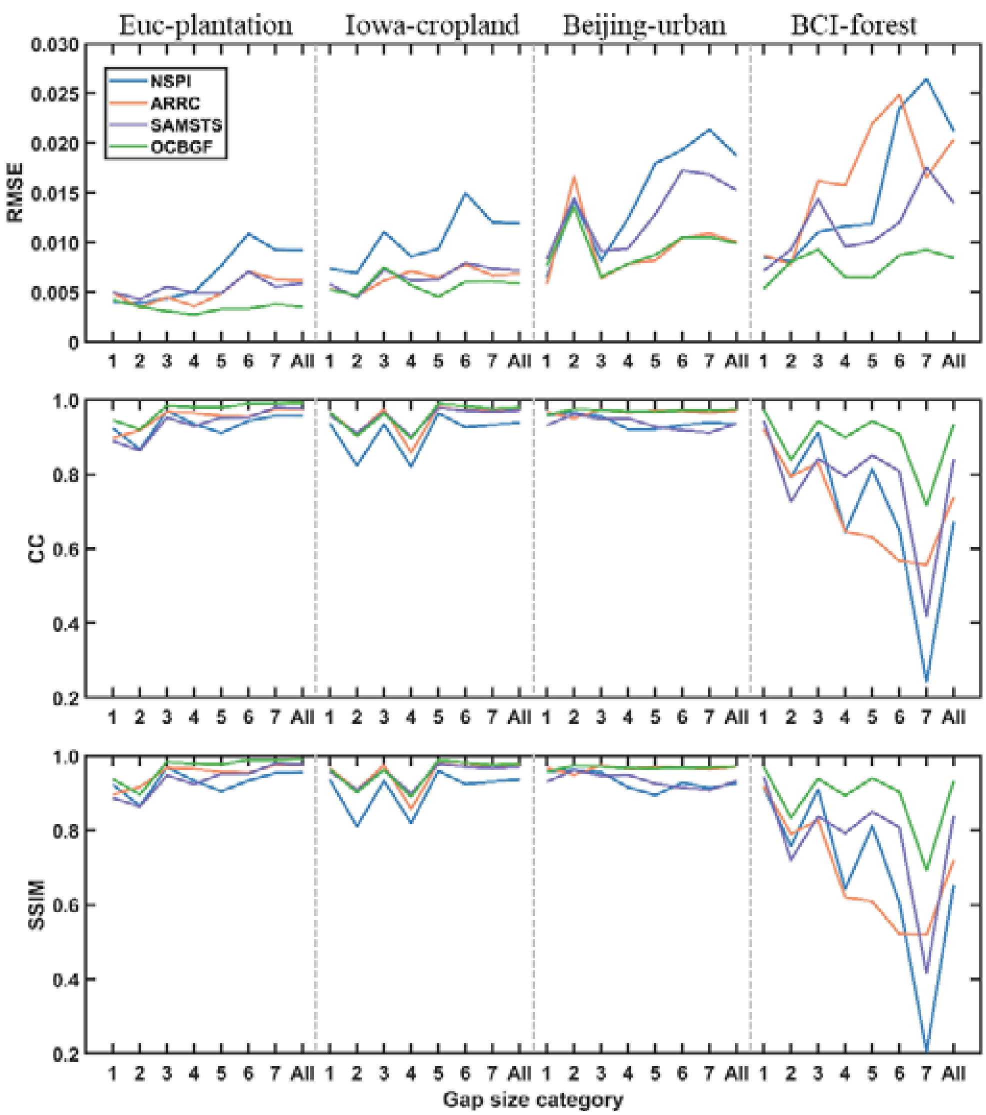

Figure 4. Cross-method comparison of all seven gap size categories across four sites using original images as benchmarks. Three metrics were used for the cross-method comparison, including RMSE, CC, and SSIM. The four methods examined here include NSPI, ARRC, SAMSTS, and OCBGF.

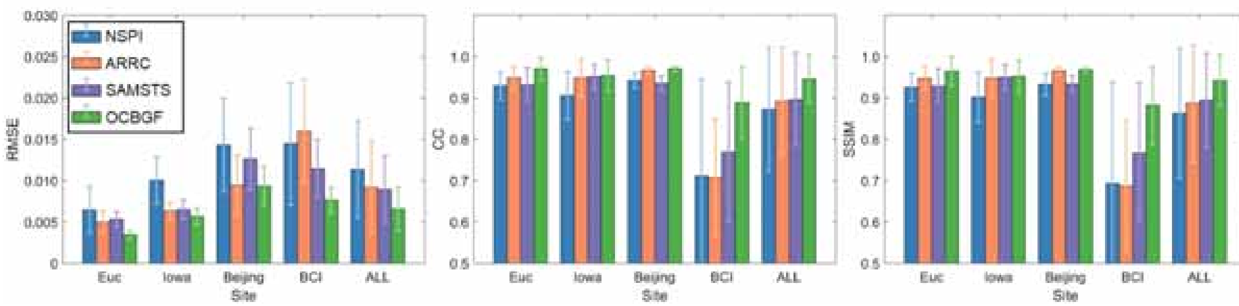

Figure 5. Assessing the effect of different gap sizes (Table 2) on the gap-filled results derived from NSPI, ARRC, SAMSTS, and OCBGF methods using original images as benchmarks. False color composites (RGB=NIR-Red-Green) of selected PlanetScope images and corresponding magnified areas (yellow squares in the first column) of gap-filling results using NSPI, ARRC, SAMSTS, and OCBGF are shown below. Blue circles highlight blocking (a, b, and d) and strip artifacts (c and e) in gap-filled results of NSPI, magenta circles highlight salt-and-pepper noises in gap-filled results of ARRC, and yellow circles highlight blocking artifacts (i) and errors (j and k) in gap-filled results of SAMSTS.

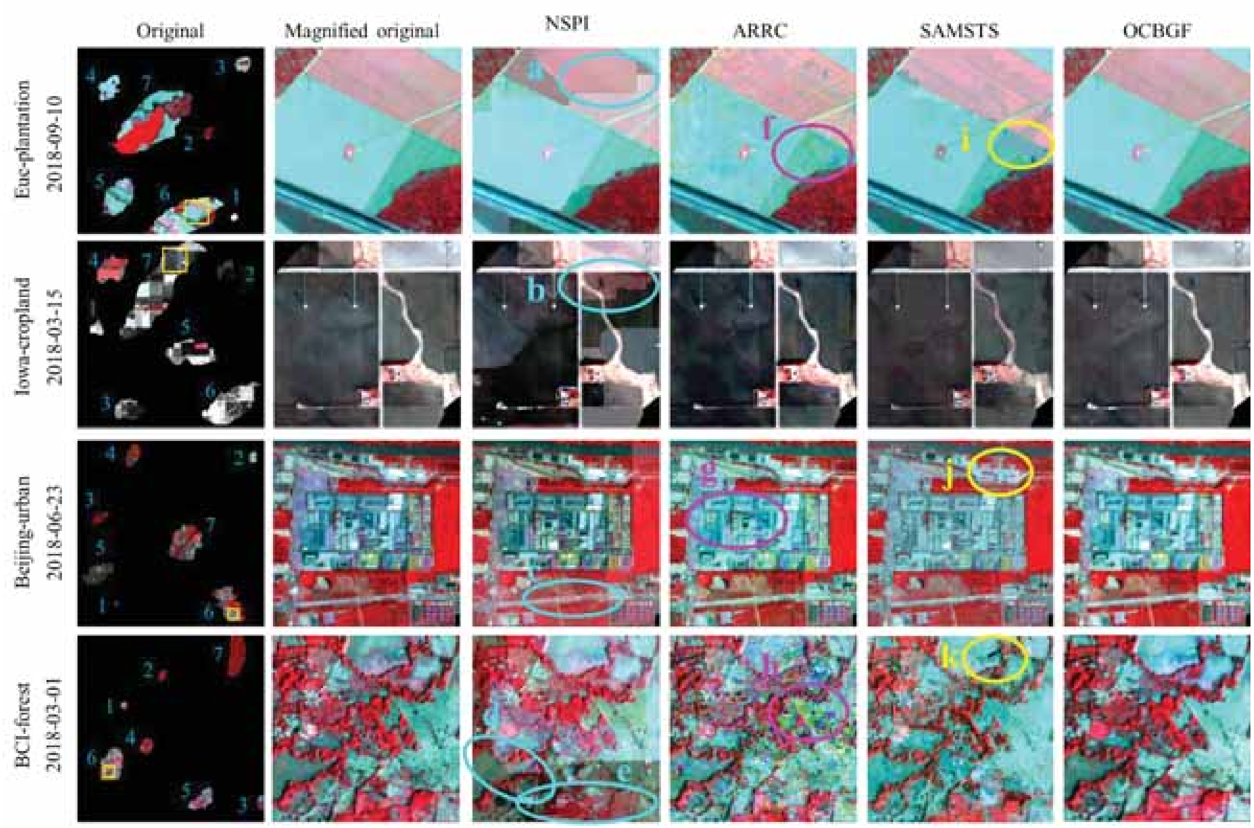

Figure 6. Cross-method comparison of gap-filled results across six gap percentage categories (a) at the Euc-plantation and Iowa-cropland sites using original images (the same as Fig. 5) as benchmarks. These categories have a circle shape with a radius (R) ranging from 500 to 3000 (pixels) with an interval of 500 (pixels), resulting in a percent data missing of 2%, 7%, 16%, 28%, 44%, and 64%, respectively. The results of cross-method comparison include category-specific accuracy (b) and average accuracy of all six categories (c) across the two sites. Three metrics were used for the cross-method comparison, including RMSE, CC, and SSIM. The four methods examined here include NSPI, ARRC, SAMSTS, and OCBGF methods.

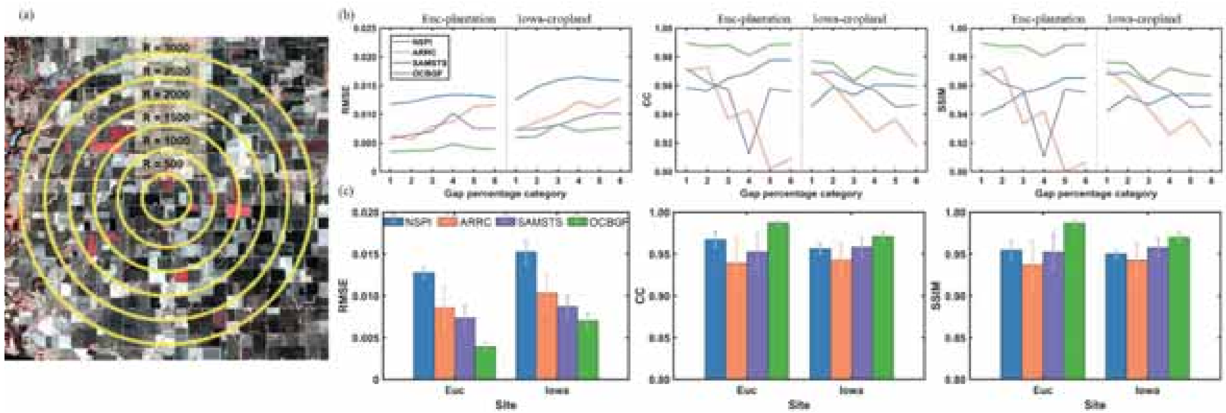

Figure 7. Cross-method comparison of gap-filled results across areas with land cover changes in different seasons at four sites using original images as benchmarks. The results of cross-method comparison include four-band average values of three indices: RMSE, CC, and SSIM across all four sites. The four methods examined here include NSPI, ARRC, SAMSTS, and OCBGF.

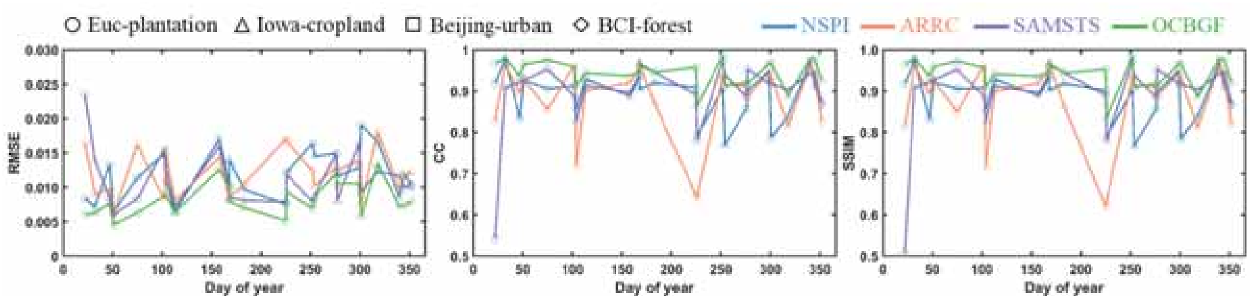

Figure 8. Cross-method comparison of all areas with land cover changes across four sites using original images as benchmarks. Three metrics were used for the cross-method comparison, including RMSE, CC, and SSIM. The four methods examined here include NSPI, ARRC, SAMSTS, and OCBGF.

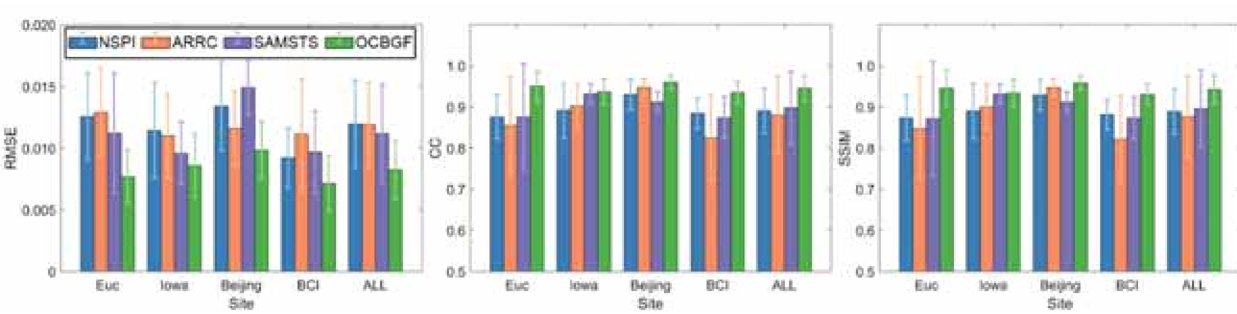

Figure 9. Assessing the effect of various land cover changes on the gap-filled results derived from NSPI, ARRC, SAMSTS, and OCBGF methods using original images as benchmarks. Eight representative land cover types from selected images of the four sites were used. False color composites (RGB=NIR-Red-Green) of selected PlanetScope images (including the date before the target date, the target date (the red frame), and the date after the target date in the first, second, and third column, respectively) and corresponding gap-filling results on the target date (red frame) using NSPI, ARRC, SAMSTS, and OCBGF are shown. Yellow circles highlight the areas experiencing various land cover change scenarios, with (i.e. b, h, and i) and without rapid land cover change (i.e. a, c, d, e, f, and g) over time.

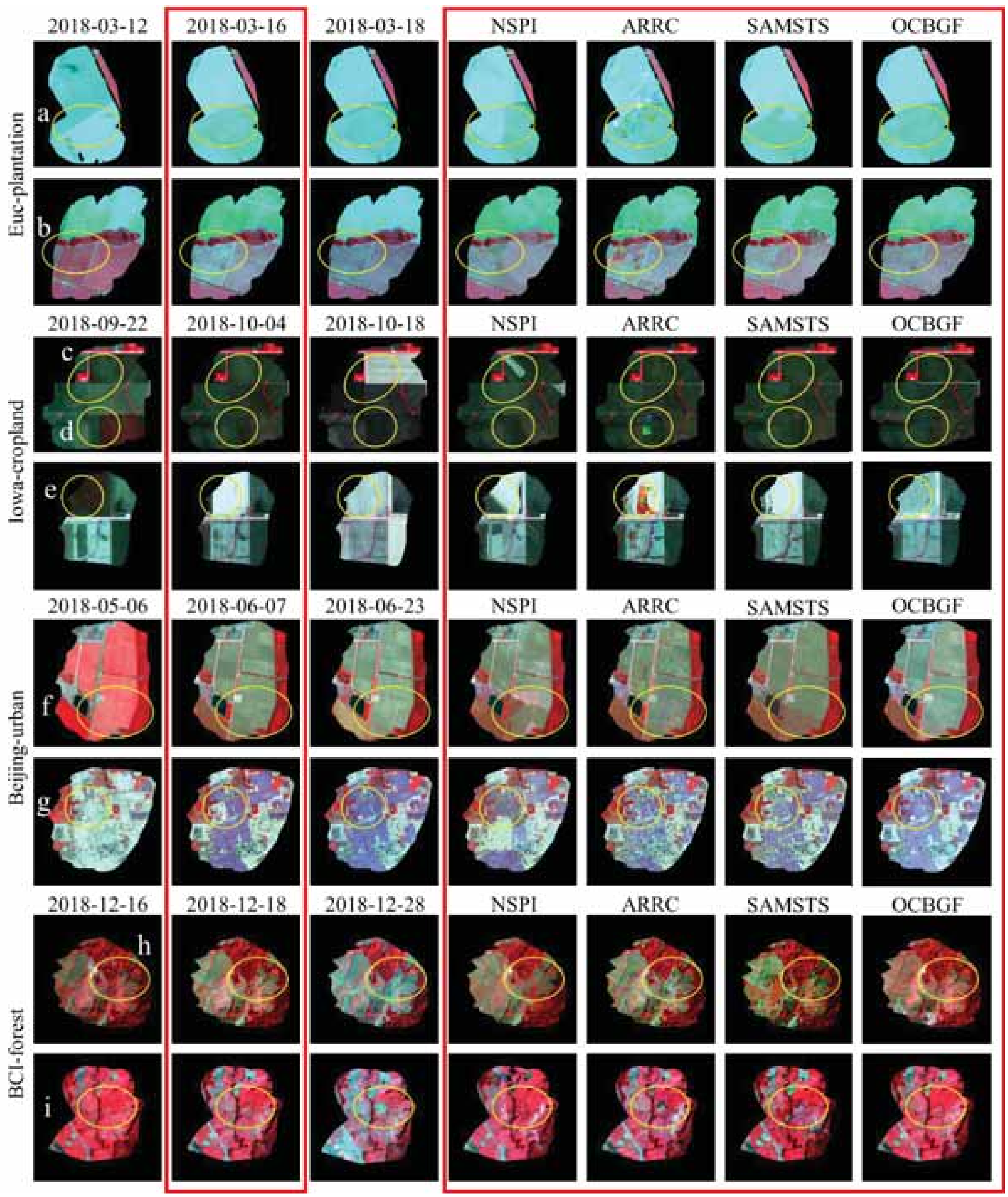

Figure 10. Assessing the performance of gap-filling methods on four real cloud-contaminated images (target image, the second column) at all four sites. Four methods, NSPI, ARRC, SAMSTS, and OCBGF, are examined here. A (nearly) cloud-free image (the first column) close in time to the target image is used as the reference. The cloud masks of target images (the third column) are derived from STI-ACSS method (Wang et al., 2021).

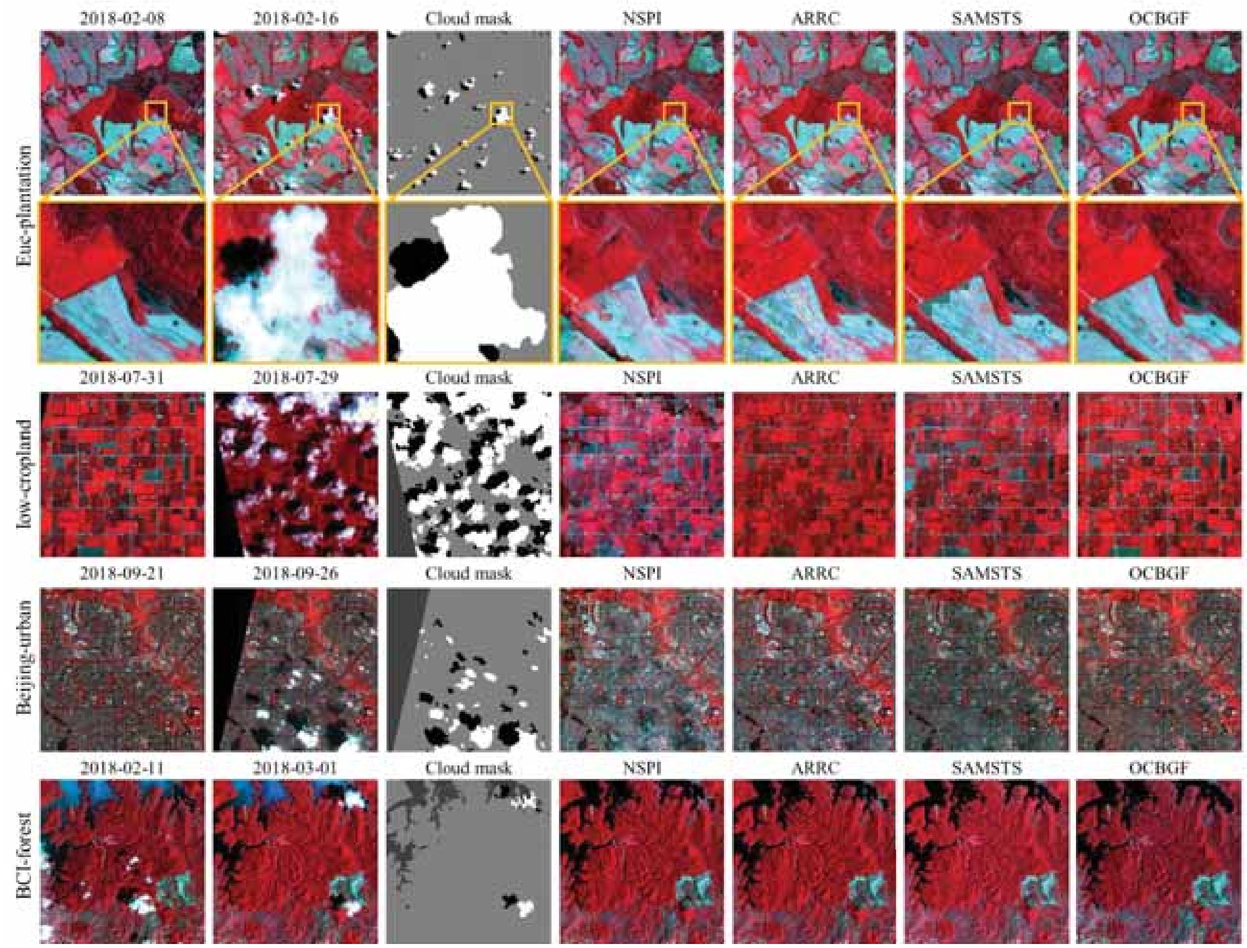

Table 1. Detailed information of the four testing sites, i.e. Euc-plantation, Iowa-cropland, Beijing-urban, and BCI-forest sites, including the location, precipitation, dry season period (< 100 mm of monthly precipitation), spatial coverage, temporal coverage, and the number of accessed PlanetScope images. The precipitation data is assessed from the Tropical Rainfall Measuring Mission (TRMM) data from 2000 to 2019.

Site	Location	Precipitation (mm yr^{-1})	Dry season period	Spatial coverage (km^2)	Temporal coverage	Number of accessed PlanetScope images
Euc-plantation	22°58'S, 48°44'W	1488	Apr-Sep	20×20	Jan-Dec, 2018	129
Iowa-cropland	42°20'N, 92°57'W	1043	NA			52
Beijing-urban	40°01'N, 116°29'E	492	NA			39
BCI-forest	9°06'N, 79°50'W	2052	Jan- Apr			51

Table 2. The number of pixels within each of the seven gap size categories over the four study sites.

Gap size category	Euc-plantation	Iowa-cropland	Beijing-urban	BCI-forest
#1	12,798	43,563	37,565	47,504
#2	27,131	72,346	85,952	78,936
#3	57,988	140,179	169,222	158,086
#4	114,959	181,624	307,798	214,020
#5	251,583	286,482	537,434	381,917
#6	544,319	443,563	815,649	744,642
#7	938,344	937,216	1,101,519	918,274

Table 3. Accuracy assessments (mean and standard deviation) across all seven gap size categories and all four sites using original images as benchmarks. RMSE, CC, and SSIM are used to evaluate the accuracy of gap-filling results respectively derived from NSPI, ARRC, SAMSTS, and OCBGF.

Method	**NSPI**		**ARRC**		**SAMSTS**		**OCBGF**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RMSE	0.0113	0.0058	0.0092	0.0056	0.0090	0.0041	0.0065	0.0027
CC	0.87	0.15	0.89	0.13	0.90	0.11	0.95	0.06
SSIM	0.86	0.16	0.89	0.14	0.89	0.11	0.94	0.06

Table 4. Accuracy assessments (mean and standard deviation) across all the areas with land cover changes and all four sites using original images as benchmarks. RMSE, CC, and SSIM are used to evaluate the accuracy of gap-filling results respectively derived from NSPI, ARRC, SAMSTS, and OCBGF.

Method	NSPI		ARRC		SAMSTS		OCBGF	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RMSE	0.0120	0.0035	0.0119	0.0034	0.112	0.0040	0.0082	0.0023
CC	0.89	0.055	0.88	0.094	0.90	0.087	0.95	0.031
SSIM	0.89	0.055	0.88	0.098	0.90	0.094	0.94	0.035

Figure S1. Example demonstration of the spatial heterogeneity of land cover in nature, including the original image (November 27, 2018 at the Iowa-cropland site), segmentation map, and classification map. The red squares highlight the situation of multiple objects within the same class and the magenta squares highlight multiple classes within the same object.

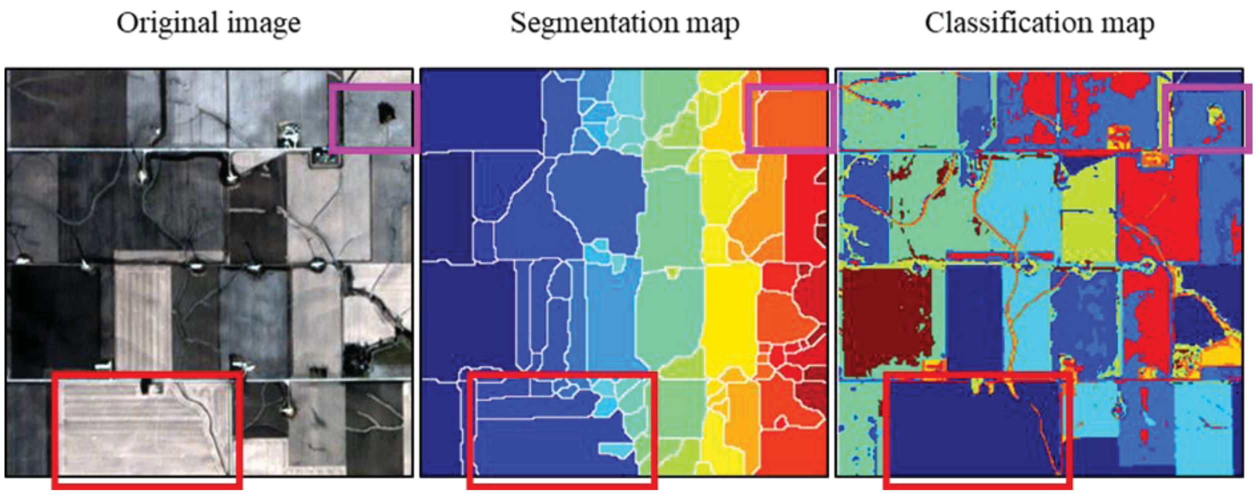

Figure S2. Example demonstration of how the composited cloud-free time series I_{ts} is generated for the automatic object segmentation and classification, which represents Step 1 of Task 2 as shown in Fig. 2. I_{ts} is automatically determined by searching the clearest images (n=2) from each 5-image group before/after the target image I_{t0}.

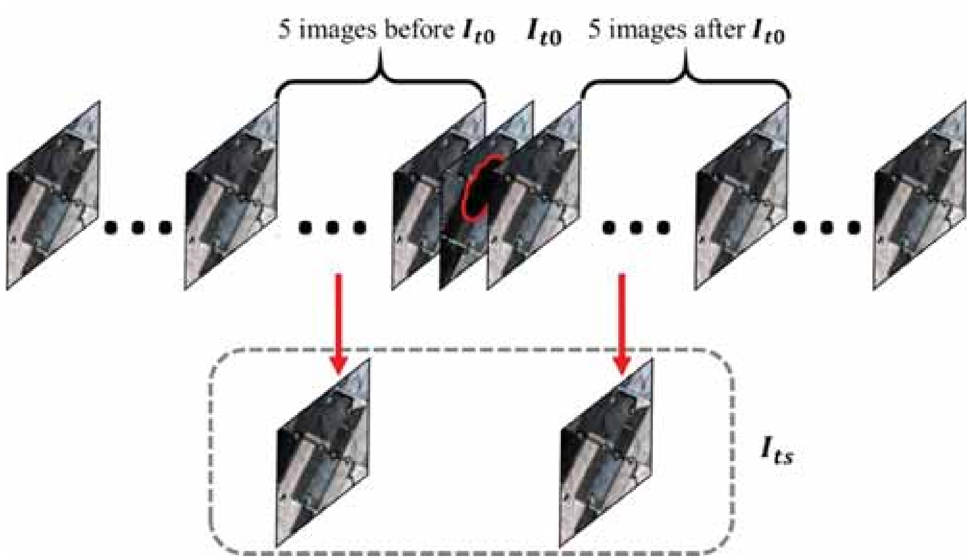

Figure S3 Sensitivity analysis for determining the optimal range of the class number (*K*) using the Iowa-cropland site as an example. Specifically, the five different *K* ranges were set, including 1-5, 3-8, 5-10, 7-12, and 10-15. We used seven testing images acquired on the same dates as Fig. 7 and masked out areas using the simulated mask with a percent data missing of 64% as shown in Fig. 6a.

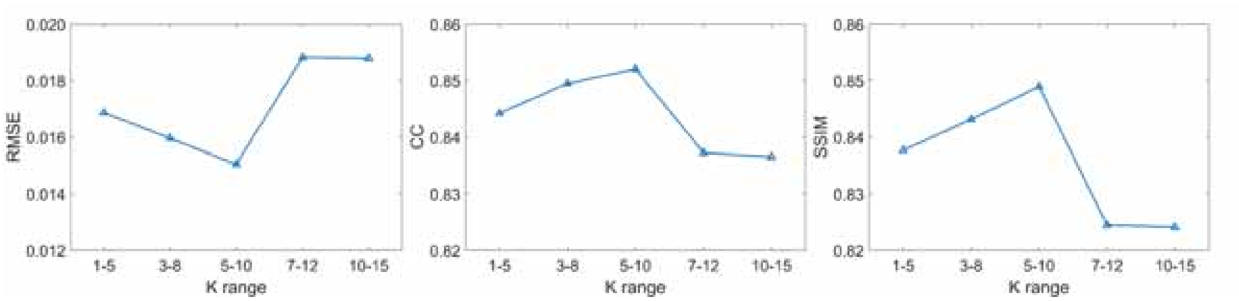

Figure S4. Sensitivity analysis for determining the optimal threshold value that helps best separate the two gap-filling scenarios as shown in Task 3 of Fig. 2. Specifically, the threshold value was set to vary from 0.65 to 0.95, and accuracies were assessed respectively on a site basis (color lines) and across all sites (grey color line). And the two independent scenarios (i.e. S1 and S2 shown in shading areas) were also separately tested. Three metrics for accuracy assessments include RMSE, CC, and SSIM. The data used for this test is the same as the data shown in Fig. 7 in the main text.

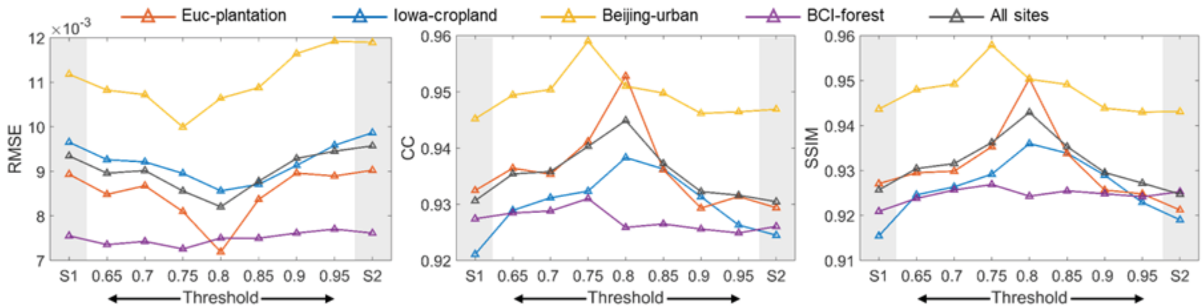

Figure S5. Comparison of gap-filled results across all sites using a fixed (i.e. 0.8) and site-specific optimal threshold value. The overall accuracy is very comparable across three metrics, including RMSE (0.0082 vs. 0.0081), CC (0.94 vs. 0.95), and SSIM (0.94 vs. 0.94).

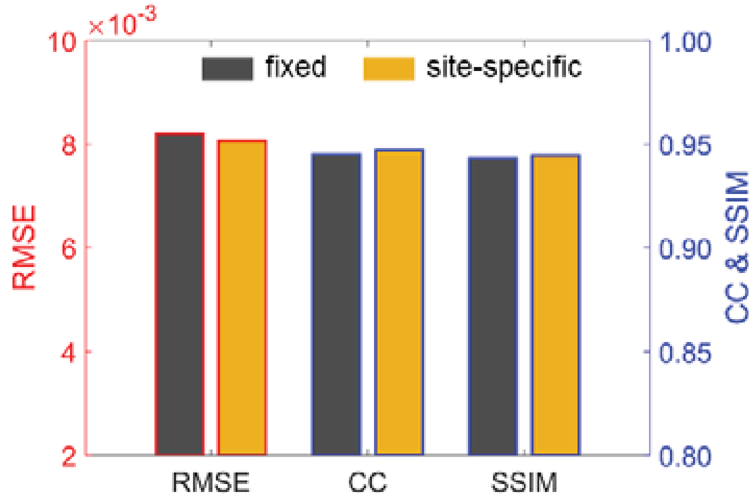

Figure S6. Example demonstrating the effectiveness of the post-image-processing (Task 4; Fig. 2). The original image (a) and the gap-filled results derived before (b) and after applying the guided filter (c) on November 14, 2018 at the Iowa-cropland site are shown in the upper panel. The corresponding magnified areas in the red squares of the upper panel are shown in the lower panel.

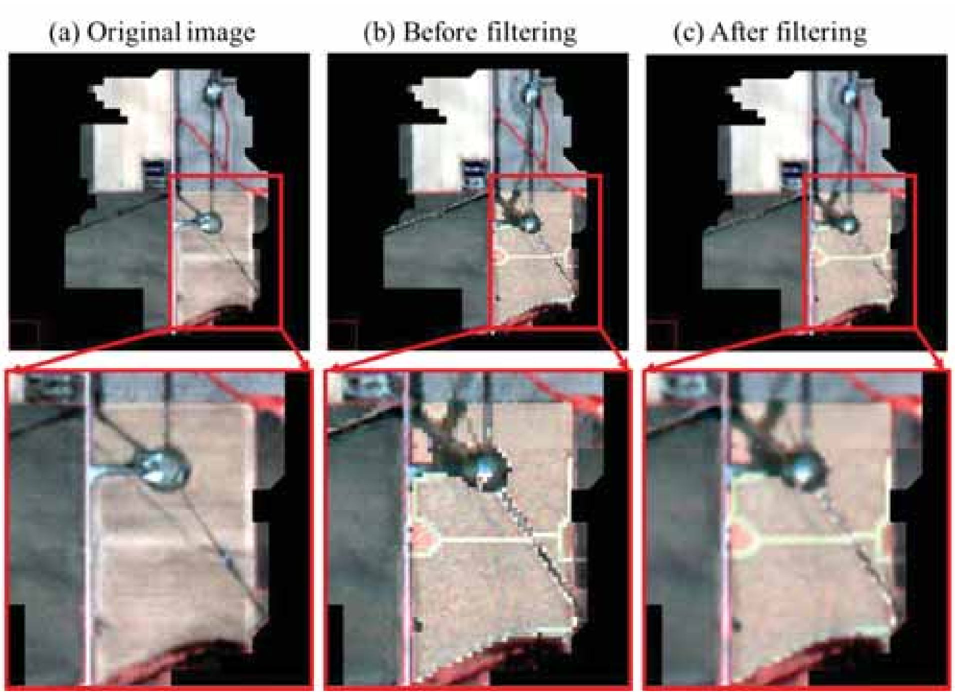

Figure S7. Results assessing the accuracy of gap-filled results derived before and after applying the guided filter, using the Iowa-cropland site as an example. All seven testing images covering areas with land cover changes and across different seasons in 2018 were used here.

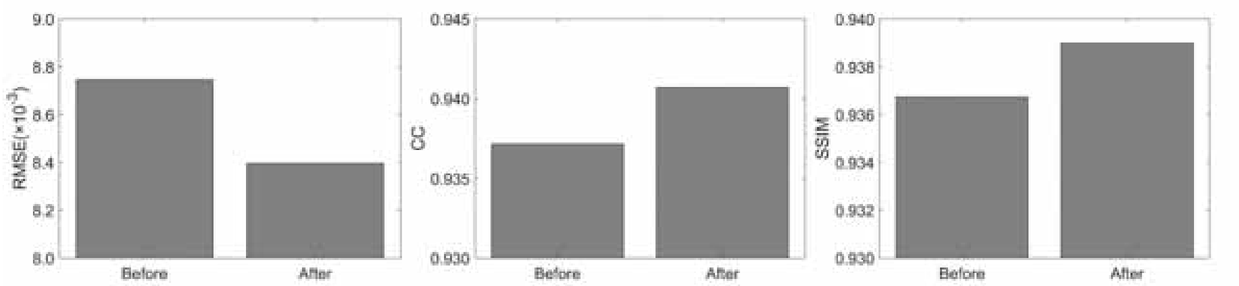

Figure S8. Assessing the effect of different gap sizes (Table 2) on the gap-filled results derived from NSPI, ARRC, SAMSTS, and OCBGF methods using original images as benchmarks. False color composites (RGB=NIR-Red-Green) of selected PlanetScope images and corresponding gap-filling results using NSPI, ARRC, SAMSTS, and OCBGF are shown below. Yellow squares are magnified areas shown in Fig. 5 in the main text.

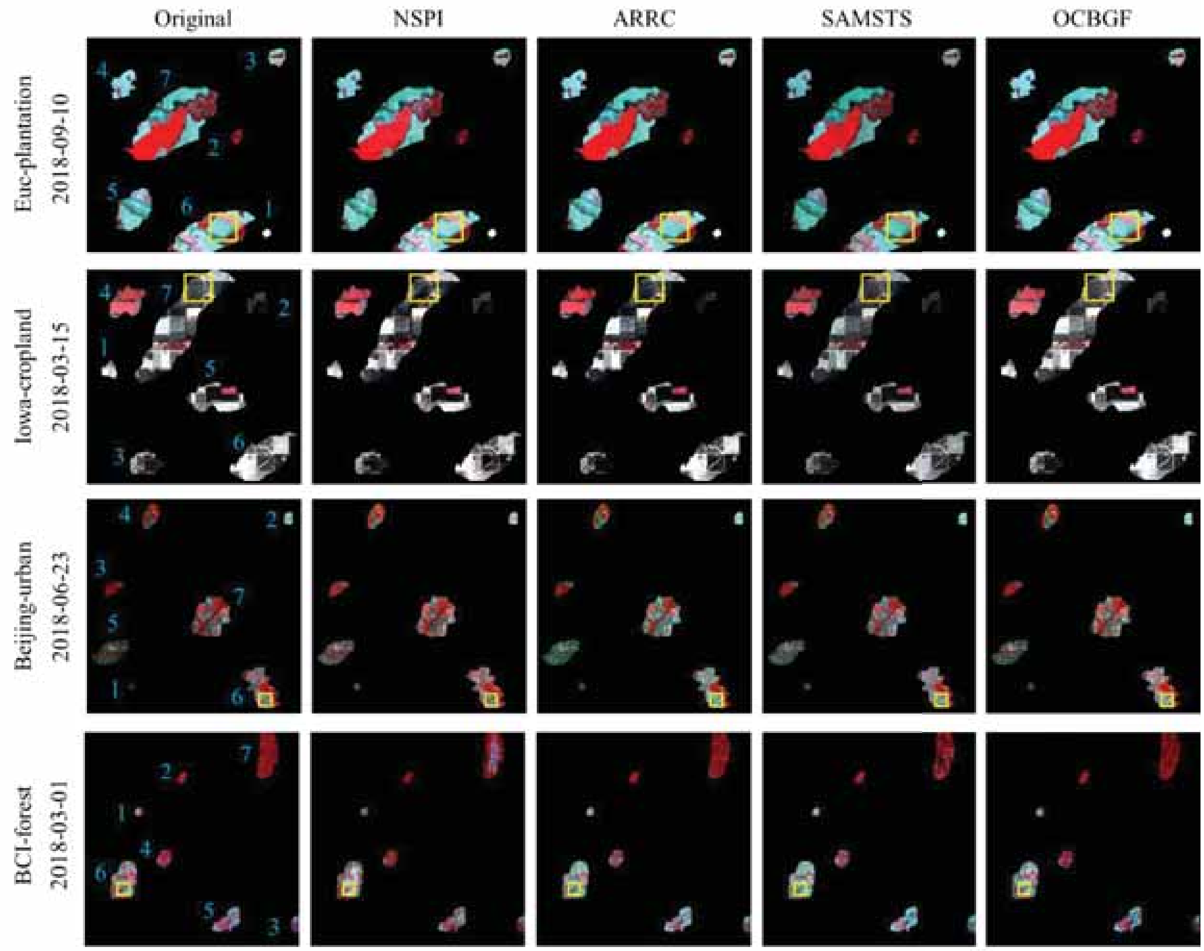

Figure S9. An example demonstration of whether gap-filled PlanetScope time series could boost the patch scale (12m×12m) leaf phenology monitoring in a 50-ha plot of a moist forest landscape at the BCI-forest site. (a) True color composites (RGB=Red-Green-Blue) of drone time-series images (*n*=8; Araujo et al., 2021), (b) false color composites (RGB=NIR-Red-Green) of original PlanetScope (PS) time-series images (*n*=51) with masked clouds/cloud shadows (black areas), and (c) gap-filled PS time-series images using OCBGF. Yellow squares are three representative patches that are affected by serious cloud contaminations. (d) Seasonal variability in patch-scale leaf abundance (gray) derived from the BCI drone images (Park et al. 2019), and enhanced vegetation index (EVI) derived respectively from original (pink) and gap-filled (blue) PS. (e) The histogram and probability distribution function (PDF) of the correlation coefficients (CCs) between the drone-derived leaf abundance seasonality and the two versions of EVI seasonality derived respectively from original and gap-filled PS time-series images, across all patches. The considerably higher average CCs value of gap-filled PS suggests that gap-filled PS could better track leaf phenology at this site compared with drone image time series.

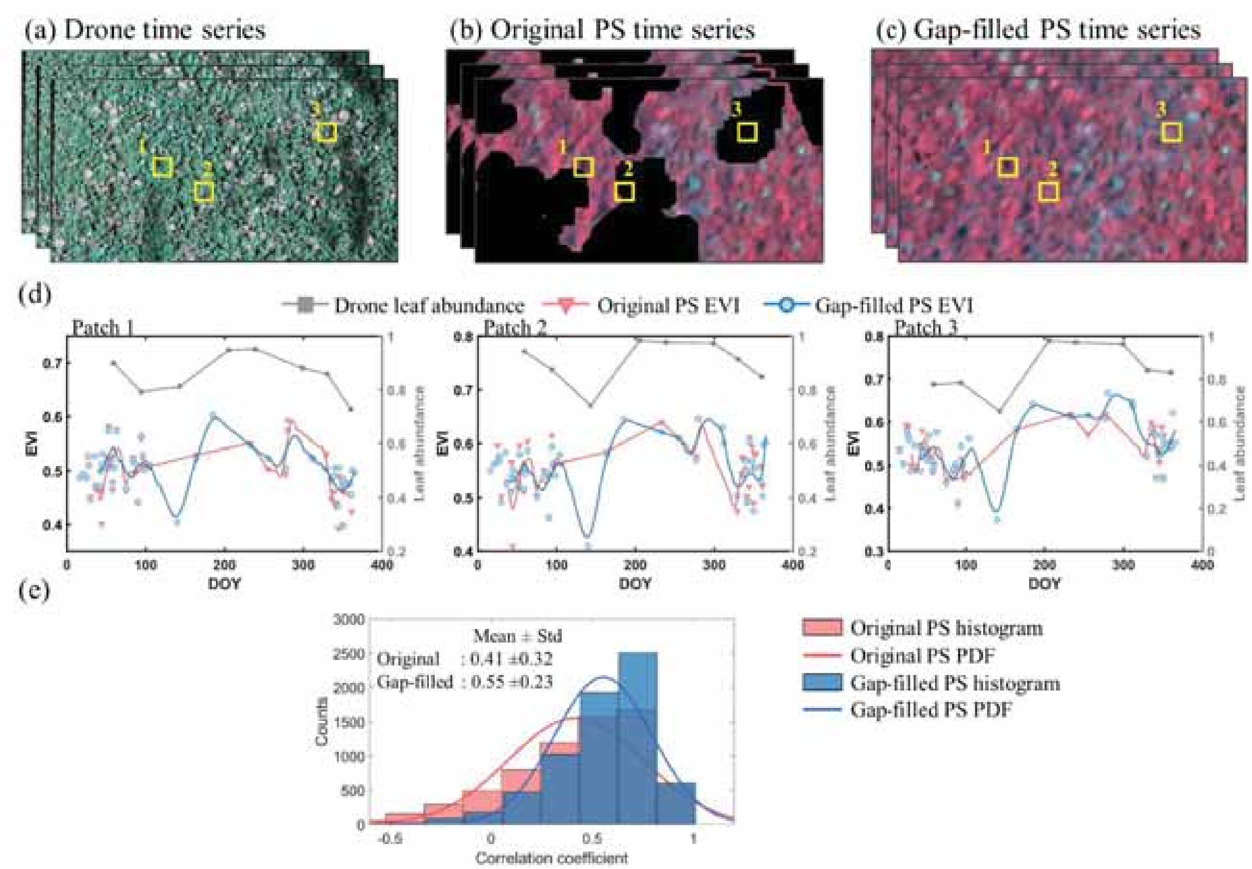

Figure S10. Effectiveness assessments on the three approaches (i.e. object alone, class alone, object-class) for gap-filling at Euc-plantation and Iowa-cropland sites, with three accuracy metrics of RMSE, CC, and SSIM.

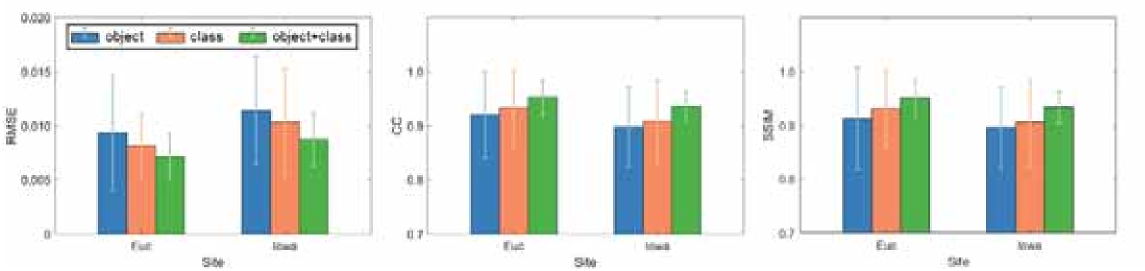

Figure S11. Sensitivity analysis on examining the effect of the frequency of low percent valid data (LPVD; with percent valid data less than 10%) on the gap-filled results, using the Euc-plantation site as an example. The image acquired on September 10, 2018 (the same as the first image in Fig. 5) was selected as the target image for gap-filling, and nine temporally-adjacent images (i.e. five before and four after) were selected from the whole image time series. The sensitivity analysis was performed in the six cases with the frequency of LPVD ranging from 0% to 50% with an interval of 10%.

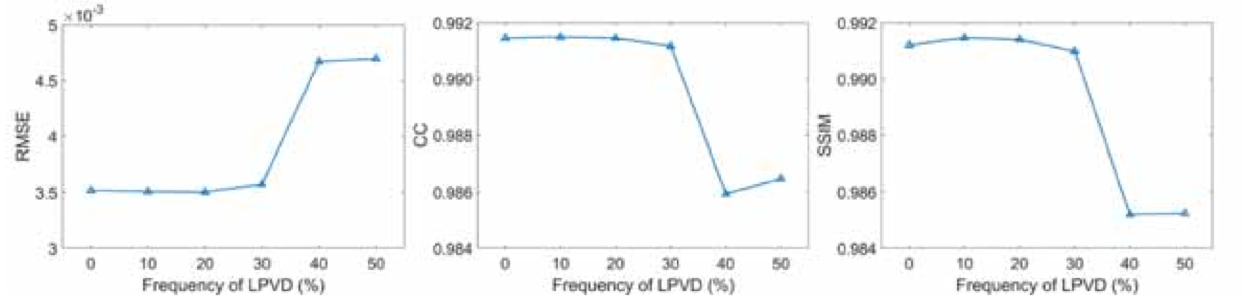

Figure S12. Sensitivity analysis on examining the effect of clear-sky observation frequency on the OCBGF gap-filled results, using the Iowa-cropland site as an example. A total of 35 clear-sky images in 2018 were identified, and the middle (acquired on July 02, 2018) of this time series was selected for gap-filling. The sensitivity analysis was performed in the following 7 cases, which respectively selected once every *N* images throughout the entire image time series, where *N*=1, 2, 3, 4, 6, 8, and 16, resulting in an average time interval of 10, 19, 27, 36, 54, 73, and 146 days, respectively. Our results show that the gap-filled accuracies remain high when the average time interval is less than 19 days, but rapidly decrease afterward with a reduction in clear-sky observation frequency.