Journal: Remote Sensing of Environment

Title: Automatic cloud and cloud shadow detection in tropical areas for PlanetScope satellite images

Author List: Jing Wang[1], Dedi Yang[2], Shuli Chen[3], Xiaolin Zhu[4], Shengbiao Wu[1], Marc Bogonovich[1], Zhengfei Guo[1], Zhe Zhu[5], Jin Wu[1*]

Author Affiliations:
(1) Division of Ecology and Biodiversity, School of Biological Sciences, University of Hong Kong, Hong Kong, China
(2) Department of Ecology and Evolutionary Biology, Stony Brook University, Stony Brook, NY, USA
(3) Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA
(4) Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China
(5) Department of Natural Resources and the Environment, University of Connecticut, Storrs, CT, USA

*** Corresponding Author:** Jin Wu
Division of Ecology and Biodiversity, School of Biological Sciences, University of Hong Kong, Pokfulam Road, Hong Kong (email: jinwu@hku.hk; phone: +852 2299-0655)

Figures/Tables Record:
Number of figures/tables: 8 figures and 3 tables
Supporting materials: 12 figures, 4 tables, and 1 method

Abstract (300 words max)

PlanetScope satellite data with a 3-meter resolution and near-daily global coverage have been increasingly used for land surface monitoring, ranging from land cover change detection to vegetative biophysics characterization and ecological assessments. Similar to other satellite data, effective screening of clouds and cloud shadows in PlanetScope images is a prerequisite for these applications, yet remains challenging as PlanetScope has 1) fewer spectral bands than other satellites hindering the use of traditional methods, and 2) inconsistent radiometric calibration across satellite sensors making the cloud/shadow detection using fixed thresholds unrealistic. To address these challenges, we developed a SpatioTemporal Integration approach for Automatic Cloud and Shadow Screening ('STI-ACSS'), including two steps: (1) generating initial masks of clouds/shadows by integrating both spatial (i.e. cloud/shadow indices of an individual PlanetScope image) and temporal (i.e. reflectance outliers in PlanetScope image time series) information with an adaptive threshold approach; (2) a two-step fine-tuning on these initial masks to derive final masks by integrating morphological processing with an object-based cloud and cloud shadow matching. We tested STI-ACSS at six tropical sites representative of different land cover types (e.g. forest, urban, cropland, savannah, and shrubland). For each site, we evaluated the performance of STI-ACSS with reference to the manual masks of clouds/shadows, and compared it with four state-of-the-art methods, namely Function of mask (Fmask), Automatic Time-Series Analysis (ATSA), Iterative Haze Optimized Transformation (IHOT) and the default PlanetScope quality control layer. Our results show that, across all sites, STI-ACSS 1) has the highest average overall accuracy (98.03%), 2) generates an average producer accuracy of 95.53% for clouds and 89.48% for cloud shadows, and 3) is robust across sites and seasons. These results suggest the effectiveness of using STI-ACSS for cloud/shadow detection for PlanetScope satellites in the tropics, with potential to be extended to other satellite sensors with limited spectral bands.

Keywords: PlanetScope, pixel quality control, cloud and cloud shadow detection, land cover types, tropical areas, ecological and environmental monitoring

1. Introduction

Optical satellite sensors, such as Landsat and Moderate Resolution Imaging Spectroradiometer (MODIS), have long been used for land surface monitoring, ranging from mapping land cover changes (Byrne et al., 1980; Chen et al., 2019a; Friedl et al., 2010) to quantifying ecosystem responses to climate change (Braswell et al., 1997; Jeong et al., 2011; Piao et al., 2006; Yuan et al., 2019). Over recent years, the increasing availability of high spatial and temporal resolution satellites offers an unprecedented opportunity to monitor the rapid and fine-scale dynamics of the Earth's surface. As an example of these new generations of satellites, the PlanetScope constellation of approximately 130 micro satellites (CubeSats) provides a near-daily global coverage at a 3-meter spatial resolution with four spectral bands (i.e. red, green, blue, and near-infrared (NIR)) (Planet, 2020). With such advantages, PlanetScope has been increasingly advocated as a powerful means to advance land surface monitoring, with accumulated many applications, such as disaster monitoring (Ganci et al., 2020; Wang et al., 2019), fine-scale land cover change detection (Halls and Magolan, 2019; Zeng et al., 2018), precision agriculture (Breunig et al., 2020), and tree-crown scale phenology quantifications (Chen et al., 2019b; Wu et al., 2021).

Similar to other commonly-used satellite observations, PlanetScope data are also subject to artifacts associated with the Bidirectional Reflectance Distribution Function (BRDF) effect and cloud/cloud shadow contamination. Consequently, algorithms that can help minimize these artifacts are needed before using PlanetScope data for land surface monitoring. To correct the BRDF effect, an approach that cross-calibrates PlanetScope to BRDF-adjusted MODIS has recently been developed and evaluated in both tropical (Wang et al., 2020) and temperate (Wu et al., 2021) forests, with consistently high accuracy at both MODIS and PlanetScope pixel scales. However, an accurate and automatic approach to detect and mask out clouds/cloud shadows in PlanetScope remains lacking. Although a default quality control layer, i.e. Unusable Data Mask (UDM) (Planet, 2018) and UDM2 (Planet, 2020), has been provided by Planet Lab Inc. to flag candidate PlanetScope pixels contaminated by clouds/cloud shadows, this default layer is mainly used as a cloud coverage filter to aid users to search and order PlanetScope images and remains insufficient for accurate cloud/cloud shadow removal, which is important for subsequent quantitative analysis. For example, after applying the UDM quality control on PlanetScope images in a homogeneous tropical forest landscape, Wang et al. (2020) observed that there remained many unidentified, fragmented clouds/cloud shadows in PlanetScope images, and applied an additional processing that involved labor-intensive visual assessments to carefully mask out the remaining cloud/cloud shadow pixels. Although UDM2 that relies on a supervised machine learning model is provided as an improved quality control layer relative to its earlier version (i.e. UDM) (Planet, 2020), the overall accuracy of cloud and cloud shadow detection using UDM2 has not yet been thoroughly assessed (Cheng et al., 2020), and is not available for dates prior to

August 2018 (Planet, 2020). Therefore, accurate and automatic screening of clouds and cloud shadows remains an essential first step to make full use of PlanetScope data for improved land surface monitoring (Shendryk et al., 2019; Wang et al., 2020).

To minimize the contamination effects from clouds/cloud shadows, many cloud/cloud shadow screening algorithms have been developed for MODIS, Landsat, and Sentinel-2, but remain lacking for PlanetScope. These methods can be grouped into two categories (Zhu et al., 2018): 1) single-image-based methods (Choi and Bindschadler, 2004; Fisher, 2014; Irish et al., 2006; Li et al., 2015; Luo et al., 2008; Roy et al., 2010; Wilson and Oreopoulos, 2013; Zhu and Woodcock, 2012) and 2) multi-temporal-based methods (Chen et al., 2016; Goodwin et al., 2013; Hagolle et al., 2010; Jin et al., 2013; Wang et al., 1999; Zhu and Helmer, 2018; Zhu and Woodcock, 2014). Single-image-based methods mainly rely on certain unique physical characteristics of clouds/cloud shadows, which, together with empirical or adaptive thresholds, can differentiate clouds/cloud shadows from other backgrounds in the same satellite image. These physical characteristics normally include apparent differences between clouds/cloud shadows and clear pixels in their spectral reflectance (Luo et al. 2008), surface temperature (Huang et al., 2010; Irish et al., 2006), and geometric and texture features (Li et al., 2017). For example, with the unique spectral signature of clouds/cloud shadows and more than 20 predefined or adaptive thresholds, Zhu and Woodcock (2012) developed a method (now known as function of mask, Fmask) that can effectively detect clouds/cloud shadows in Landsat Thematic Mapper (TM) and Enhanced TM Plus (ETM+) data and is later adopted to the applications in both Landsat 8 and Sentinel-2 data (Qiu et al., 2019; Zhu et al., 2015). Fmask has now been widely used by the U.S. Geological Survey (USGS) to provide Quality Assessment (QA) for Landsats 4-8 Collection level-1 and level-2 products. Besides these threshold-based methods, there are also some other methods employing machine-learning algorithms such as decision trees (Scaramuzza et al., 2012; Wei et al., 2020), neural networks (Chai et al., 2019; Hughes and Hayes, 2014; Jeppesen et al., 2019; Segal-Rozenhaimer et al., 2020), and support vector machines (Li et al., 2015) to train complex models for automatic cloud and cloud shadow detection.

The second category of cloud/cloud shadow detection algorithms is multi-temporal-based methods that detect clouds/cloud shadows using two or more satellite images of the same region. The central hypothesis underlying these methods is that cloud/cloud shadow pixels have a larger temporal variation in their spectral reflectance or brightness than clear pixels in an image time series, and consequently, accounting for such temporal information could improve the accuracy of cloud/cloud shadow detection. For example, Wang et al. (1999) and Hagolle et al. (2010) independently showed that a pair of two-images (i.e. cloudy vs. cloud-free Landsat/FORMOSAT-2 images) together with a threshold value can help differentiate cloud/cloud shadow pixels from clear ones. Chen et al. (2016) used a clear Landsat image as reference, and developed

an iterative haze optimized transformation (IHOT) for automatic differentiation of cloudy/hazy pixels from clear ones. However, the requirement of a clear reference image constrains the broad applications of these two-image-based methods, as a clear image is not always available. To resolve this issue, Zhu and Woodcock (2014) firstly proposed a multi-Temporal mask (Tmask) based on Landsat image time series to identify clouds and cloud shadows without any predefined clear satellite image as input, though it is often criticized for being computationally expensive. Based on this earlier success, Zhu and Helmer (2018) further developed an Automatic Time-Series Analysis (ATSA) method that automatically screens clouds/cloud shadows with a time-series analysis on a cloud/shadow index. This method circumvents the requirement for a clear reference image and also has greater computational efficiency. Additionally, this multi-temporal-based method has also been shown to have improved accuracy in screening clouds/cloud shadows than the single-image-based method (e.g. Fmask), especially in areas with persistent cloud cover (Zhu and Helmer, 2018).

Despite the fact that these two categories of cloud/cloud shadow screening algorithms have been developed and successfully implemented for MODIS, Landsat, and Sentinel-2, their applicability to PlanetScope data is challenged by two factors. First, most of these algorithms (e.g. Fmask, Tmask, and ATSA) rely on input from more spectral bands, including the shortwave infrared (SWIR) or thermal infrared (TIR) bands that are particularly sensitive to clouds/cloud shadows (Irish et al., 2006; Zhu and Helmer, 2018; Zhu and Woodcock, 2012) but which are absent from the PlanetScope data. Consequently, these algorithms cannot be directly applied to PlanetScope. Second, these algorithms use predefined or adaptive thresholds that are often optimized to the sensors of specific spatial resolution and radiometric characteristics. These default threshold values may not work for PlanetScope due to the fundamental difference in its spatial resolution and potentially also in the radiometric calibration associated with the use of different atmospheric correction methods and radiometric inconsistencies among PlanetScope images acquired by different CubeSats (Houborg and McCabe, 2018). Here we hypothesize that the integration of spatial and temporal information may provide an alternative way for effective cloud/cloud shadow screening, despite there being fewer spectral bands available in PlanetScope data. Similar approaches integrating spatial and temporal information have been developed for Landsat data by previous methods such as Tmask (Zhu and Woodcock, 2014) and ATSA (Zhu and Helmer, 2018). These methods take spatially-derived cloud/shadow indices from single-date images combined with temporally-spectral variations among images in a time series, and are able to effectively screen clouds and cloud shadows, even with fewer input spectral bands (i.e. not including TIR bands) and fewer predefined thresholds. However, whether a similar approach can be adapted to PlanetScope data remains unknown.

To evaluate the feasibility of our hypothesis, we used the tropics as a testbed, for three reasons. First, there is a practical and urgent need, as tropical areas are characterized by frequent precipitation and high cloud cover, resulting in heavy cloud contamination in satellite images (Maeda et al., 2016; Samanta et al., 2010). Second, the tropics are undergoing fast and fine-scale changes in land cover types due to selective logging (Asner et al., 2005), agricultural expansion (Laurance et al., 2014), rapid urbanization (Mohan and Kandya, 2015), and fire disturbance (Brando et al., 2020), and high spatial- and temporal-resolution satellites like PlanetScope (Zeng et al., 2018) is critical for capturing such changes. Third, tropical forests, which account for around one-third of annual terrestrial photosynthesis (Beer et al., 2010), are globally important biomes but are vulnerable to climate change (Aleixo et al., 2019; Phillips et al., 2009), with large potential impacts on global biogeochemical cycles of carbon, water, and energy fluxes and vegetation-mediated climate feedbacks (Fu et al., 2013; Stark et al., 2016; Wright et al., 2017).

The goal of this study is to develop an effective and accurate method for automatic cloud and cloud shadow screening of PlanetScope satellite imagery. Given the challenges of the cloud and cloud shadow detection in PlanetScope, we highlight three major novelties of this work as follows. First, we combined both spatial and temporal information to compensate for the limited number of spectral bands of PlanetScope; second, we developed a novel adaptive threshold approach to overcome the radiometric inconsistencies among PlanetScope time series images acquired by different CubeSats; third, we integrated a two-step fine-tuning, i.e. morphological processing and object-based cloud and cloud shadow matching to reduce random/omission/commission errors that commonly occur in cloud/shadow masks generated for high spatial resolution images. We name this method the SpatioTemporal Integration approach for Automatic Cloud and Shadow Screening (STI-ACSS). By this means, we hope to effectively screen clouds and cloud shadows in PlanetScope imagery and reduce observational uncertainty for fine-scale tropical land-surface monitoring. For this proof-of-concept, we focus on six sites in the cloudy tropics, which include the main tropical land cover types, span a large gradient in annual precipitation, and have a wide range of percent cloud coverage.

2. Testing sites and materials

2.1 Testing sites

To test the proposed STI-ACSS method, we selected six tropical sites that are representative of different land cover types across the three continents in the tropics (Fig. 1). These include 1) a moist forest landscape from the Barro Colorado Island in Panama (henceforth 'BCI-forest'), 2) a metropolitan urban area from

Manaus city in North Brazil (henceforth 'Manaus-urban'), 3) a primary intact moist forest from Tapajos National Forest k67 site in North Brazil (henceforth 'k67-forest'), 4) a managed forest landscape from a Eucalyptus plantation site in South Brazil (henceforth 'Euc-plantation'), 5) a savannah farm grazed by managed cattle in Machakos County, Kenya in Eastern Africa (henceforth 'Machakos-savannah'), and 6) a shrubland landscape from a Mulga woodland in Alice Springs, Australia (henceforth 'Alice Springs Mulga- or ASM-shrubland').

We selected these testing sites for four reasons. First, these six sites represent different land cover types with homogeneous or heterogeneous landscapes and span a large range of annual precipitation from 310-2600 mm per year (Table 1). For example, the k67-forest site is a homogeneous forest site while the Euc-plantation site has mixed land cover types (i.e. commercial Eucalyptus, forest patches, buildings, roads, and dark moist soil).

Second, these sites include bright and dark land surfaces that can be easily confused with clouds and cloud shadows. For example, the BCI-forest site is an island surrounded by coastal areas with sands and has roughly 163 m change in elevation (Knight, 1975) where abundant topographic shadows are formed associated with mountains. The Manaus-urban site represents the largest city in the Amazonas state, Manaus, which contains a large quantity of bright man-made structures, such as airports, highways, and concrete roofs, with many building shadows caused by tall building shelters.

Third, there is modest to strong temporal variability in surface reflectance at these sites. For most multi-temporal-based cloud detection methods, large temporal variations in surface reflectance caused by leaf phenology (i.e. the periodic cycle in leaf flushing, development, senescence, and abscission with consequences on leaf and canopy surface reflectance changes; Chavana-Bryant et al., 2017; Wu et al., 2018) and abrupt land cover change can be easily confused with those arising from the occurrence of clouds and cloud shadows (Zhu and Helmer, 2018; Zhu and Woodcock, 2014). The BCI-forest, k67-forest, and ASM-shrubland sites exhibit modest to large seasonal reflectance variability caused by the different leaf phenology pattern of various vegetation types, such as evergreen and deciduous trees, shrubs, and grass (Detto et al., 2018; Ma et al., 2013; Park et al., 2019; Wang et al., 2020; Wu et al., 2018). The Euc-plantation site experiences strong seasonal reflectance changes caused by harvesting and land cover changes that are associated with deforestation and reforestation (Ceccon and Miramontes, 2008; Lopez-Poma et al., 2020; Qin et al., 2019). The Machakos-savannah site also presents large temporal and spatial variability in reflectance induced by the complex vegetation composition (e.g. woody plants and herbaceous vegetation) of semi-arid rangelands and herbivory through grazing by livestock and wildlife (Cheng et al., 2020).

Fourth, the six sites have been widely used for ecological and environmental research, with great scientific value (Araujo et al., 2002; Cleverly et al., 2016; Comita et al., 2010; Kim et al., 2012; Silva et al., 1995; Vigna et al., 2020; Volkov et al., 2003; Wang et al., 2020; Wu et al., 2016). However, substantial cloud contamination has largely hindered the utility of remote sensing data to monitor the surface dynamics at these and other tropical sites. This makes it particularly important to test our method at these sites.

Details regarding the location, precipitation, dry season duration, and spatial extent of these six sites are shown in Table 1. For more details about the ecological, hydrological, and topographic characteristics of these six sites, refer to previous studies (Campoe et al., 2012; Cleverly et al., 2016; de Souza and Alvala, 2014; Holdridge, 1947; Leigh, 1999; Rice et al., 2004; Vigna et al., 2020; Wu et al., 2016).

2.2 Materials

The four-band PlanetScope data from Planet Lab Inc. (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 Lab Inc. In this study, we downloaded the PlanetScope data for the above-mentioned six testing sites that span both wet and dry seasons with various extents of cloud/cloud shadow covers for each site (Fig. S1). The level 3B surface reflectance product of PlanetScope was used, which has been orthorectified and pre-processed (including geometric, radiometric, and atmospheric corrections) (Planet, 2020). Across the six sites and both seasons in 2018 and 2019, we accessed a total of 822 days of PlanetScope land surface reflectance images (Table 1). In addition to the surface reflectance data, the corresponding default quality control layer (henceforth 'the default QC', i.e. UDM and UDM2) provided by the Planet Lab Inc. was also accessed and was later used for comparison with the developed STI-ACSS method (see Section 3.3 below). Notably, since the BCI-forest and Manaus-urban sites have some water surfaces occupied by lakes and rivers, prior to testing our proposed method, we manually masked out these elements in the corresponding PlanetScope images with the commercial software ENVI 5.3 (Exelis Visual Information Solutions, Boulder, Colorado), by following the same masking procedure as Zhu and Helmer (2018).

3. Methods

The proposed STI-ACSS uses two main steps composed of five tasks (A-E) to achieve an effective screening of clouds and cloud shadows in a cohort of PlanetScope image time series (Fig. 2). In **Step 1**, the initial cloud and shadow masks are produced for each image in the time series by combining both single-image-based (Tasks A and B) and multi-temporal-based (Task C) cloud/cloud shadow screening algorithms.

Tasks A and B respectively rely on an adaptive threshold on a haze optimized transformation (HOT; Zhang et al. 2002) -based cloud index and a flood-fill transformation approach (Soille, 1999) on a shadow index for each single image. Task C is performed with a novel multi-temporal-based cloud and cloud shadow screening algorithm, where we developed an iterative procedure to identify reflectance outliers exceeding a pair of statistical percentiles derived from the entire image time series. In **step 2**, the derived initial cloud and shadow masks are further refined, using morphological processing (Soille, 1999; Zhu et al., 2015; Task D) and object-based cloud and cloud shadow matching (Zhu and Woodcock, 2012; Task E), to remove the "salt-and-pepper noises", as well as to correct under- or over-estimated clouds and cloud shadows. The details for each Task are described below.

3.1 Step 1: generating initial cloud and shadow masks

3.1.1 Single-image-based cloud detection (Task A)

In this section we developed an adaptive threshold to screen out candidate cloud pixels with the HOT index—a commonly-used cloud index for differentiating cloud and clear pixels (Zhang et al., 2002; Zhu and Helmer, 2018; Zhu and Woodcock, 2012). HOT is designed based on the fact that the blue and red bands of clear-sky pixels are highly correlated regardless of land cover types, which tightly follows a linear regression line (synchronically called 'clear line') in the blue vs. red two-dimensional (2-D) space (Zhang et al., 2002). By contrast, pixels contaminated by clouds/haze in this 2-D space display a significant departure from the 'clear-line' (More details see Zhang et al., 2002). Following this idea, the HOT index is calculated as the perpendicular distance of a pixel to the clear line:

$$\mathrm{HOT} = \frac{|a \times B_{blue} - B_{red} + b|}{\sqrt{1+a^2}} \quad (1)$$

where B_{blue} and B_{red} are the blue and red reflectances of a PlanetScope pixel, respectively; a and b are the slope and intercept of the 'clear-line', respectively. To obtain a and b of the 'clear line', we applied a bin-based approach to automatically calculate the HOT index per pixel of any given PlanetScope image (More details see Zhu and Helmer, 2018).

Since the HOT values of cloud pixels are generally greater than clear-sky pixels, we classified an image pixel (x) as a cloud pixel if its HOT value (HOT_x) is greater than a given threshold value (henceforth, 'HOT threshold', T_{HOT}); otherwise, we classified it as a clear pixel. The potential cloud mask (pCM) at pixel x (pCM_x) can be computed as

$$\mathrm{pCM}_x = \begin{cases} 1, & if\ \mathrm{HOT}_x \geq T_{HOT} \\ 0, & if\ \mathrm{HOT}_x < T_{HOT} \end{cases} \quad (2)$$

Here, we proposed an adaptive approach to determine T_{HOT} for each PlanetScope image, considering that it may vary with land cover types, percent cloud cover, and plant phenology. The approach includes four steps:

1) Generate potential HOT thresholds: For this, the HOT indices at 2.5% ($HOT^{2.5}$) and 97.5% ($HOT^{97.5}$) for each PlanetScope image are calculated. In between $HOT^{2.5}$ and $HOT^{97.5}$, a series of potential HOT thresholds ($T_i, i \in [1, N]$) are also generated by equally dividing this range into N (e.g. 50 in this study) intervals.
2) Record the number of pCM pixels (n_i) for each T_{HOT} using T_i according to Eq. 2.
3) Display all pairs of (T_i, n_i) on a 2-D space, resulting in a monotonically decreasing curve (Fig. 3).
4) Determine the optimal HOT threshold: Since the L-shaped "corner" of the above-derived 2-D curve (Hanson, 1992) has been previously shown for the best separation of cloud pixels from those cloud-free pixels (Liu and Liu, 2013), here we automatically determined the L-shaped "corner" to derive the optimal HOT threshold. Specifically, we connected the end point (T_N, n_N) and the start point (T_1, n_1) of the curve and calculated the perpendicular distance to the connected line for each point (T_i, n_i). We then identified the point with the largest distance as the L-shaped "corner" (Fig. 3), and the corresponding threshold value is used as the optimal HOT threshold. It is worthy to note that the way proposed here to identify the L-shaped "corner" is just one solution, and interested readers can refer to Hansen and Oleary (1993) to learn about other solutions as well.

3.1.2 Single-image-based cloud shadow detection (Task B)

To screen out all candidate cloud shadow pixels in a given PlanetScope image, we first calculated a shadow index (SI) and then applied a flood-fill approach (Soille, 1999) on the derived SI image. The SI is derived using Eq. 3, which is similar to Luo et al. (2008).

$$\mathrm{SI} = \sqrt{\left(\frac{B_{red}}{mean_{red}} \times \frac{B_{NIR}}{mean_{NIR}}\right)} \quad (3)$$

Where B_{red} and B_{NIR} are the reflectances of a PlanetScope pixel at red and NIR bands, respectively, and $mean_{red}$ and $mean_{NIR}$ are the mean values of red and NIR bands of the entire image, respectively.

Since cloud shadow pixels typically have much lower SI values compared with their surroundings (Fig. 2), we here applied a commonly-used flood-fill transformation approach (Soille, 1999), which aims to bring

the intensity values of darker areas (with lower SI values) up to the same intensity level as their surrounding lighter pixels (with higher SI values) while removing regional minima that are not connected to the SI image border. With this approach, it would first generate a flood-filled SI image without cloud shadows (i.e. floodfill_SI), and then identifies all candidate cloud shadow pixels as a potential shadow mask (pSM) with an empirical threshold on a SI image difference (the floodfill_SI image - the original SI image). For details about the flood-fill approach, please refer to Li et al. (2017) and Zhu and Woodcock (2012). Consequently, for a given image pixel (x), its pSM value (pSM_x) can be computed as:

$$\mathrm{pSM}_x = \begin{cases} 1, & if\ \mathrm{floodfill\_SI}_x - \mathrm{SI}_x \geq 0.1 \\ 0, & if\ \mathrm{floodfill\_SI}_x - \mathrm{SI}_x < 0.1 \end{cases} \quad (4)$$

3.1.3 Multi-temporal-based cloud and cloud shadow detection (Task C)

To further detect candidate cloud/cloud shadow pixels, we searched 'outliers' with a significant difference in surface reflectance from clear pixels within a cohort of image time series. Specifically, we developed an iterative workflow to repeatedly refine the detected 'outliers' until the cloud and shadow masks stabilize for each image. We initially performed the workflow for each of the four PlanetScope bands individually and then aggregated these band-specific outliers to derive the final outlier layer. The workflow includes the following five steps (Fig. 2).

1) Centralize PlanetScope image time series: To minimize the data inconsistency on cloud and cloud shadow detection associated with inconsistent cross-sensor calibration among different PlanetScope sensors (Houborg and MaCabe, 2018; Wang et al., 2020), we centralized PlanetScope images using the equation $B_{is} - m_{is},\ i \in [1, 4]$, where i is the PlanetScope spectral band, B_{is} is the spectral reflectance of band i of a given image s, and m_{is} is the mean value of band i of image s.
2) Divide PlanetScope image time series into discrete cubic blocks: We divided the image time series into discrete cubic blocks with a user-specified size (e.g. 160×160×M pixels, where M is the number of images in its time series) and calculated the mean and standard deviation for each band (i) of each block (b) as m_{bi} and σ_{bi}, respectively. Users can set the block size with two considerations: (1) block size should be relatively smaller than cloud patches so that the cloud information does not significantly overlap with clear land cover; and (2) block size should encompass a patch size of land surface that represents sufficient spatial variability of clear land cover. In this study, a size of 160×160 pixels (corresponding to a near 500 m × 500 m area) is appropriate according to the above two considerations.
3) Identify candidate outlier blocks: The blocks with high temporal variability in surface reflectance are more likely to be contaminated by clouds/cloud shadows. We thus labelled those blocks with σ_{bi} greater than the average of all blocks ($\bar{\sigma}_{bi}$) as candidate outlier blocks.

4) Detect outliers with a pair of percentile thresholds: As clouds/cloud shadows often have much higher/lower reflectance in visible and NIR bands than other backgrounds, we expected that clouds/cloud shadows are located in the upper/lower ends of the reflectance distribution of each spectral band. With this, a pair of thresholds of 5 and 95 percentiles (henceforth 'outlier threshold', see Section 3.4 for the determination of this threshold pair) were introduced and applied to each candidate outlier block, by which we screened out cloud/cloud shadow pixels for each candidate outlier block.
5) Iterate the above four steps to optimize outlier detection: For each iteration, we calculated the difference of relative standard deviation of the image time series (i.e. $\nabla\frac{\bar{\sigma}_i}{\bar{m}_i}$, where $\bar{m}_i$ and $\bar{\sigma}_i$ are the mean and standard deviation of spectral band *i* of the time series) before and after the removal of the labelled cloud and shadow outliers, and repeated the above four steps until the relative difference (i.e. $\nabla\frac{\bar{\sigma}_i}{\bar{m}_i}$) is smaller than 0.01 for each band. After several iterations, $\nabla\frac{\bar{\sigma}_i}{\bar{m}_i}$ becomes stable (<0.01), and all potential cloud and shadow outliers for each given band *i* (pCSO_i) are derived. To be conservative in labelling all potential cloud/cloud shadow pixels in the image time series, we unioned the band-specific pCSO_i to generate the final potential cloud and shadow outliers (pCSO):

$$\mathrm{pCSO} = \bigcup_{i=1}^{4} \mathrm{pCSO}_i \quad (5)$$

Finally, we intersected all the candidate cloud/cloud shadow pixels detected respectively by single-image-based methods (i.e. pCM for Task A and pSM for Task B) and multi-temporal-based method (pCSO for Task C) to derive the initial cloud (iCM; Eq. 6) and shadow (iSM; Eq. 7) masks, following the assumption that the cloud/cloud shadow pixels simultaneously detected by these two methods would help reduce the omission and commission errors for the cloud/cloud shadow detection.

$$\mathrm{iCM} = \mathrm{pCM} \cap \mathrm{pCSO} \quad (6)$$

$$\mathrm{iSM} = \mathrm{pSM} \cap \mathrm{pCSO} \quad (7)$$

3.2 Step 2: refining initial cloud and shadow masks

3.2.1 Fine-tuning on the cloud and shadow masks using morphological processing (Task D)

Although the initial cloud and shadow masks derived from **Step 1** can assist in identifying the most likely cloud and cloud shadow pixels, two main issues remain, including 1) the "salt-and-pepper noises" and 2) the underestimation of the low-density border areas surrounding clouds/cloud shadows (Li et al., 2017; Zhu et al., 2015). To address these issues, for each PlanetScope image, we integrated morphological opening,

closing, and dilation with a structure element on the initial cloud and shadow masks to remove "salt-and-pepper noises" and to include the cloud and cloud shadow border areas (Soille, 1999; Soille and Pesaresi, 2002). More specifically, these morphological operations use a moving structure element (i.e. a shape including both the target pixel and its neighbors) to probe and interact with the input image (Gonzalez and Woods, 2006), with different purposes. Morphological opening and closing are useful for removing small objects and filling small holes and gaps, while preserving the shape and size of larger objects in the input image. Morphological dilation is useful for expanding the shapes and smoothing boundary of objects in the input image. Details regarding these morphological operations are also shown in Eqs. 8-9 below.

$$\mathrm{rCM} = (\mathrm{iCM} \circ \mathrm{E}) \cdot \mathrm{E}) \oplus \mathrm{E} \tag{8}$$

$$\mathrm{rSM} = (\mathrm{iSM} \circ \mathrm{E}) \cdot \mathrm{E}) \oplus \mathrm{E} \tag{9}$$

Where iCM and iSM are the initial cloud and shadow masks, respectively; rCM and rSM are the refined cloud and shadow masks, respectively; $\circ$, $\cdot$ and $\oplus$ are the morphological operations of opening, closing, and dilation, respectively; and E is the structure element of morphological processing. The structure element is a disk-shaped matrix (Zheng, 1995) with a user-specified size (e.g. 7×7 pixels in this study).

3.2.2 Fine-tuning on the cloud and shadow masks using object-based cloud and cloud shadow matching (Task E)

After Task D, some commission errors remain, particularly associated with dark or bright surface objects (e.g. moist soil and roads; Fig. 2). To further reduce these errors, we applied an object-based cloud and cloud shadow matching approach (Zhu and Woodcock, 2012; Li et al., 2017), which utilizes the cloud-cloud shadow linkage whereby clouds and their corresponding shadows usually appear concomitantly in an image with similar geometric shapes. More specifically, we first calculated the projection direction of the cloud shadows relative to their corresponding clouds by using the geometric position of the sun (Zhu and Woodcock, 2012). We then calculated the best matching distance between cloud and cloud shadows by iteratively moving the cloud mask along the projection direction until a maximum overlap area is achieved between the moved cloud mask and original cloud shadow mask. Lastly, we took an additional step as used in Fmask to exclude those cloud pixels without matched cloud shadows and vice versa. For more details, see Zhu and Woodcock (2012).

It is worthy to note that there are two scenarios where the object-based cloud and cloud shadow matching approach is not applicable because clouds do not always have corresponding shadows in the same image. These include 1) not all clouds have corresponding shadows, including but not limited to thin clouds/haze,

and 2) the projected cloud/cloud shadow objects are partly at the edges of the image or fully outside the image. In order to minimize the uncertainty associated with the first scenario, we implemented an additional process as follows: for any given PlanetScope image, when the size of cloud objects is far larger than the corresponding shadow objects (e.g. more than double) and the ratio of the overlapping area of all matching cloud and cloud shadow objects to the area of all cloud objects is less than 0.5, we tended to keep the refined cloud and shadow masks generated by Task D. To address the second issue, we first determined the range of the image edge on *x* and *y* axes based on the best matching distance between clouds and cloud shadows, and kept those objects with their projected clouds/shadows partly or fully outside the image.

3.3 Method evaluation and cross-method comparison

We here adopted two approaches for method evaluation and cross-method comparisons. First, we selected 47 representative PlanetScope images, carefully generated manual masks of clouds and cloud shadows for each selected image using ENVI 5.3 (Exelis Visual Information Solutions, Boulder, Colorado), and used these manual masks for quantitative accuracy assessments across all methods/products, including STI-ACSS (our method), the default PlanetScope QCs (i.e. UDM and UDM2), Fmask (Zhu and Woodcock, 2012; 2015), IHOT (Chen et al., 2016), and ATSA (Zhu and Helmer, 2018). Specifically, these 47 representative images cover all tropical sites and span a large gradient in percent cloud cover (0%-96%), including 9 from BCI-forest, 7 from Manaus-urban, 7 from k67-forest, 10 from Euc-plantation, 7 from Machakos-savannah, and 7 from ASM-shrubland. We used five accuracy metrics for assessments, namely overall accuracy (OA), user's accuracy (UA, equal to 100%-percent commission error) of shadows and clouds, respectively, and producer's accuracy (PA, equal to 100%-percent omission error) of shadows and clouds, respectively (Foody, 2002). Second, since STI-ACSS and the default QCs are both developed specifically for the high spatial and temporal resolution PlanetScope data with fewer spectral bands, we used the default QCs (i.e. UDM and UDM2), particularly the more developed UDM2 (available since August 2018), as a benchmark to cross-evaluate STI-ACSS and other methods (i.e. Fmask, IHOT, and ATSA). This step helps evaluate the temporal consistency/inconsistency among different methods across the full image time series.

It is noted that 1) the default QCs are available from https://www.planet.com/; 2) Fmask and ATSA methods were originally designed for multispectral data such as Landsat and Sentinel-2, and thus some rules or filters associated with SWIR or TIR bands are not applicable to the four-band PlanetScope data; 3) IHOT was developed for detecting hazes and clouds, not for cloud shadows. With these considerations, to make the cross-method comparison meaningful and rigorous, we carefully adjusted the parameters of Fmask, ATSA, and IHOT for each site to achieve their best performance, following Chen et al. (2016), Zhu and Helmer

(2018), and Zhu and Woodcock (2012). The results of these parameters are shown in Table S1, while the details regarding how we tuned these methods to derive site-specific optimal parameters are shown in Supplementary Method 1. Additionally, ATSA was conducted in Interactive Data Language (IDL 8.5, Exelis Visual Information Solutions, Boulder, Colorado), while the other three methods (Fmask, IHOT, and STI-ACSS) were conducted in Matlab R2019a (MathWorks, Natick, MA, USA).

3.4 Sensitivity analysis: Assessing the effects of the outlier thresholds in Task C on the derived cloud and shadow masks using STI-ACSS

The outlier thresholds in Task C, which are used to determine reflectance outliers in the multi-temporal-based method, are critical for deriving the final cloud and shadow masks. Here we assessed the effects of different thresholds on the final results with the Euc-plantation site as an example, including the PlanetScope image time series in 2018 (n=35). Specifically, we used the manual masks of clouds and shadows as reference and evaluated the results that were respectively derived using the following five thresholds pairs, i.e. 1) the 1st and 99th percentiles, 2) 2.5th and 97.5th percentiles, 3) 5th and 95th percentiles, 4) 7.5th and 92.5th percentiles, and 5) 10th and 90th percentiles. Three representative dates covering both wet and dry seasons were selected for generating manual masks and cross-scenario model performance assessments, including February 22, July 19, and November 13 of 2018. The sensitivity analysis suggests that the different threshold pairs tend to generate comparable model performance with reference to the manual masks (Fig. S2), despite that the threshold pair of the 5th and 95th percentiles obtained a slightly higher OA than other threshold pairs (Table S2). We thus used the threshold pair of the 5th and 95th percentiles in our STI-ACSS.

4. Results

4.1 Cross-method comparison across representative images

To evaluate the accuracy of all methods (STI-ACSS, the default QCs including UDM and UDM2, Fmask, IHOT, and ATSA), we conducted accuracy assessments on 47 representative images using manual masks as benchmarks. STI-ACSS yields a consistently high level of accuracy (OA>95%) across all images covering all sites and seasons (Figs. 4a and S3), while other comparative methods generate large variations in OA. On average, STI-ACSS obtains the highest OAs and lowest standard deviations across all six sites (Fig. 4b), with an OA of 98.03% (mean) ± 1.30% (standard deviation) across all images, followed by UDM2, ATSA, Fmask, UDM, and IHOT (Table 2). Additionally, we quantified the performance improvement of STI-ACSS from Step 1 (including Tasks A-C; Section 3.1) to Step 1+Step 2 (including Tasks D-E; Section

3.2). Our results demonstrate an average OA of 92.31% when implementing Step 1 alone, with an additional average increase by 5.72% in OA when also implementing Step 2 (Fig. 5). Taken together, these results demonstrate that: 1) STI-ACSS is the most accurate among all comparative methods and is robust across all land cover types and a wide range of percent cloud cover; 2) UDM2 obtains a significantly improved OA in comparison to UDM on average (an OA increase of 8.30%; Table 2); 3) Fmask and ATSA obtain higher OA values than UDM; and 4) IHOT obtains the lowest OA partly due to the lack of a cloud shadow detection module.

STI-ACSS also performs the best in comparison to other methods when using the PA (100%-percent omission error) and UA (100%-percent commission error) of cloud/cloud shadow detection as evaluation statistics, with the highest average PAs (95.53% for clouds and 89.48% for shadows; Table 2) and highest average UAs (93.70% for clouds and 91.55% for shadows; Table 2). In contrast, the default PlanetScope UDM has much larger omission and commission errors in the cloud and cloud shadow detection as revealed respectively by much lower PAs (56.67%±27.05% for clouds and 2.11%±3.68% for shadows; Table 2) and a very low UA of shadows (18.37%±27.80%; Table 2). Compared with UDM, UDM2 has significant improvements in both PAs (a PA increase of 26.97% for clouds and 27.94% for shadows, respectively; Table 2) and the shadow UA (an increase of 72.06%; Table 2), but is still insufficient for the cloud shadow detection (shadow PA=30.05%; Table 2). Relative to UDM2, Fmask, and ATSA have higher PAs of clouds and shadows with a cost of reducing UAs of clouds and shadows (Table 2). Compared with all other methods, IHOT has the lowest UA of clouds with the largest variation across all images covering all sites and seasons (51.20%±33.06%; Table 2). Collectively, these results again demonstrate that STI-ACSS performs the best in screening clouds and cloud shadows regardless of the accuracy metrics used, with high consistency across OA, PAs of clouds and cloud shadows, and UAs of clouds and cloud shadows.

To aid visual interpretation of these cross-method comparisons, we next presented 12 representative images across six sites spanning a large gradient in percent cloud cover (2%-86%) (Fig. 6). These results clearly show that STI-ACSS generates the highest accuracies in screening clouds and cloud shadows compared with all other methods. Meanwhile, STI-ACSS is not 100% accurate. For example, in certain cases, small portions of moist plantation are misclassified as cloud shadows while some thin clouds/cloud shadows are often omitted (e.g. shadow PA=87.19% and shadow UA=81.42% on March 12, 2018 of the Euc-plantation site, and cloud PA=78.40% on September 05, 2018 of the ASM-shrubland site; Table 3). Compared with STI-ACSS, UDM generates much larger omission and commission errors in cloud/cloud shadow detection. For example, the large omission errors are consistently observed across all sites using UDM, leading to much lower PAs of clouds and cloud shadows (cloud PA<86.33% and shadow PA<7%; Table 3). Meanwhile, for a certain case, e.g. at the Euc-plantation site, it misclassifies a large part of plantation as

cloud shadows (shadow UA<5%; Table 3). Compared with UDM, UDM2 has significant improvements in its accuracies, but still omits some clouds and cloud shadows, especially at the ASM-shrubland site (cloud PA<79.17% and shadow PA<34.32%; Table 3). There are also large omission and commission errors when using Fmask for cloud/cloud shadow detection. For example, Fmask omits some cloud shadows located at the image edges (e.g. the k67-forest site), and sometimes erroneously detects cloud shadows at an incorrect distance from their corresponding clouds (e.g. December 13, 2019 of the BCI-forest site and March 12, 2018 of the Euc-plantation sites), resulting in a low PA of shadows (<68.08%; Table 3). Additionally, Fmask misclassifies a significant amount of bright buildings/bare soil/building shadows/tree shadows as clouds/cloud shadows on June 02, 2018 of the Manaus-urban and September 05, 2018 of the ASM-shrubland sites, leading to lower UA values (shadow UA<45.51% and cloud UA<33.99%; Table 3). Similar to Fmask, ATSA also introduces some omission errors in cloud shadow detection and commission errors in cloud detection. For example, ATSA omits some cloud shadows on March 12, 2018 of the Euc-plantation site and September 05, 2018 of the ASM-shrubland site (shadow PA<24.69%), and also overestimates some bright buildings/branches/building shadows/canopy gaps as clouds/cloud shadows, especially for June 02, 2018 of the Manaus-urban site and September 12, 2019 of k67-forest site (shadow UA< 52.72% and cloud UA<78.34%; Table 3). Last, IHOT is observed with large omission errors in cloud detection, and for some cases, e.g. on August 18, 2019 of the Manaus-urban site, it omits a large part of clouds in the image (cloud PA=60.53%; Table 3). Collectively, these results demonstrate that STI-ACSS performs the best while UDM and IHOT perform the worst in cloud/cloud shadow detection across the representative images shown here (Fig. 6).

Finally, we also observed that traditional methods (i.e. Fmask, IHOT, and ATSA) often misclassify the whole image as clouds (Fig. 6), resulting in a low cloud UA (<70%; Table 3). In particular, IHOT introduces the largest commission errors in cloud detection at the BCI-forest, k67-forest, and Machakos-savannah sites (Fig. 6 and Table 3). A likely explanation for this is that IHOT uses an iterative model to derive a 'clear-line' (i.e. a fitted line for clear pixels in the 2-D space of red vs. blue bands; Fig. S4) for separating cloud pixels from those cloud-free ones based on their distances to the 'clear-line', while for the BCI-forest, k67-forest, and Machakos-savannah sites, both cloud and cloud-free pixels follow a similar linear relationship in the red-blue space with a slope close to 1 (e.g. 0.998-1.187; Fig. S4), resulting in low separability in these cases. Meanwhile, we observed that Fmask, IHOT, and ATSA tend to overestimate clouds in some cases, e.g. December 28, 2018 of the BCI-forest site, March 22, 2019 of the Euc-plantation site, and February 12, 2018 of the Machakos-savannah site (Fig. 6; Table 3). This might be related to the inconsistent cross-sensor calibration among different PlanetScope sensors, and as a result, there is a significant difference in the distribution of HOT values between two PlanetScope images adjacent in time (including the range and median of HOT values; Fig. S5). Also because of this, Fmask, IHOT, and ATSA, which rely

on either predefined or statistical HOT thresholds, are observed to work well for some PlanetScope images but not all (Fig. 6). Moreover, we observed that STI-ACSS, Fmask, and ATSA are all more effective at detecting clouds than IHOT (Fig. 6). This is because these methods introduce a virtual 'clear line' with a larger slope than 1 in the red-blue space (e.g. slope=2 for Fmask and STI-ACSS, and slope=1.5 for ATSA), which increases the separability between cloud and cloud-free pixels. Lastly, we observed that STI-ACSS has the highest accuracy in cloud detection compared with all other methods (Fig. 6). This is largely because STI-ACSS adopted an adaptive HOT threshold rather than a fixed or statistical threshold, and consequently, it helps to address the issue caused by inconsistent cross-sensor calibration among different PlanetScope sensors (Figs. S6-S8).

4.2 Cross-method comparison across time series images

To examine the temporal consistency of our STI-ACSS and the three other methods (Fmask, IHOT, and ATSA) across the full image time series in 2018 and 2019, we used the default PlanetScope QC layer (particularly UDM2) as references. As shown in Fig. 7, across all testing sites and image time series, STI-ACSS almost consistently yields the highest overall agreements with the default QCs among all methods, with the highest average OA of 92.49% and the smallest OA standard deviation of 7.70%.

4.3 The effect of seasonal and interannual reflectance changes on the STI-ASS-derived cloud and shadow masks

To evaluate how temporal (seasonal and interannual) reflectance changes impact the results of STI-ACSS, we used the full time series in 2018 and 2019 (n=24 months) at the Euc-plantation site as an example, as temporal reflectance changes at this site are considerable (see the colored circles of Fig. 8 as examples). Overall, our results show that STI-ACSS accurately detects clouds and cloud shadows across these images, without any significant commissions that misclassify those clear pixels with large temporal reflectance variations caused by plant phenology (e.g. leaf flushing, development, senescence, abscission, flowering, and fruiting) or land cover changes as clouds/cloud shadows, particularly for the areas highlighted with colored circles. Meanwhile, we also observed that STI-ACSS occasionally tends to misclassify clouds/cloud shadows. For example, as shown in those areas highlighted with red squares in Fig. 8, some small areas of bright surfaces/bare soil/moist plantation/wet soil are misclassified as clouds/cloud shadows.

5. Discussion

The high spatial and temporal resolution PlanetScope data have shown great potential to advance land surface monitoring, but a critical challenge remains with the lack of an effective and accurate approach for cloud and cloud shadow detection, particularly in the tropics (Shendryk et al., 2019; Wang et al., 2020). To address this challenge, we developed a SpatioTemporal Integration approach for Automatic Cloud and Shadow Screening, STI-ACSS, and evaluated the effectiveness of this method across six tropical sites with various land cover types and percent cloud covers. Our results demonstrate that STI-ACSS can effectively detect clouds and cloud shadows with the highest accuracy (OA=98.03%; Table 2 and Fig. 4) relative to the other four state-of-the-art methods (Fmask, IHOT, ATSA, and the default PlanetScope QCs) and high temporal consistency across the image time series of two full years (Fig. 7). Taken together, these results suggest that STI-ACSS is an effective and accurate approach for automatic cloud and cloud shadow detection in the tropics.

In addition to the effectiveness assessment, the robustness of our developed STI-ACSS method was also rigorously assessed with three types of tests. First, the test sites spanned a large gradient in land cover types (e.g. forest, urban, cropland, savannah, and shrubland; Fig. 1), annual precipitation (e.g. 310-2600 mm per year; Table 1), and percent cloud covers (e.g. ranging from 0% to 100%; Fig. S1). Second, the testing images included various temporal changes in land surface reflectance, e.g. associated with plant phenology at the BCI-forest, k67-forest, and ASM-shrubland sites, seasonal land cover change at the Euc-plantation site, and grazing at the Machakos-savannah site (see Figs. 6 and 8, or refer to Cheng et al., 2020; Ma et al., 2013; Park et al., 2019; Qin et al., 2019; Wu et al., 2018). Third, a rigorous sensitivity analysis was conducted on the key parameter of the STI-ACSS algorithm, i.e. the outlier threshold that is used to determine cloud and cloud shadow outliers in the multi-temporal-based method (Fig. S2 and Table S2). With the above tests, STI-ACSS is demonstrated to be robust, and is able to detect clouds and cloud shadows with consistently high accuracies (OA=95%-100%; Figs. 4 and S3) across a wide range of land cover types, annual precipitation, and cloud coverage. STI-ACSS is also shown to be robust across the annual timescale, with consistently low commission errors (shadow UA =91.55%±14.21% and cloud UA =93.70%±14.20%; Table 2, and Fig. 8) in misclassifying those temporal changes associated with either plant phenology or land cover changes as the cloud/cloud shadow classes. Additionally, our sensitivity analysis demonstrates that STI-ACSS is insensitive to the outlier threshold used (Fig. S2) and can consistently detect clouds/cloud shadows with high accuracy (Table S2). All these results demonstrate the robustness of using STI-ACSS for automatic and accurate detection of clouds and cloud shadows.

The success of STI-ACSS relies on the following three strengths. First, to compensate for the limited spectral bands in PlanetScope images, STI-ACSS integrates both spatial and temporal spectral variability

to enhance cloud and cloud shadow detection. Single-image-based methods that only use the spatial-wide spectral information for cloud and cloud shadow detection often require more spectral bands as input, including SWIR and TIR bands (Huang et al., 2010; Irish et al., 2006; Luo et al., 2008; Zhu and Woodcock, 2012), which are not applicable to the PlanetScope data. Meanwhile, the multi-temporal-based methods that rely on the temporal-wide spectral variability for cloud/cloud shadow detection have been shown with difficulty in differentiating those temporal variations caused by clouds/cloud shadows from those associated with plant phenology or rapid intra-year land cover changes (Zhu and Woodcock, 2014). By assimilating the advantages of the two approaches while making full use of both spatial and temporal spectral variability, STI-ACSS not only compensates for the limited spectral bands of PlanetScope, but also improves the accuracy in the cloud and cloud shadow detection with reduced commission errors caused by plant phenology and intra-year land cover changes (Fig. 8). Although a similar spatiotemporal integration approach has been successfully tested in ATSA and Tmask in Landsat and Sentinel-2 images with a limited number of input spectral bands (i.e. 3~5 bands) (Zhu and Helmer; 2018; Zhu and Woodcock, 2014), the use of this approach in high-resolution PlanetScope satellites has not been explored previously. Such integration (i.e. STI-ACSS Step 1) with observed very high accuracy (OA=92.31%; Fig. 5) thus not only complements these previous studies by demonstrating the feasibility of utilizing the spatiotemporal integration approach for effective cloud/cloud shadow detection in PlanetScope images, but also suggests the potential to extend this approach to other high-resolution satellite sensors with limited spectral bands but frequent measurements, e.g. Gaofen-1, Skysat-2, SPOT-7 (Gaofen-1, 2020; SkySat, 2020; SPOT-7, 2020).

Second, to minimize the effect of cross-sensor inconsistency in cloud detection of PlanetScope images, STI-ACSS includes a novel adaptive threshold approach for cloud detection. As shown in Fig. S6, due to the radiometric inconsistency of multi-date and multi-sensor imagery, the optimal threshold that differentiates clouds from other backgrounds often varies from one image to another. This observation is particularly apparent in the PlanetScope constellation that is composed of more than 130 CubeSats, which vary in overpass times, atmospheric conditions, spectral response functions, and absolute radiometric accuracies (Houborg and McCabe, 2018). Similar to previous adaptive thresholds (Huang et al., 2010; Zhu and Woodcock, 2012; Zhu and Helmer, 2018), STI-ACSS also employs an adaptive method that optimizes the HOT threshold based on the surface and atmospheric conditions of each image. Unlike these previous approaches, which use a HOT threshold calculated from a fixed multiplier of the standard deviation or a fixed percentile of the HOT index across the entire image (or image time series), we allowed the threshold to be automatically optimized for each image without any predefined parameters. For demonstration purpose, we compared the results derived using our adaptive threshold approach with that derived using 6

other threshold approaches, including a fixed threshold (i.e. the average value of the HOT index across full image time series), and five fixed percentiles (i.e. 10^{th}, 30^{th}, 50^{th}, 70^{th} and 90^{th} percentiles) of each HOT index image (Fig. S6). As shown in Figs. S7 and S8, our adaptive threshold is the only approach that generates consistently high accuracies in cloud and cloud shadow detection across all the testing images (Table S3). These results also highlight that our proposed adaptive threshold approach can likely be extended to the cloud detection in other CubeSat constellations, e.g. Lingque and Landmapper-HD (Landmapper-HD, 2020; Lingque, 2020), which are also composed of many satellite sensors and have similar cross-sensor inconsistency issues as PlanetScope.

Third, to further refine the derived cloud and cloud shadow masks, STI-ACSS involves a two-step fine-tuning (i.e. morphological processing and object-based cloud and cloud shadow matching). Despite the above-mentioned advantages, two main issues remain with the spatiotemporal integration approach: (1) the "salt-and-pepper noises" that remain with the derived masks (Fig. 2), especially considering there is large spatial heterogeneity in the high spatial resolution imagery (Chen et al., 2018; Pu et al., 2011) such as PlanetScope, and the omission errors of cloud/cloud shadow edges that are the most common errors in cloud and cloud shadow detection methods (Foga et al., 2017); and (2) the commission errors that misclassify those non-cloud/cloud shadow pixels (e.g. bright or dark surfaces) into the cloud/cloud shadow classes (Fig. 2). To address these issues, we applied an integration of morphological processing (Soille, 1999; Zhu et al., 2015) and object-based cloud and cloud shadow matching (Zhu and Woodcock, 2012) to minimize these noises and errors. Importantly, we applied the morphological processing first so that it could increase the connectivity of objects, help sieve tiny objects, and ultimately improve the accuracy and efficiency for the subsequent use of the object-based matching approach for paired cloud and cloud shadow detections. Our two-step fine-tuning was built upon the previous approach (Fisher, 2014; Luo et al., 2008; Zhu and Helmer, 2018; Zhu and Woodcock, 2012), but, to our knowledge, has rarely been applied to high spatial-resolution satellite images, such as PlanetScope. Relative to the spatiotemporal integration approach (i.e. STI-ACSS Step 1), the performance improvement by an additional 5.72% in OA when involving this two-step fine-tuning (i.e. STI-ACSS Step 2; Fig. 5) further suggests the necessity and effectiveness of including this additional step on cloud/cloud shadow detection for PlanetScope satellite images.

Our study also identifies three important next steps that need to be considered for future advances. First, a considerable omission of low-density clouds/haze/cloud shadows, especially those surrounding thick clouds/deep cloud shadows or in an image with heavy haze contamination (Fig. S9), is found in STI-ACSS. This is very likely because the spectral reflectances of thin clouds/cloud shadows are much closer to clear pixels than thick clouds/deep cloud shadows, making their correct identification challenging when using

649 both single-image-based and multi-temporal-based methods (Li et al., 2020; Shao et al., 2019). Since recent
650 studies (He et al., 2013; Li et al., 2017) demonstrate that a guided filtering approach that uses the statistical
651 features of their neighboring pixels is effective for improving the detection of these low-density
652 clouds/cloud shadows, to resolve the considerable omission issue of low-density clouds/haze/cloud
653 shadows, we thus recommend integrating our STI-ACSS with the guided filtering approach in future
654 practices. Second, although the effectiveness of STI-ACSS for cloud/cloud shadow detection is evaluated
655 in this study, its performance might be contaminated by the high aerosol loading, which is a common
656 phenomenon in tropical areas due to biomass burning, deforestation, dust storms, and others (Cardil et al.,
657 2020; Silva et al., 2018). To investigate such impacts, we examined an additional site (7°29'S, 64°27'W)
658 in Canutama, Amazonas, Brazil, which is dominated by croplands and forests, with a reported wild fire
659 event in August 2019 (Lizundia-Loiola et al., 2020). Our results from this demonstration site (Fig. S10)
660 show that STI-ACSS can accurately screen smoke and smoke shadows, suggesting that STI-ACSS might
661 be also used for aerosol and aerosol shadow detection. Since there is large variability in spectral
662 characteristics and spatial distributions of aerosol episodes across the pan-tropics (Martins et al., 2002; Su
663 et al., 2017; Zhou et al., 2020), therefore more comprehensive future tests regarding aerosol detection and
664 impacts on STI-ACSS's performance remain needed but is beyond the scope of this paper. Third, since
665 STI-ACSS was mainly developed for the land surface monitoring in the tropics with water bodies at the
666 BCI-forest and Manaus-urban sites being manually masked out, the method may encounter challenges when
667 extending to the areas mixed with water surfaces (e.g. tropical wetlands, swamps, and flood plains) and
668 regions beyond the tropics (e.g. temperate and boreal regions). To address these issues, for broader
669 geographical applications that include water bodies, we recommend integrating our STI-ACSS with similar
670 approaches as used in ATSA and Fmask (Frazier and Page, 2000; Zhu and Helmer, 2018; Zhu and
671 Woodcock, 2012) to either automatically mask out water bodies that simultaneously account for the
672 seasonal flooding impacts or enable automatic cloud/cloud shadow detection over water bodies. While for
673 temperate/boreal regions, distinguishing between snow/ice and clouds would be a major task but remains
674 challenging for PlanetScope that does not include the commonly-used spectral bands (i.e. SWIR and TIR)
675 for separating snow/ice from clouds (Zhu and Woodcock, 2012). Since the use of texture features has been
676 shown with effectiveness for separating snow/ice from cloud/cloud shadow in some previous studies (Hu
677 et al., 2015; Li et al., 2017), we thus recommend for future attempts that integrate our STI-ACSS with the
678 use of the texture feature for the effectiveness assessments of extending STI-ACSS to the temperate/boreal
679 regions. Nonetheless, STI-ACSS is broadly effective on land within the tropics, a sector of the Earth where
680 previous efforts at automatic cloud and cloud shadow detection methods met with mixed success (Bodart
681 et al., 2011; Shendryk et al., 2019). STI-ACSS will prove valuable for tropical land monitoring, regardless

of whether STI-ACSS can be extended to other regions with additional subalgorithms or complementary tailor methods made for those contexts.

6. Conclusions

We have just presented a novel method, STI-ACSS, for automatic cloud and cloud shadow detection in tropical areas for PlanetScope imagery. This method firstly integrates spatial and temporal spectral variability in the image time series to obtain initial cloud/shadow masks, and then uses a two-step fine-tuning (i.e. morphological processing and object-based cloud and cloud shadow matching) to refine the initial masks and derive the final masks. The accuracy of STI-ACSS was evaluated at the six contrasting tropical sites over the two years (i.e. 2018 and 2019). Compared with the other four state-of-the-art cloud/cloud shadow detection methods (i.e. the default PlanetScope QC, Fmask, ATSA, and IHOT), STI-ACSS obtains the highest accuracies (average OA=98.03%, cloud PA=95.53%, and shadow PA=89.48%) with reference to manual cloud and shadow masks. Additionally, the robustness of STI-ACSS was also rigorously assessed under a wide range of scenarios, i.e. testing sites spanning large gradients in land cover types, annual precipitation, percent cloud covers, and temporal changes (associated with both plant phenology and land cover types), and sensitivity analysis on important parameters of STI-ACSS. With both improved accuracy and robustness assessments, we thus concluded that STI-ACSS is an accurate and robust approach for cloud and cloud shadow detection for PlanetScope images in the tropics, and is likely extendable to other satellite sensors with limited spectral bands, e.g. Gaofen-1, Skysat, and SPOT-7 (Gaofen-1, 2020; SkySat, 2020; SPOT-7, 2020).

Acknowledgements: We would thank the editors and the four anonymous reviewers for providing valuable suggestions and comments, which are greatly helpful in improving this manuscript. The work was supported by multiple funding agencies, including J.Wu, J.Wang, S.Wu, and Z.Guo supported by National Natural Science Foundation of China (#31922090) and the Research Grants Council Early Career Scheme (#27306020); J.Wang in part supported by the Division of Ecology and Biodiversity PDF research award; and D.Yang. supported by the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory. The PlanetScope data used in this study were accessed through the Education and Research Program, contracted between Planet Labs Inc. and HKU.

References:

Aleixo, I., Norris, D., Hemerik, L., Barbosa, A., Prata, E., Costa, F., & Poorter, L. (2019). Amazonian rainforest tree mortality driven by climate and functional traits. *Nature Climate Change, 9*, 384-388.

Araujo, A.C., Nobre, A.D., Kruijt, B., Elbers, J.A., Dallarosa, R., Stefani, P., von Randow, C., Manzi, A.O., Culf, A.D., Gash, J.H.C., Valentini, R., & Kabat, P. (2002). Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: The Manaus LBA site. *Journal of Geophysical Research-Atmospheres, 107*. https://doi.org/10.1029/2001JD000676.

Asner, G.P., Knapp, D.E., Broadbent, E.N., Oliveira, P.J.C., Keller, M., & Silva, J.N. (2005). Selective logging in the Brazilian Amazon. *Science, 310*, 480-482.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., Rodenbeck, C., Arain, M.A., Baldocchi, D., Bonan, G.B., Bondeau, A., Cescatti, A., Lasslop, G., Lindroth, A., Lomas, M., Luyssaert, S., Margolis, H., Oleson, K.W., Roupsard, O., Veenendaal, E., Viovy, N., Williams, C., Woodward, F.I., & Papale, D. (2010). Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate. *Science, 329*, 834-838.

Bodart, C., Eva, H., Beuchle, R., Rasi, R., Simonetti, D., Stibig, H.J., Brink, A., Lindquist, E., & Achard, F. (2011). Pre-processing of a sample of multi-scene and multi-date Landsat imagery used to monitor forest cover changes over the tropics. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*, 555-563.

Brando, P.M., Soares-Filho, B., Rodrigues, L., Assuncao, A., Morton, D., Tuchschneider, D., Fernandes, E.C.M., Macedo, M.N., Oliveira, U., & Coe, M.T. (2020). The gathering firestorm in southern Amazonia. *Science Advances, 6*. https://doi.org/10.1126/sciadv.aay1632.

Braswell, B.H., Schimel, D.S., Linder, E., & Moore, B. (1997). The response of global terrestrial ecosystems to interannual temperature variability. *Science, 278*, 870-872.

Breunig, F.M., Galvao, L.S., Dalagnol, R., Dauve, C.E., Parraga, A., Santi, A.L., Della Flora, D.P., & Chen, S.S. (2020). Delineation of management zones in agricultural fields using cover crop biomass estimates from PlanetScope data. *International Journal of Applied Earth Observation and Geoinformation, 85*. https://doi.org/10.1016/j.jag.2019.102004.

Byrne, G.F., Crapper, P.F., & Mayo, K.K. (1980). Monitoring Land-Cover Change by Principal Component Analysis of Multitemporal Landsat Data. *Remote Sensing of Environment, 10*, 175-184.

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.

Cardil, A., de-Miguel, S., Silva, C.A., Reich, P.B., Calkin, D., Brancalion, P.H.S., Vibrans, A.C., Gamarra, J.G.P., Zhou, M., Pijanowski, B.C., Hui, C., Crowther, T.W., Herault, B., Piotto, D., Salas-Eljatib, C., Broadbent, E.N., Zambrano, A.M.A., Picard, N., Aragao, L.E.O.C., Bastin, J.F., Routh, D., van den Hoogen, J., Peri, P.L., & Liang, J.J. (2020). Recent deforestation drove the spike in Amazonian fires. *Environmental Research Letters, 15*. https://doi.org/10.1088/1748-9326/abcac7.

Ceccon, E., & Miramontes, O. (2008). Reversing deforestation? Bioenergy and society in two Brazilian models. *Ecological Economics, 67*, 311-317.

Chai, D., Newsam, S., Zhang, H.K.K., Qiu, Y., & Huang, J.F. (2019). Cloud and cloud shadow detection in Landsat imagery based on deep convolutional neural networks. *Remote Sensing of Environment, 225*, 307-316.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anastasiou, A., Enquist, B.J., Caravasi, E.G.C., Doughty, C.E., Saleska, S.R., Martin, R.E., & Gerard, F.F. (2017). Leaf aging of Amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist, 214*, 1049-1063.

761 Chen, C., Park, T., Wang, X.H., Piao, S.L., Xu, B.D., Chaturvedi, R.K., Fuchs, R., Brovkin, V., Ciais, P.,
762 Fensholt, R., Tommervik, H., Bala, G., Zhu, Z.C., Nemani, R.R., & Myneni, R.B. (2019a). China
763 and India lead in greening of the world through land-use management. *Nature Sustainability, 2*,
764 122-129.
765 Chen, B., Jin, Y.F., & Brown, P. (2019b). An enhanced bloom index for quantifying floral phenology
766 using multi-scale remote sensing observations. *ISPRS Journal of Photogrammetry and Remote*
767 *Sensing, 156*, 108-120.
768 Chen, G.Z., Zhang, X.D., Wang, Q., Dai, F., Gong, Y.F., & Zhu, K. (2018). Symmetrical Dense-Shortcut
769 Deep Fully Convolutional Networks for Semantic Segmentation of Very-High-Resolution
770 Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and*
771 *Remote Sensing, 11*, 1633-1644.
772 Chen, S.L., Chen, X.H., Chen, J., Jia, P.F., Cao, X., & Liu, C.Y. (2016). An Iterative Haze Optimized
773 Transformation for Automatic Cloud/Haze Detection of Landsat Imagery. *IEEE Transactions on*
774 *Geoscience and Remote Sensing, 54*, 2682-2694.
775 Cheng, Y., Vrieling, A., Fava, F., Meroni, M., Marshall, M., & Gachoki, S. (2020). Phenology of short
776 vegetation cycles in a Kenyan rangeland from PlanetScope and Sentinel-2. *Remote Sensing of*
777 *Environment, 248*. https://doi.org/10.1016/j.rse.2020.112004.
778 Choi, H., & Bindschadler, R. (2004). Cloud detection in Landsat imagery of ice sheets using shadow
779 matching technique and automatic normalized difference snow index threshold value decision.
780 *Remote Sensing of Environment, 91*, 237-242.
781 Cleverly, J., Eamus, D., Coupe, N.R., Chen, C., Maes, W.H., Li, L.H., Faux, R., Santini, N.S., Rumman,
782 R., Yu, Q., & Huete, A. (2016). Soil moisture controls on phenology and productivity in a semi-
783 arid critical zone. *Science of the Total Environment, 568*, 1227-1237.
784 Comita, L.S., Muller-Landau, H.C., Aguilar, S., & Hubbell, S.P. (2010). Asymmetric Density
785 Dependence Shapes Species Abundances in a Tropical Tree Community. *Science, 329*, 330-332.
786 de Souza, D.O., & Alvala, R.C.D. (2014). Observational evidence of the urban heat island of Manaus
787 City, Brazil. *Meteorological Applications, 21*, 186-193.
788 Detto, M., Wright, S.J., Calderon, O., & Muller-Landau, H.C. (2018). Resource acquisition and
789 reproductive strategies of tropical forest in response to the El Nino-Southern Oscillation. *Nature*
790 *Communications, 9*. https://doi.org/10.1038/s41467-41018-03306-41469.
791 Fisher, A. (2014). Cloud and Cloud-Shadow Detection in SPOT5 HRG Imagery with Automated
792 Morphological Feature Extraction. *Remote Sensing, 6*, 776-800.
793 Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L.,
794 Hughes, M.J., & Laue, B. (2017). Cloud detection algorithm comparison and validation for
795 operational Landsat data products. *Remote Sensing of Environment, 194*, 379-390.
796 Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of*
797 *Environment, 80*, 185-201.
798 Frazier, P.S., & Page, K.J. (2000). Water body detection and delineation with Landsat TM data.
799 *Photogrammetric Engineering and Remote Sensing, 66*, 1461-1467.
800 Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X.M.
801 (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of
802 new datasets. *Remote Sensing of Environment, 114*, 168-182.
803 Fu, R., Yin, L., Li, W.H., Arias, P.A., Dickinson, R.E., Huang, L., Chakraborty, S., Fernandes, K.,
804 Liebmann, B., Fisher, R., & Myneni, R.B. (2013). Increased dry-season length over southern
805 Amazonia in recent decades and its implication for future climate projection. *Proceedings of the*
806 *National Academy of Sciences of the United States of America, 110*, 18110-18115.
807 Ganci, G., Cappello, A., Bilotta, G., & Del Negro, C. (2020). How the variety of satellite remote sensing
808 data over volcanoes can assist hazard monitoring efforts: The 2011 eruption of Nabro volcano.
809 *Remote Sensing of Environment, 236*. https://doi.org/10.1016/j.rse.2019.111426.

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). Upper Saddle River, N.J.: Prentice-Hall, Inc., USA.

Goodwin, N.R., Collett, L.J., Denham, R.J., Flood, N., & Tindall, D. (2013). Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM plus time series. *Remote Sensing of Environment, 134*, 50-65.

Hagolle, O., Huc, M., Pascual, D.V., & Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VEN mu S, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment, 114*, 1747-1755.

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.

Hansen, P.C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM review, 34*, 561-580.

Hansen, P.C., & Oleary, D.P. (1993). The Use of the L-Curve in the Regularization of Discrete Iii-Posed Problems. *Siam Journal on Scientific Computing, 14*, 1487-1503.

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

Holdridge, L.R. (1947). Determination of world plant formations from simple climatic data. *Science, 105*, 367-368.

Houborg, R., & McCabe, M.F. (2018). A Cubesat enabled Spatio-Temporal Enhancement Method (CESTEM) utilizing Planet, Landsat and MODIS data. *Remote Sensing of Environment, 209*, 211-226.

Hu, X.Y., Wang, Y., & Shan, J. (2015). Automatic Recognition of Cloud Images by Using Visual Saliency Features. *IEEE Geoscience and Remote Sensing Letters, 12*, 1760-1764.

Huang, C.Q., Thomas, N., Goward, S.N., Masek, J.G., Zhu, Z.L., Townshend, J.R.G., & Vogelmann, J.E. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing, 31*, 5449-5464.

Hughes, M.J., & Hayes, D.J. (2014). Automated Detection of Cloud and Cloud Shadow in Single-Date Landsat Imagery Using Neural Networks and Spatial Post-Processing. *Remote Sensing, 6*, 4907-4926.

Irish, R.R., Barker, J.L., Goward, S.N., & Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm. *Photogrammetric Engineering and Remote Sensing, 72*, 1179-1188.

Jeong, S.J., Ho, C.H., Gim, H.J., & Brown, M.E. (2011). Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology, 17*, 2385-2399.

Jeppesen, J.H., Jacobsen, R.H., Inceoglu, F., & Toftegaard, T.S. (2019). A cloud detection algorithm for satellite imagery based on deep learning. *Remote Sensing of Environment, 229*, 247-259.

Jin, S.M., Homer, C., Yang, L.M., Xian, G., Fry, J., Danielson, P., & Townsend, P.A. (2013). Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA. *International Journal of Remote Sensing, 34*, 1540-1560.

Kim, Y., Knox, R.G., Longo, M., Medvigy, D., Hutyra, L.R., Pyle, E.H., Wofsy, S.C., Bras, R.L., & Moorcroft, P.R. (2012). Seasonal carbon dynamics and water fluxes in an Amazon rainforest. *Global Change Biology, 18*, 1322-1334.

Knight, D.H. (1975). Phytosociological Analysis of Species-Rich Tropical Forest on Barro Colorado Island, Panama. *Ecological Monographs, 45*, 259-284.

Landmapper-HD (2020). Landmapper-HD 1, ..., 20 In: https://space.skyrocket.de/doc_sdat/landmapper-hd.htm (accessed 03 October 2020).

Laurance, W.F., Sayer, J., & Cassman, K.G. (2014). Agricultural expansion and its impacts on tropical nature. *Trends in Ecology & Evolution, 29*, 107-116.
Leigh, E.G. (1999). *Tropical forest ecology: a view from Barro Colorado Island*. New York, NY: Oxford University Press.
Li, J., Wu, Z.C., Hu, Z.W., Zhang, J.Q., Li, M.L., Mo, L., & Molinier, M. (2020). Thin cloud removal in optical remote sensing images based on generative adversarial networks and physical model of cloud distortion. *ISPRS Journal of Photogrammetry and Remote Sensing, 166*, 373-389.
Li, P.F., Dong, L.M., Xiao, H.C., & Xu, M.L. (2015). A cloud image detection method based on SVM vector machine. *Neurocomputing, 169*, 34-42.
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.
Lingque (2020). Lingque 1A, 1B, 1C. In: https://space.skyrocket.de/doc_sdat/lingque-1.htm (accessed 03 October 2020).
Liu, R.G., & Liu, Y. (2013). Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sensing of Environment, 133*, 21-37.
Lizundia-Loiola, J., Pettinari, M.L., & Chuvieco, E. (2020). Temporal Anomalies in Burned Area Trends: Satellite Estimations of the Amazonian 2019 Fire Crisis. *Remote Sensing, 12*. https://doi.org/10.3390/rs12010151.
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.
Luo, Y., Trishchenko, A.P., & Khlopenkov, K.V. (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment, 112*, 4167-4185.
Ma, X.L., Huete, A., Yu, Q., Coupe, N.R., Davies, K., Broich, M., Ratana, P., Beringer, J., Hutley, L.B., Cleverly, J., Boulain, N., & Eamus, D. (2013). Spatial patterns and temporal dynamics in savanna vegetation phenology across the North Australian Tropical Transect. *Remote Sensing of Environment, 139*, 97-115.
Maeda, E.E., Moura, Y.M., Wagner, F., Hilker, T., Lyapustin, A.I., Wang, Y.J., Chave, J., Mottus, M., Aragao, L.E.O.C., & Shimabukuro, Y. (2016). Consistency of vegetation index seasonality across the Amazon rainforest. *International Journal of Applied Earth Observation and Geoinformation, 52*, 42-53.
Martins, J.V., Tanre, D., Remer, L., Kaufman, Y., Mattoo, S., & Levy, R. (2002). MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters, 29*. https://doi.org/10.1029/2001GL013252.
Mohan, M., & Kandya, A. (2015). Impact of urbanization and land-use/land-cover change on diurnal temperature range: A case study of tropical urban airshed of India using remote sensing data. *Science of the Total Environment, 506*, 453-465.
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.
Phillips, O.L., Aragao, L.E.O.C., Lewis, S.L., Fisher, J.B., Lloyd, J., Lopez-Gonzalez, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C.A., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T.R., Banki, O., Blanc, L., Bonal, D., Brando, P., Chave, J., de Oliveira, A.C.A., Cardozo, N.D., Czimczik, C.I., Feldpausch, T.R., Freitas, M.A., Gloor, E., Higuchi, N., Jimenez, E., Lloyd, G., Meir, P., Mendoza, C., Morel, A., Neill, D.A., Nepstad, D., Patino, S., Penuela, M.C., Prieto, A., Ramirez, F., Schwarz, M., Silva, J., Silveira, M., Thomas, A.S., ter Steege, H., Stropp, J., Vasquez, R., Zelazowski, P., Davila, E.A., Andelman, S.,

Andrade, A., Chao, K.J., Erwin, T., Di Fiore, A., Honorio, E., Keeling, H., Killeen, T.J., Laurance, W.F., Cruz, A.P., Pitman, N.C.A., Vargas, P.N., Ramirez-Angulo, H., Rudas, A., Salamao, R., Silva, N., Terborgh, J., & Torres-Lezama, A. (2009). Drought Sensitivity of the Amazon Rainforest. *Science, 323*, 1344-1347.

Piao, S.L., Fang, J.Y., Zhou, L.M., Ciais, P., & Zhu, B. (2006). Variations in satellite-derived phenology in China's temperate vegetation. *Global Change Biology, 12*, 672-685.

Planet (2018). Planet Imagery Product Specification [WWW Document]. In. https://assets.planet.com/docs/Combined-Imagery-Product-Spec-Dec-2018.pdf (accessed 03 March 2020).

Planet (2020). Planet Imagery Product Specification [WWW Document]. In. https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf (accessed 03 October 2020).

Pu, R.L., Landry, S., & Yu, Q. (2011). Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing, 32*, 3285-3308.

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.

Rice, A.H., Pyle, E.H., Saleska, S.R., Hutyra, L., Palace, M., Keller, M., de Camargo, P.B., Portilho, K., Marques, D.F., & Wofsy, S.C. (2004). Carbon balance and vegetation dynamics in an old-growth Amazonian forest. *Ecological Applications, 14*, S55-S71.

Roy, D.P., Ju, J.C., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E., & Zhang, C.S. (2010). Web-enabled Landsat Data (WELD): Landsat ETM plus composited mosaics of the conterminous United States. *Remote Sensing of Environment, 114*, 35-49.

Samanta, A., Ganguly, S., Hashimoto, H., Devadiga, S., Vermote, E., Knyazikhin, Y., Nemani, R.R., & Myneni, R.B. (2010). Amazon forests did not green-up during the 2005 drought. *Geophysical Research Letters, 37*. https://doi:10.1029/2009GL042154.

Scaramuzza, P.L., Bouchard, M.A., & Dwyer, J.L. (2012). Development of the Landsat Data Continuity Mission Cloud-Cover Assessment Algorithms. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 1140-1154.

Segal-Rozenhaimer, M., Li, A., Das, K., & Chirayath, V. (2020). Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (CNN). *Remote Sensing of Environment, 237*. https://doi.org/10.1016/j.rse.2019.111446.

Shao, Z.F., Pan, Y., Diao, C.Y., & Cai, J.J. (2019). Cloud Detection in Remote Sensing Images Based on Multiscale Features-Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing, 57*, 4062-4076.

Shendryk, Y., Rist, Y., Ticehurst, C., & Thorburn, P. (2019). Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 157*, 124-136.

Silva, C.H.L., Aragao, L.E.O.C., Fonseca, M.G., Almeida, C.T., Vedovato, L.B., & Anderson, L.O. (2018). Deforestation-Induced Fragmentation Increases Forest Fire Occurrence in Central Brazilian Amazonia. *Forests, 9*. https://doi.org/10.3390/f9060305.

Silva, M.C.D., Scarano, F.R., & Cardel, F.D.S. (1995). Regeneration of an Atlantic Forest Formation in the Understorey of a Eucalyptus-Grandis Plantation in South-Eastern Brazil. *Journal of Tropical Ecology, 11*, 147-152.

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).

962 Soille, P. (1999). *Morphological image analysis: Principles and applications* Springer-Verlag Berlin
963 Heidelberg.
964 Soille, P., & Pesaresi, M. (2002). Advances in mathematical morphology applied to geoscience and
965 remote sensing. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 2042-2055.
966 SPOT-7 (2020). SPOT-6 and SPOT-7 Commercial Imaging Constellation. In.
967 https://directory.eoportal.org/web/eoportal/satellite-missions/s/spot-6-7 (accessed 03 October
968 2019).
969 Stark, S.C., Breshears, D.D., Garcia, E.S., Law, D.J., Minor, D.M., Saleska, S.R., Swann, A.L.S.,
970 Villegas, J.C., Aragao, L.E.O.C., Bella, E.M., Borma, L.S., Cobb, N.S., Litvak, M.E.,
971 Magnusson, W.E., Morton, J.M., & Redmond, M.D. (2016). Toward accounting for ecoclimate
972 teleconnections: intra- and inter-continental consequences of altered energy balance after
973 vegetation change. *Landscape Ecology, 31*, 181-194.
974 Su, Q., Sun, L., Di, M., Liu, X., & Yang, Y. (2017). A method for the spectral analysis and identification
975 of Fog, Haze and Dust storm using MODIS data. *Atmospheric Measurement Techniques*
976 *Discussions*, 1-20. https://doi.org/10.5194/amt-2017-306.
977 Vigna, I., Bigi, V., Pezzoli, A., & Besana, A. (2020). Comparison and Bias-Correction of Satellite-
978 Derived Precipitation Datasets at Local Level in Northern Kenya. *Sustainability, 12*.
979 https://doi.org/10.3390/su12072896.
980 Volkov, I., Banavar, J.R., Hubbell, S.P., & Maritan, A. (2003). Neutral theory and relative species
981 abundance in ecology. *Nature, 424*, 1035-1037.
982 Wang, B., Ono, A., Muramatsu, K., & Fujiwara, N. (1999). Automated detection and removal of clouds
983 and their shadows from landsat TM images. *IEEE Transactions on Information and Systems,*
984 *E82d*, 453-460.
985 Wang, F.R., Fan, X.M., Yunus, A.P., Subramanian, S.S., Alonso-Rodriguez, A., Dai, L.X., Xu, Q., &
986 Huang, R.Q. (2019). Coseismic landslides triggered by the 2018 Hokkaido, Japan (M-w 6.6),
987 earthquake: spatial distribution, controlling factors, and possible failure mechanism. *Landslides,*
988 *16*, 1551-1566.
989 Wang, J., Yang, D.D., Detto, M., Nelson, B.W., Chen, M., Guan, K.Y., Wu, S.B., Yan, Z.B., & Wu, J.
990 (2020). Multi-scale integration of satellite remote sensing improves characterization of dry-
991 season green-up in an Amazon tropical evergreen forest. *Remote Sensing of Environment, 246*.
992 https://doi.org/10.1016/j.rse.2020.111865.
993 Wei, J., Huang, W., Li, Z.Q., Sun, L., Zhu, X.L., Yuan, Q.Q., Liu, L., & Cribb, M. (2020). Cloud
994 detection for Landsat imagery by combining the random forest and superpixels extracted via
995 energy-driven sampling segmentation approaches. *Remote Sensing of Environment, 248*.
996 https://doi.org/10.1016/j.rse.2020.112005.
997 Wilson, M.J., & Oreopoulos, L. (2013). Enhancing a Simple MODIS Cloud Mask Algorithm for the
998 Landsat Data Continuity Mission. *IEEE Transactions on Geoscience and Remote Sensing, 51*,
999 723-731.
1000 Wright, J.S., Fu, R., Worden, J.R., Chakraborty, S., Clinton, N.E., Risi, C., Sun, Y., & Yin, L. (2017).
1001 Rainforest-initiated wet season onset over the southern Amazon. *Proceedings of the National*
1002 *Academy of Sciences of the United States of America, 114*, 8481-8486.
1003 Wu, J., Albert, L.P., Lopes, A.P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K.T., Guan, K.Y., Stark,
1004 S.C., Christoffersen, B., Prohaska, N., Tavares, J.V., Marostica, S., Kobayashi, H., Ferreira, M.L.,
1005 Campos, K.S., da Silva, R., Brando, P.M., Dye, D.G., Huxman, T.E., Huete, A.R., Nelson, B.W.,
1006 & Saleska, S.R. (2016). Leaf development and demography explain photosynthetic seasonality in
1007 Amazon evergreen forests. *Science, 351*, 972-976.
1008 Wu, J., Kobayashi, H., Stark, S.C., Meng, R., Guan, K.Y., Tran, N.N., Gao, S.C., Yang, W., Restrepo-
1009 Coupe, N., Miura, T., Oliviera, R.C., Rogers, A., Dye, D.G., Nelson, B.W., Serbin, S.P., Huete,
1010 A.R., & Saleska, S.R. (2018). Biological processes dominate seasonality of remotely sensed
1011 canopy greenness in an Amazon evergreen forest. *New Phytologist, 217*, 1507-1520.

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.

Yuan, W.P., Zheng, Y., Piao, S.L., Ciais, P., Lombardozzi, D., Wang, Y.P., Ryu, Y., Chen, G.X., Dong, W.J., Hu, Z.M., Jain, A.K., Jiang, C.Y., Kato, E., Li, S.H., Lienert, S., Liu, S.G., Nabel, J.E.M.S., Qin, Z.C., Quine, T., Sitch, S., Smith, W.K., Wang, F., Wu, C.Y., Xiao, Z.Q., & Yang, S. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science Advances, 5*. https://doi.org/10.1126/sciadv.aax1396.

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.

Zhang, Y., Guindon, B., & Cihlar, J. (2002). An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sensing of Environment, 82*, 173-187.

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.

Zhou, Y.P., Levy, R.C., Remer, L.A., Mattoo, S., Shi, Y.X., & Wang, C.X. (2020). Dust Aerosol Retrieval Over the Oceans With the MODIS/VIIRS Dark-Target Algorithm: 1. Dust Detection. *Earth and Space Science, 7*. https://doi.org/10.1029/2020EA001222.

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, Z., Qiu, S., He, B., & Deng, C. (2018). Cloud and cloud shadow detection for Landsat images: The fundamental basis for analyzing Landsat time series. *Remote Sensing Time Series Image Processing* (pp. 25-46): CRC Press.

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.

Zhu, Z., & Woodcock, C.E. (2014). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment, 152*, 217-234.

Figure 1. Study sites. (a) locations (the red stars) and (b) false color composites (RGB=NIR-Red-Green) of PlanetScope images of the six testing sites (i.e. S1-S6) in the tropics, including Barro Colorado Island (BCI) in Panama (spatial extent: 6km×7.26km), Manaus city in North Brazil (spatial extent: 20km×20km), Tapajós National Forest k67 in North Brazil (spatial extent: 10km×10km), Eucalyptus plantation in South Brazil (spatial extent: 10km×10km), Machakos County in Kenya (spatial extent: 10km×10km), and Alice Springs in Australia (spatial extent: 10km×10km). The temporal coverage of PlanetScope observations at these sites spans January-December in 2018-2019. The main land cover type of the testing sites are a moist forest at the BCI site (henceforth BCI-forest), a metropolitan urban area at the Manaus site (henceforth Manaus-urban), a primary intact moist forest at the k67 site (henceforth k67-forest), a managed ecosystem composed of plantation, forest, and buildings at the Eucalyptus plantation site (henceforth Euc-plantation), a savannah farm at the Machakos site (henceforth Machakos-savannah), and a shrubland ecosystem from a Mulga woodland at the Alice Springs site (henceforth Alice Springs Mulga- or ASM-shrubland). The map in panel (a) is adapted from National Geographic, ESRI.

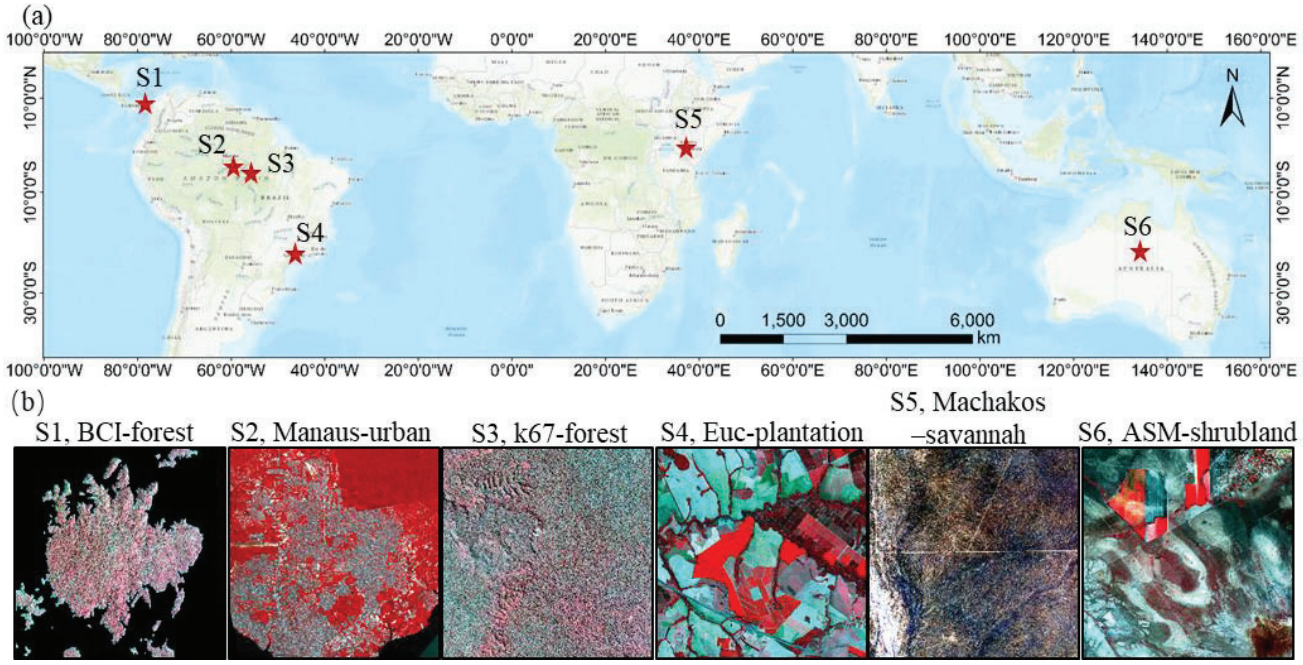


Figure 2. Flowchart of the SpatioTemporal Integration approach for Automatic Cloud and Shadow Screening (STI-ACSS) method. It includes two key steps composed of five tasks (A-E, more details see Section 3.1-3.2). Step 1: generating initial cloud and shadow masks through the integration of single-image-based methods based on cloud and shadow indices (Task A and B) and a multi-temporal-based method with an iterative outlier detection (Task C); Step 2: performing a first-round fine tuning through morphological processing (Task D) and a second-round fine tuning through an object-based cloud and cloud shadow matching approach (Task E) on the initial cloud and shadow masks. Thus, the final cloud and shadow masks are generated.

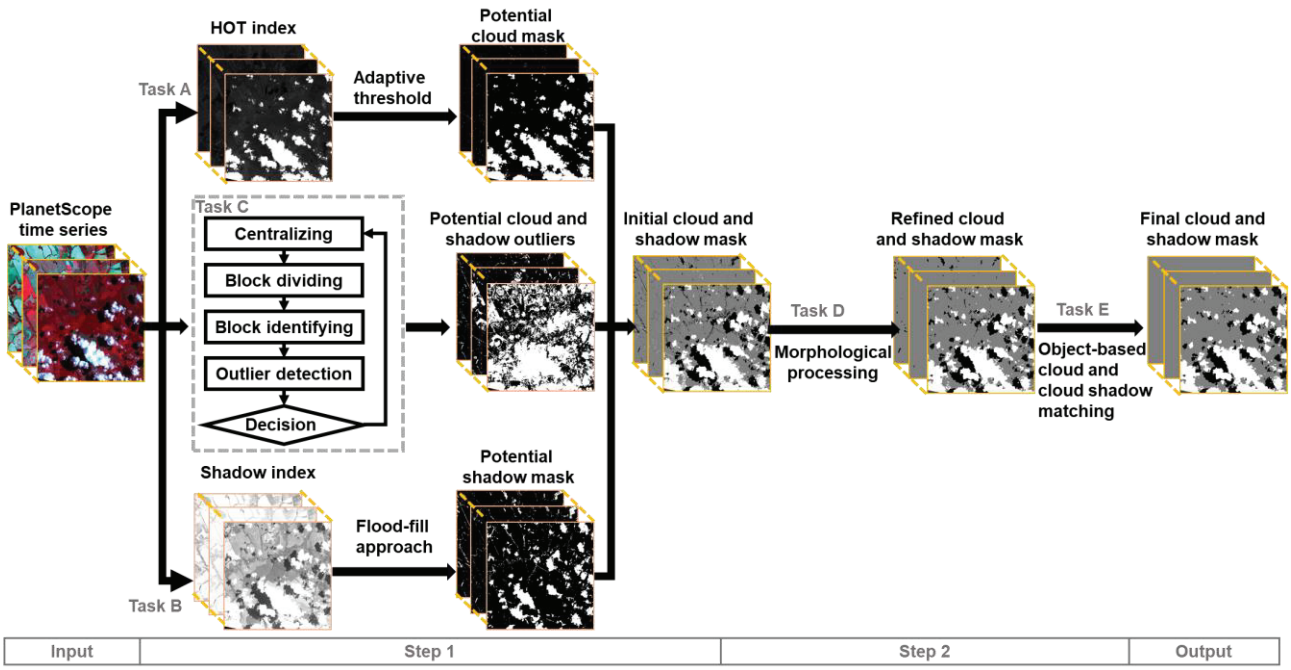


Figure 3. Example demonstration of how to determine the optimal threshold for differentiating cloud pixels from cloud-free pixels using an adaptive approach (Task A, Fig. 2). The blue line shows the number of potential cloud pixels (Y-axis) that have haze optimized transformation (HOT) values larger than a given HOT threshold (X-axis). The brown dashed line connects the end point and start point of this curve. The optimal threshold is determined as the L-shaped "corner", which has the largest distance to the brown dashed line (e.g. 180 in this case on April 06, 2018 at the Euc-plantation site).

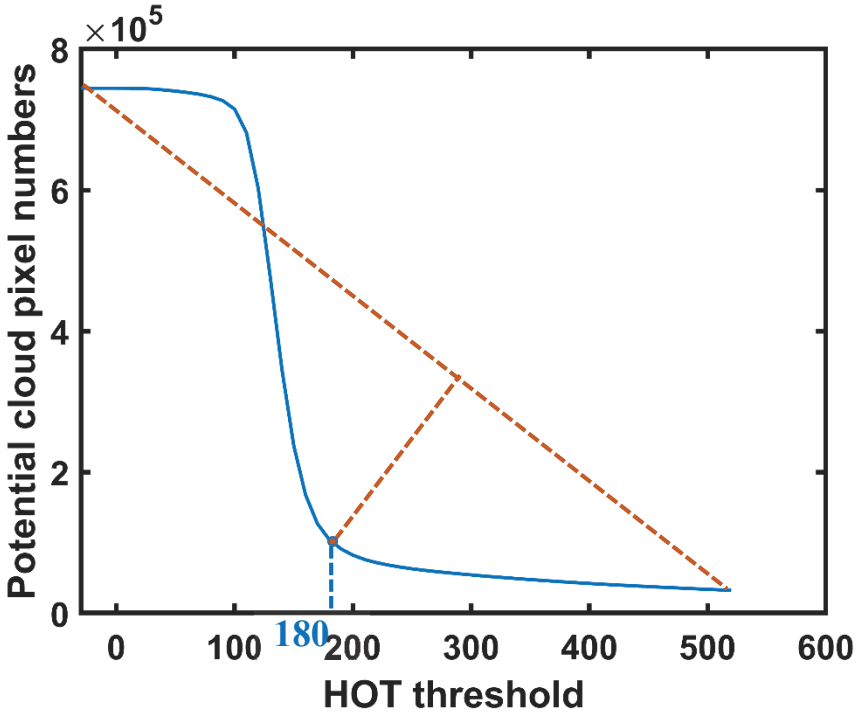


Figure 4. Cross-method comparison across a total of 47 representative images in 2018-2019 at all six sites using manual masks as benchmarks. The results of cross-method comparison include (a) the overall accuracy (OA) and (b) the average OA value across the six sites. The five methods examined here include STI-ACSS, the default PlanetScope quality control layers (i.e. UDM and UDM2), Function of Mask (Fmask), Iterative Haze-Optimized Transformation (IHOT), and Automatic Time-Series Analysis (ATSA); error bars indicate one standard deviation.

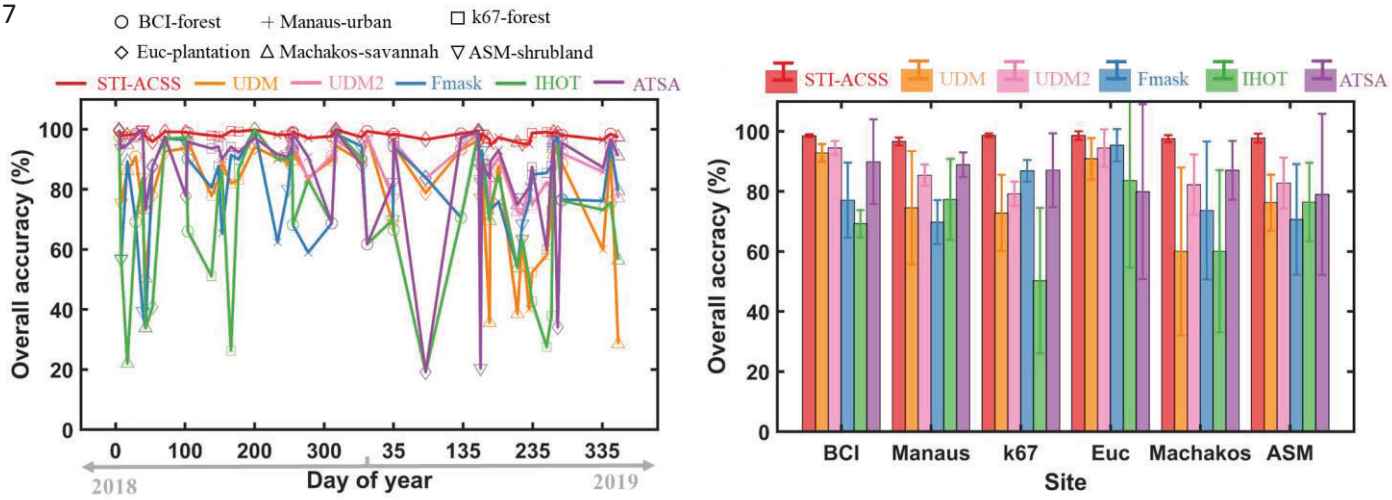


Figure 5. Comparison of two STI-ACSS steps across a total of 47 representative images in 2018-2019 at all six sites using manual masks as benchmarks. The results of this comparison include (a) the OA and (b) the average OA value across the six sites. The two steps examined here include STI-ACSS Step 1: generating initial cloud and shadow masks (including Tasks A-C of Fig. 2; see Section 3.1) and STI-ACSS Step 2: refining initial cloud and shadow masks (including Tasks D-E of Fig. 2; see Section 3.2); error bars indicate one standard deviation. The results show that the average OA of STI-ACSS was improved by 5.72% from Step 1 to Step 1+Step 2 (Step 1: 92.31%; Step 1+Step 2: 98.03%).

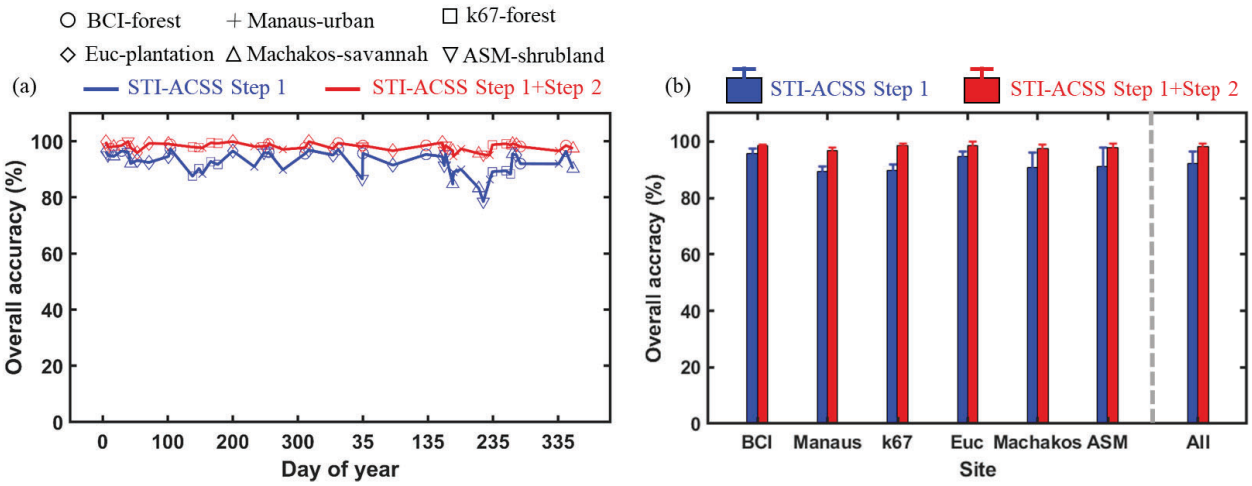
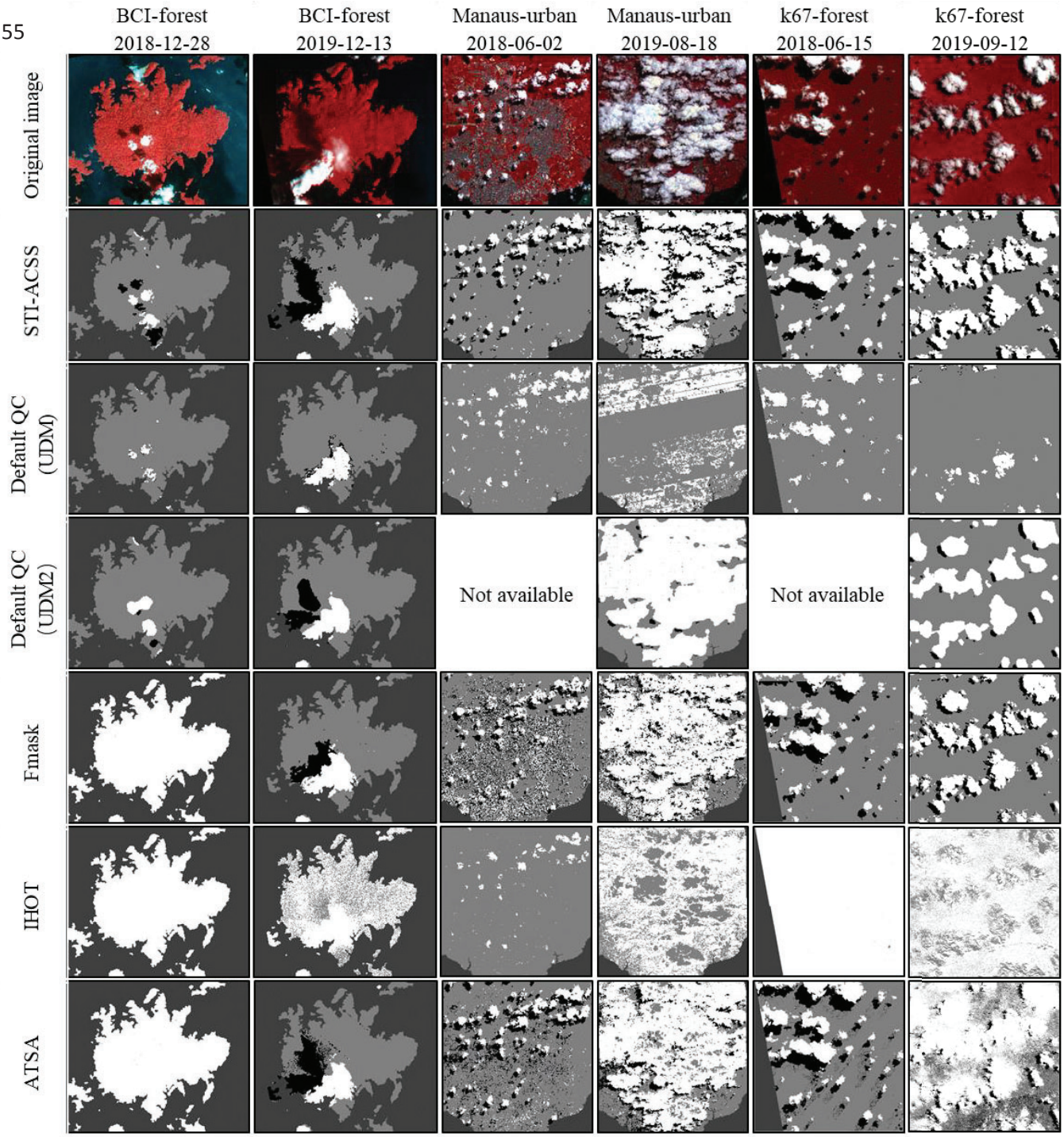


Figure 6. Comparisons of cloud and shadow masks derived from STI-ACSS and the other four state-of-the-art methods using manual cloud and shadow masks (Fig. S11) as benchmarks. Twelve representative images from all six sites spanning a large gradient in percent cloud cover ranging from 2% to 86% were used for demonstration purposes. False color composites (RGB=NIR-Red-Green) of selected PlanetScope images (the first row) and corresponding cloud and shadow masks using STI-ACSS (the second row), the default QCs (including UDM and UDM2 in the third and fourth row, respectively), Fmask (the fifth row), IHOT (the sixth row), and ATSA (the seventh row) are shown below. Color legends include clouds in white, cloud shadows in black, clear pixels in light grey, and backgrounds (covering water bodies or missing data) in dark grey.



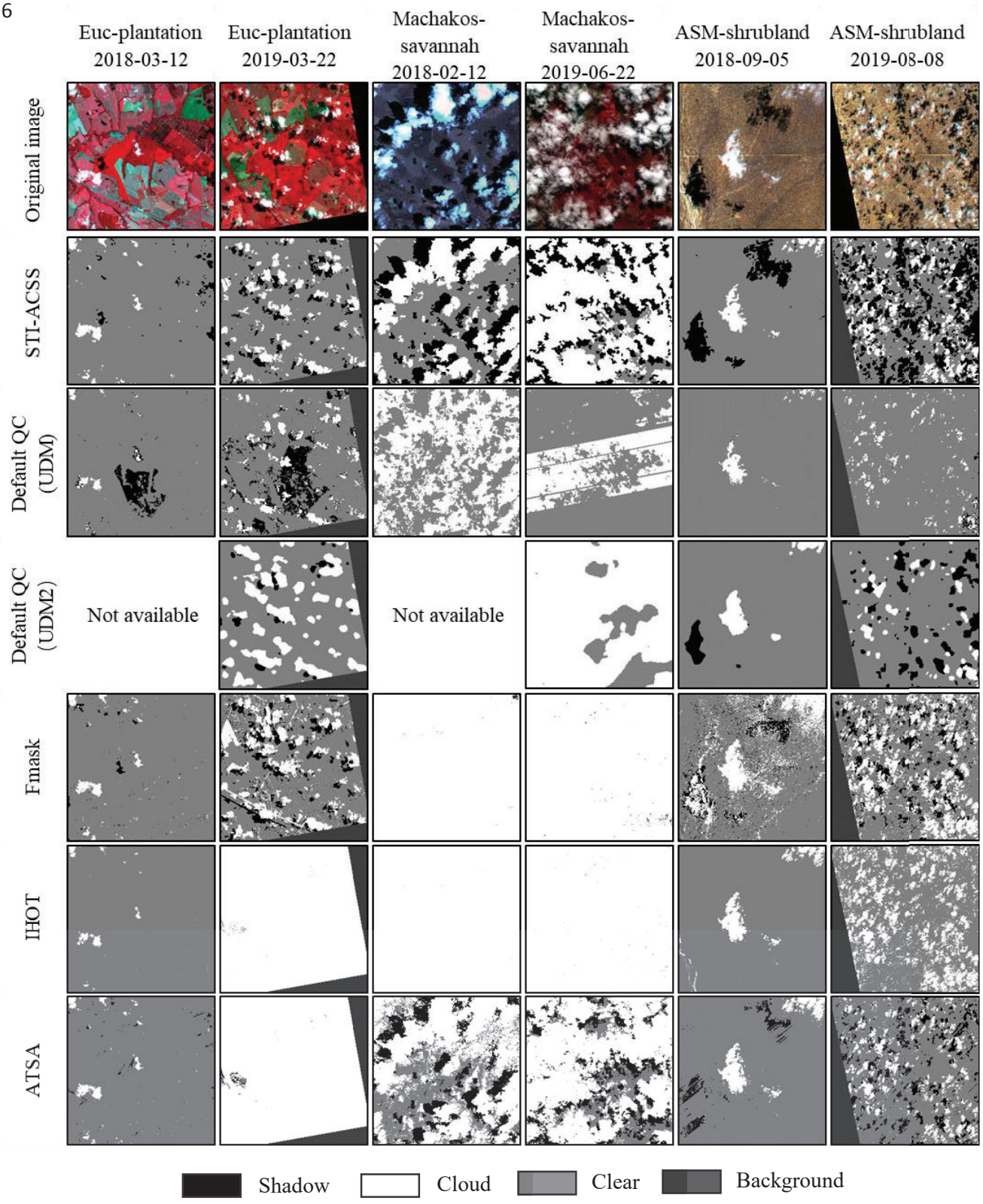

Euc-plantation 2018-03-12
Euc-plantation 2019-03-22
Machakos-savannah 2018-02-12
Machakos-savannah 2019-06-22
ASM-shrubland 2018-09-05
ASM-shrubland 2019-08-08
Original image
STI-ACSS
Default QC (UDM)
Default QC (UDM2)
Not available
Not available
Fmask
IHOT
ATSA
Shadow
Cloud
Clear
Background

Figure 7. Cross-method comparison across the full image time series in 2018-2019 at all six sites using the default QC as the benchmark. The comparison includes the average OA across the six sites (the left of the grey dashed line) and the average OA value of all sites (the right of the grey dashed line) among the four methods, i.e. STI-ACSS, Fmask, IHOT, and ATSA. Error bars indicate one standard deviation.

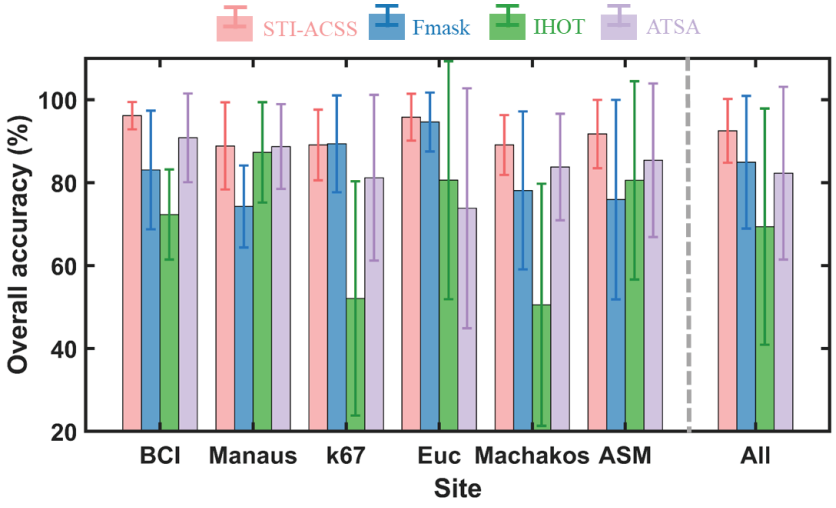


Figure 8. Assessing the effect of temporal reflectance changes on the derived cloud and shadow masks of STI-ACSS across full seasons of 2018-2019 at the Ecu-plantation site. Colored circles indicate regions with large apparent temporal reflectance changes that were possibly caused by harvesting/land cover changes (e.g. deforestation and reforestation). Red squares show the commission errors in cloud and shadow masks of STI-ACSS.

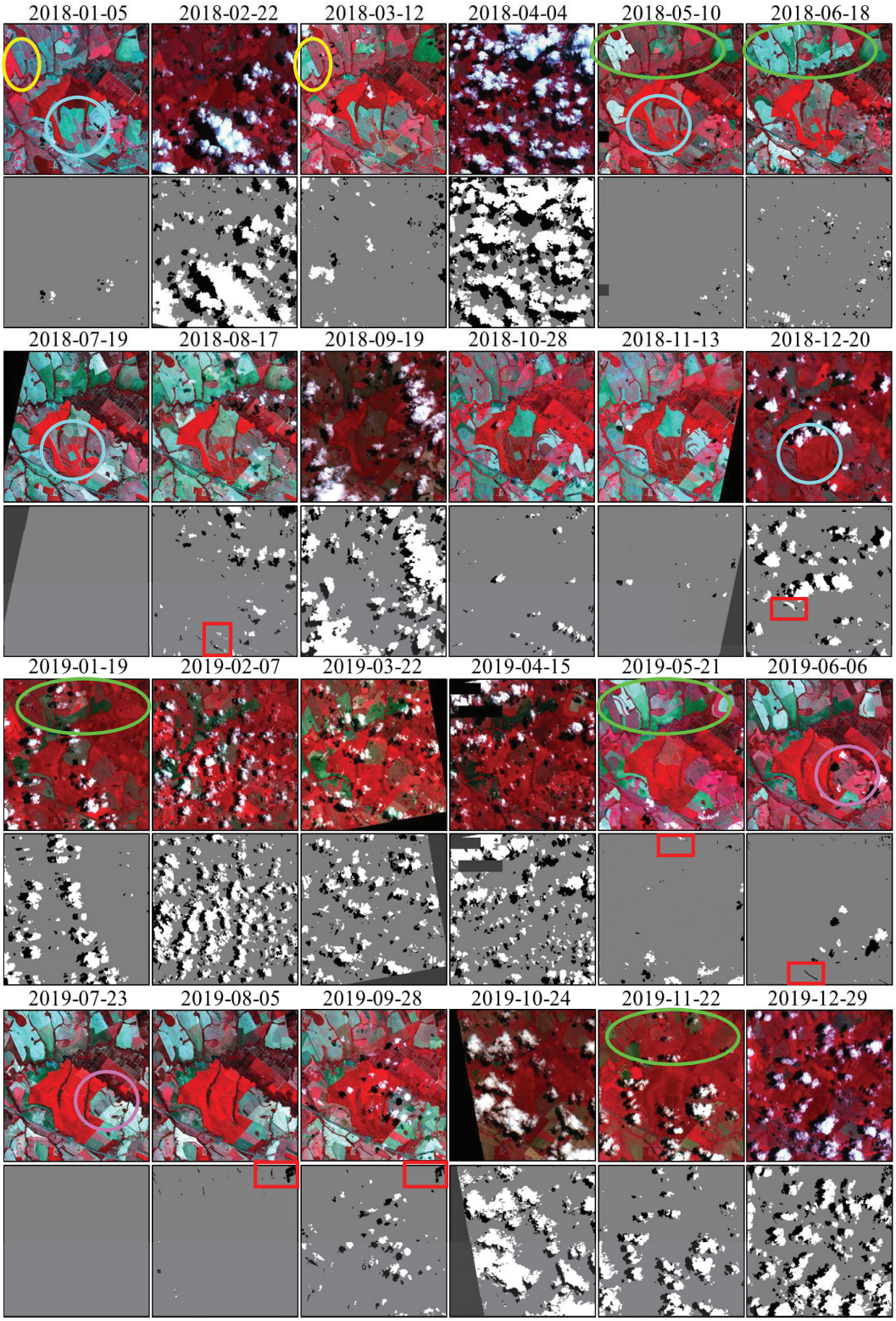


Table 1. Detailed information of the six testing sites, i.e. BCI-forest, Manaus-urban, k67-forest, Euc-plantation, Machakos-savannah, and ASM-shrubland sites, including the location, precipitation, dry season period, spatial coverage, temporal coverage, and the number of accessed PlanetScope images.

Site	Location	Precipitation (mm yr^{-1})	Dry season period	Spatial coverage (km^2)	Temporal coverage	Number of images
BCI-forest	9°09'N, 79°51'W	2600	Mid Dec-Mid Apr	6×7.26	Jan-Dec, 2018-2019	144
Manaus-urban	3°06'S, 60°01'W	2300	May-Dec	20×20		82
k67-forest	2°51'S, 54°58'W	2022	Jun-Dec	10×10		111
Euc-plantation	22°58'S, 48°44'W	1400	Mar-Oct	10×10		247
Machakos-savannah	1°36'S, 37°06'E	730	Jan-Feb, May-Oct	10×10		180
ASM-shrubland	22°17'S, 133°15'E	310	Jan-Dec	10×10		58

Table 2. Accuracy assessments (mean and standard deviation) across a total of 47 representative images in 2018-2019 at all six sites using manual masks as benchmarks. Five indices, i.e. overall accuracy (OA), producer's accuracy (PA) of shadows and clouds, and user's accuracy (UA) of shadows and clouds are used to evaluate the accuracy of cloud and shadow masks respectively derived from STI-ACSS, the PlanetScope default QCs (including UDM and UDM2), Fmask, IHOT, and ATSA.

Method	OA (%)		PA (%)				UA (%)			
			shadow		cloud		shadow		cloud	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
STI-ACSS	98.03	1.30	89.48	15.01	95.53	14.52	91.55	14.21	93.70	14.20
default QC-UDM	79.37	18.00	2.11	3.68	56.67	27.05	18.37	27.80	87.77	22.22
default QC-UDM2	87.67	8.38	30.05	19.33	83.64	16.63	90.43	11.53	75.42	17.02
Fmask	79.89	15.82	36.48	27.13	90.79	15.84	54.31	29.04	60.67	29.89
IHOT	70.38	22.65	0.00	0.00	80.44	29.09	0.00	0.00	51.20	33.06
ATSA	85.16	18.67	64.86	30.37	86.76	17.87	69.08	24.48	73.12	31.39

80 **Table 3**. Accuracy assessments of the 12 representative images in Fig. 6 using manual masks as benchmarks.
81 Five indices, i.e. OA, PA of shadows and clouds, and UA of shadows and clouds, are used to evaluate the
82 accuracy of cloud and shadow masks respectively derived from STI-ACSS, the PlanetScope default QCs
83 (including UDM and UDM2), Fmask, IHOT, and ATSA.

84

Site Date	Method	OA (%)	PA (%)		UA (%)	
			shadow	cloud	shadow	cloud
BCI-forest	STI-ACSS	99.35	90.32	99.51	92.92	92.08
2018-12-28	default QC-UDM	97.73	0.12	57.24	1.59	98.95
	default QC-UDM2	97.67	14.44	85.23	94.25	61.10
	Fmask	61.73	0.01	99.82	7.48	3.44
	IHOT	61.73	0.00	99.82	0.00	3.44
	ATSA	61.76	0.81	99.82	79.85	3.44
BCI-forest	STI-ACSS	98.47	87.10	99.38	92.85	97.83
2019-12-13	default QC-UDM	91.81	0.72	66.27	5.87	99.65
	default QC-UDM2	95.57	54.26	74.27	96.98	96.23
	Fmask	94.38	33.37	90.01	69.13	97.46
	IHOT	75.62	0.00	99.39	0.00	19.26
	ATSA	96.69	77.54	79.13	84.43	99.55
Manaus-urban	STI-ACSS	97.64	89.84	94.62	93.06	93.76
2018-06-02	default QC-UDM	88.99	0.44	78.67	10.82	88.56
	default QC-UDM2					
	Fmask	65.36	59.49	87.44	40.19	33.99
	IHOT	85.92	0.00	33.44	0.00	93.85
	ATSA	89.98	94.53	91.62	52.72	78.34
Manaus-urban	STI-ACSS	95.05	95.57	96.14	91.70	97.64
2019-08-18	default QC-UDM	40.15	0.83	30.57	39.02	77.39
	default QC-UDM2	81.13	4.48	97.71	98.99	82.44
	Fmask	76.24	37.50	92.15	57.03	81.99
	IHOT	50.11	0.00	60.53	0.00	59.25
	ATSA	80.69	45.43	89.43	74.02	83.76
k67-forest	STI-ACSS	99.41	98.54	99.82	98.23	99.05
2018-06-15	default QC-UDM	82.02	0.27	64.25	6.23	99.77
	default QC-UDM2					
	Fmask	91.44	68.08	75.69	89.59	99.20
	IHOT	26.30	0.00	99.97	0.00	18.19
	ATSA	94.18	95.71	85.29	79.64	97.81
k67-forest	STI-ACSS	99.10	96.66	99.83	98.05	99.25
2019-09-12	default QC-UDM	58.07	0.21	11.07	8.72	99.94
	default QC-UDM2	82.46	15.10	78.95	98.82	92.50
	Fmask	85.49	59.36	72.77	74.52	99.35
	IHOT	27.53	0.00	64.33	0.00	27.67
	ATSA	60.00	37.67	98.54	47.26	49.15

85

86

Site Date	Method	OA (%)	PA (%)		UA (%)	
			shadow	cloud	shadow	cloud
Euc-plantation	STI-ACSS	99.20	87.19	99.68	81.42	91.82
2018-03-12	default QC-UDM	92.58	0.13	50.37	0.05	98.69
	default QC-UDM2					
	Fmask	96.90	0.96	89.57	1.84	89.99
	IHOT	96.94	0.00	48.32	0.00	98.94
	ATSA	97.56	12.69	81.49	37.44	95.61
Euc-plantation	STI-ACSS	96.61	85.45	99.06	90.77	86.90
2019-03-22	default QC-UDM	78.85	6.77	70.53	4.32	93.49
	default QC-UDM2	83.59	22.16	94.05	69.89	49.31
	Fmask	83.96	67.83	95.76	50.62	52.82
	IHOT	19.17	0.00	99.97	0.00	11.10
	ATSA	19.23	0.02	99.97	0.48	11.15
Machakos-savannah	STI-ACSS	97.69	100	97.15	95.86	98.80
2018-02-12	default QC-UDM	50.43	0.02	86.33	73.69	53.01
	default QC-UDM2					
	Fmask	33.84	0.37	99.96	85.24	33.79
	IHOT	33.77	0.00	99.98	0.00	33.77
	ATSA	73.25	67.04	98.07	93.50	58.17
Machakos-savannah	STI-ACSS	96.66	99.99	98.88	97.33	97.14
2019-06-22	default QC-UDM	35.60	0.00	37.03	0.00	89.22
	default QC-UDM2	75.38	0.00	97.13	0.00	78.16
	Fmask	69.55	0.22	99.94	24.79	69.63
	IHOT	69.52	0.00	99.95	0.00	69.55
	ATSA	83.80	57.41	90.64	83.16	90.16
ASM-shrubland	STI-ACSS	98.23	97.78	78.40	95.05	97.88
2018-09-05	default QC-UDM	89.13	0.00	56.33	0.00	99.64
	default QC-UDM2	91.70	26.60	79.17	96.64	82.87
	Fmask	80.15	38.00	98.26	45.51	25.11
	IHOT	90.10	0.00	91.20	0.00	85.76
	ATSA	90.88	24.69	92.93	67.46	78.42
ASM-shrubland	STI-ACSS	95.25	94.77	99.23	90.72	95.85
2019-08-08	default QC-UDM	63.44	1.06	31.13	39.80	97.46
	default QC-UDM2	71.74	34.32	33.44	91.18	78.94
	Fmask	68.50	24.77	95.33	64.75	55.52
	IHOT	63.28	0.00	93.66	0.00	53.89
	ATSA	77.27	34.88	70.48	95.04	87.51

87