ISPRS Journal of Photogrammetry and Remote Sensing 183 (2022) 19–33

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Monitoring leaf phenology in moist tropical forests by applying a superpixel-based deep learning method to time-series images of tree canopies

Guangqin Song [a], Shengbiao Wu [a], Calvin K.F. Lee [a], Shawn P. Serbin [b], Brett T. Wolfe [c,d], Michael K. Ng [e], Kim S. Ely [b], Marc Bogonovich [a], Jing Wang [a], Ziyu Lin [a], Scott Saleska [f], Bruce W. Nelson [g], Alistair Rogers [b], Jin Wu [a,*]

[a] *School of Biological Sciences, University of Hong Kong, Hong Kong, China*
[b] *Department of Environmental and Climate Sciences, Brookhaven National Laboratory, Upton, NY, USA*
[c] *School of Renewable Natural Resources, Louisiana State University Agricultural Center, Baton Rouge, LA, USA*
[d] *Smithsonian Tropical Research Institute, Balboa, Republic of Panama*
[e] *Department of Mathematics, University of Hong Kong, Hong Kong, China*
[f] *Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA*
[g] *National Institute for Amazon Research (INPA), Manaus, Brazil*

ARTICLE INFO

Keywords:
Residual Networks
Semantic segmentation
Proximate remote sensing
Leaf quantity
Leaf quality
Green Chromatic Coordinate
Tropical forests
Phenocam

ABSTRACT

Tropical leaf phenology—particularly its variability at the tree-crown scale—dominates the seasonality of carbon and water fluxes. However, given enormous species diversity, accurate means of monitoring leaf phenology in tropical forests is still lacking. Time series of the Green Chromatic Coordinate (GCC) metric derived from tower-based red–greenblue (RGB) phenocams have been widely used to monitor leaf phenology in temperate forests, but its application in the tropics remains problematic. To improve monitoring of tropical phenology, we explored the use of a deep learning model (i.e. superpixel-based Residual Networks 50, SP-ResNet50) to automatically differentiate leaves from non-leaves in phenocam images and to derive leaf fraction at the tree-crown scale. To evaluate our model, we used a year of data from six phenocams in two contrasting forests in Panama. We first built a comprehensive library of leaf and non-leaf pixels across various acquisition times, exposure conditions and specific phenocams. We then divided this library into training and testing components. We evaluated the model at three levels: 1) superpixel level with a testing set, 2) crown level by comparing the model-derived leaf fractions with those derived using image-specific supervised classification, and 3) temporally using all daily images to assess the diurnal stability of the model-derived leaf fraction. Finally, we compared the model-derived leaf fraction phenology with leaf phenology derived from GCC. Our results show that: 1) the SP-ResNet50 model accurately differentiates leaves from non-leaves (overall accuracy of 93%) and is robust across all three levels of evaluations; 2) the model accurately quantifies leaf fraction phenology across tree-crowns and forest ecosystems; and 3) the combined use of leaf fraction and GCC helps infer the timing of leaf emergence, maturation and senescence, critical information for modeling photosynthetic seasonality of tropical forests. Collectively, this study offers an improved means for automated tropical phenology monitoring using phenocams.

1. Introduction

In moist tropical forests, leaf phenology is an important control on the seasonality of carbon and water fluxes from local to regional scales (Manoli et al., 2018; Restrepo-Coupe et al., 2013, 2017; Wu et al., 2016; Yang et al., 2018). In contrast to temperate regions where phenology is driven by large seasonality in climate (Piao et al., 2019; Richardson et al., 2009, 2018), moist tropical forests have seasonally warm and humid climates and a rich assemblage of tree species, resulting in leaf phenology that is diverse, complex, and poorly understood (Albert et al., 2019; Lopes et al., 2016; Reich, 1995; Wang et al., 2020). Thus, an improved characterization of leaf phenology in moist tropical forests

---

* Corresponding author at: School of Biological Sciences, University of Hong Kong, Pokfulam Road, Hong Kong, China.
E-mail address: jinwu@hku.hk (J. Wu).

https://doi.org/10.1016/j.isprsjprs.2021.10.023
Received 26 April 2021; Received in revised form 18 October 2021; Accepted 29 October 2021

across scales, from individual tree-crowns to forest ecosystems, is critically needed to help understand a cryptic tropical leaf phenology and reveal how tropical forests are responding to a changing climate.

Proximate remote sensing represents the most promising method of providing the data required to characterize tropical leaf phenology across various scales. Field-based *in-situ* visual assessments are the most accurate method for monitoring tropical phenology (Brando et al., 2010; Bush et al., 2017; Singh and Kushwaha, 2016), but are time- and labor-intensive, and often limited to very small geographical coverage. Satellite remote sensing with global coverage and frequent revisits provides a potential solution for characterizing tropical leaf phenology over large areas (Guan et al., 2015; Huete et al., 2006; Tang and Dubayah, 2017). However, most satellite observations are at spatial resolutions (10 m and greater) that are too coarse to resolve tropical leaf phenology at the individual tree-crown level (Alberton et al., 2014; Berra et al., 2019). At the same time, due to frequent cloud and aerosol contamination (Samanta et al., 2010), cloud-free satellite data at fine temporal resolution are challenging to obtain, particularly in the moist tropics during wet seasons, resulting in large uncertainties in leaf phenology patterns derived using this data (Gonçalves et al., 2020; Nagai et al., 2016; Wang et al., 2020). Proximate remote sensing observations, such as tower-mounted phenocams and unoccupied aerial vehicles (UAVs), provide an important means for monitoring tropical phenology, with lower labor and time requirements compared to field observations, while offering higher spatial and temporal resolution than most satellite observations (de Moura et al., 2017; Lopes et al., 2016; Park et al., 2019). Consequently, recent studies have used proximate remote sensing to characterize tropical leaf phenology at the scales of both tree individuals and forest ecosystems (e.g. Alberton et al., 2017; Lopes et al., 2016; Park et al., 2019).

To obtain leaf phenology from these proximate remote sensing observations, the most commonly used method relies on information extracted from the red–greenblue (RGB) bands of the images to derive the Green Chromatic Coordinate (GCC) metric (Richardson et al., 2009, 2018). However, its application in moist tropical forests remains challenging for two main reasons. First, the GCC metric lacks a clear biophysical meaning, as it varies based on changes in both leaf quantity (i.e. canopy leaf fraction) and quality (i.e. leaf age and age-associated biochemistry) (Chavana-Bryant et al., 2017; Wu et al., 2017, 2018; Yang et al., 2014). Second, moist tropical forests are composed of tree species with diverse leaf phenology patterns, including evergreen, deciduous, and semi-deciduous leaf phenology habits (Detto et al., 2018; Eamus et al., 1999). Due to the seasonal variability in GCC that can be consistently observed in evergreen, deciduous, and semi-deciduous trees, it is difficult to automatically differentiate the leaf phenology habits across diverse tropical tree species with only GCC (Lopes et al., 2016).

To resolve the above limitations, canopy leaf fraction (i.e., the proportion of a given tree-crown that is displaying leaves) has recently been advocated as an alternative to GCC for tropical leaf phenology monitoring (Park et al., 2019). Specifically, machine learning approaches have been used to derive leaf fraction from proximate remote sensing observations, and in Park et al. (2019), it was shown to be successful for extracting the crown-scale leaf fraction phenology from UAV-based imagery in a tropical forest in Panama over an annual cycle. However, several issues remain with this method. First, it requires leaf and non-leaf features to be manually extracted for model training and was only tested at one site. Thus, the generalizability of the method across different sites remains unknown. Second, the machine learning results were evaluated only with reference to human visual estimations, which often vary largely among interpreters. Last, as the initial attempt was conducted using UAV-based time-series images, it is unknown whether a similar approach can be applied to fixed-location phenocam images, another important proximate remote sensing means for tropical phenology monitoring.

In this study, our main goal was to develop a rigorous method to extract leaf fraction phenology from standard RGB phenocam images that could be applied in highly diverse tropical forests. We did this through the application of a state-of-the-art superpixel-based deep learning approach to the phenocam time-series images. The superpixel here refers to regions of spectrally similar and spatially contiguous pixels within phenocam imagery. We used a superpixel-based deep learning approach for two reasons. First, deep learning approaches, especially Convolutional Neural Networks (CNNs), can automatically extract high-level features and uncover complex and hierarchical relationships from the input data, outperforming traditional machine learning methods, including the stochastic gradient tree boosting (Park et al., 2019), and other methods such as random forests and support vector machines (Cai et al., 2018; Kattenborn et al., 2021; Lecun et al., 2015; Tong et al., 2020). Second, superpixel-based deep learning approaches rely on superpixels as the functional units, greatly improving computational efficiency and reducing data redundancy of the high spatial resolution images (Gong et al., 2017; Mi and Chen, 2020; Zhao et al., 2017). We tested the accuracy and scalability of the superpixel-based deep learning approach by first differentiating leaf and non-leaf pixels from phenocam image time series and then assessing leaf fraction phenology at scales from individual trees to forest ecosystems.

We used a full year of data from six phenocams across two tropical forests in Panama spanning a large rainfall gradient. We first manually constructed a comprehensive library of leaf pixels (leaves of various ages, and exposure conditions) and non-leaf pixels (sky, canopy gaps, buildings, and other backgrounds) from phenocam images. We divided the library into training and testing components. We next built the model with the training component and rigorously evaluated the accuracy and robustness of the model at three levels: superpixel, crown, and diurnal. At the superpixel level, we evaluated the accuracy of the model with the testing component; at the crown level, we compared the model-derived leaf fractions to those derived from the image-specific supervised classifications; and at the diurnal level, we assessed the diurnal stability of model-derived leaf fraction. Lastly, we cross-compared the phenology of our model-derived leaf fraction with GCC and assessed the potential advantages of the combined use of both leaf fraction and GCC to monitor tropical leaf phenology across various scales.

2. Study sites and data

2.1. Study sites

We used two lowland seasonal moist tropical forests in the Republic of Panama as our study sites (Fig. 1). These two sites include one seasonally dry forest in the Natural Metropolitan Park (PNM; 8.9950°N, 79.5431°W) in Panama City and one wet evergreen forest in the San Lorenzo Protected Area (SLZ; 9.2810°N, 79.9745°W) near the Caribbean entrance of the Panama Canal. At these two sites, the Smithsonian Tropical Research Institute maintains two canopy crane towers with heights of 42 m at PNM and 52 m at SLZ. Within the 1-hectare area of each tower footprint, the PNM site includes treetops with 80 tree and liana species and the SLZ site includes more than 120 tree and liana species. The mean annual air temperature from 1998 to 2015 was 26.3 °C and 25.8 °C at PNM and SLZ, respectively. The mean annual precipitation from 1998 to 2015 was 1826 mm yr^{-1} and 3286 mm yr^{-1} at PNM and SLZ, respectively, with a 4-month long dry season (precipitation $<$ 100 mm per month; from January to April) at both sites. For more details about the forest composition and climate conditions of these two sites, please refer to Wright et al. (2003) and Wu et al. (2019).

2.2. Data

For both sites, three standard RGB phenocams, manufactured by Moultrie Wingscapes TimelapseCam Pro (Wingscapes, Inc., Alabaster, AL, USA), were mounted on the canopy crane structures pointing to different compass directions. These phenocams are hereafter referred to

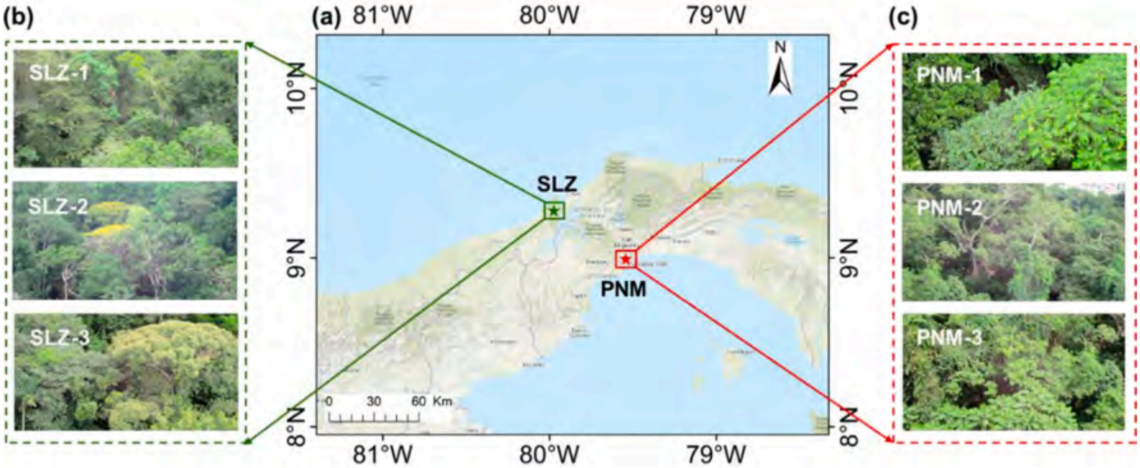

Fig. 1. Study sites and phenocam data. **(a)** Location of the two study sites, namely the San Lorenzo Protected Area (SLZ, green star) and the Natural Metropolitan Park (PNM, red star). The background figure in panel **(a)** is adapted from National Geographic, ESRI. Images from a full annual cycle from each of six phenocams were used, as labelled in **(b)** SLZ-1, SLZ-2, and SLZ-3 and in **(c)** PNM-1, PNM-2, and PNM-3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as PNM-1, PNM-2, PNM-3, and SLZ-1, SLZ-2, SLZ-3 (Fig. 1). Each phenocam was programmed to automatically acquire images at a 10-min interval from 10:00 AM to 1:50 PM local time every day. The phenocam images were saved in the JPG format (8-bit per channel and a total of 3 bands) with an interpolated resolution of 6080 × 3420 pixels. The images used in this study covered a full annual cycle from December 3, 2016 to December 31, 2017. Since the field of view (FOV) of the phenocams would change slightly when the batteries were changed every 2–6 months (given that battery access was blocked by the tower mounts), we only focused on trees that consistently appeared in the phenocam images for at least ten months. We manually generated the tree-crown masks for 35 consistent trees and 1 liana (PNM-3–2) in our study, including 12 trees and 1 liana at PNM and 23 trees at SLZ (Fig. 2). Full resolution images are available from the NGEE-Tropics data archive, where phenocams named here as PNM-1, PNM-2, PNM-3, and SLZ-1, SLZ-2, SLZ-3 correspond to camera locations PNM-C, PNM-E and PNM-H, and SLZ-H, SLZ-J and SLZ-F, respectively (Wu et al., 2021).

3. Methods

Our workflow can be divided into four main steps: 1) building a superpixel-based Residual Networks 50 (SP-ResNet50) deep learning model by integrating the convolutional neural network Residual Networks 50 (ResNet50) model with a superpixel-based segmentation algorithm; 2) training and evaluating the SP-ResNet50 model with the comprehensive superpixel library of leaves and non-leaves extracted from time-series phenocam images across three levels; 3) comparing the SP-ResNet50-derived leaf fraction time series with the GCC time series at the tree-crown scale; and 4) exploring ecosystem-scale phenology of leaf fraction across the two forest sites. Semantic segmentation methods have been used for similar purposes in other deep-learning-based studies (e.g. Zhang et al., 2020), so we also compared our SP-ResNet50 results with the commonly used semantic segmentation method, U-Net (Ronneberger et al., 2015), through which we hope to understand whether the model selection would affect our results and to make recommendations for future similar efforts. A flowchart of the workflow is shown in Fig. 3.

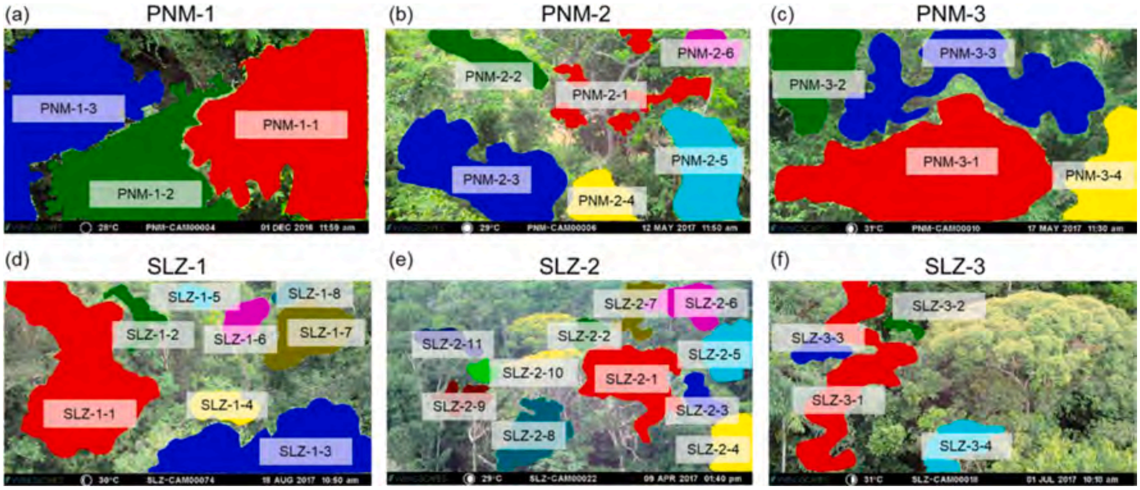

Fig. 2. Manual segmentation of upper canopy tree-crowns within the field of view of each phenocam. Each colored polygon indicates a manually identified canopy tree-crown, with corresponding tree ID, which consistently showed up in the phenocam images for at least ten months. 1 liana (PNM-3–2) and a total of 12 trees were included from the PNM site: **(a)** 3 from PNM-1; **(b)** 6 from PNM-2; and **(c)** 3 from PNM-3. A total of 23 trees were included from the SLZ site: **(d)** 8 from SLZ-1; **(e)** 11 from SLZ-2; and **(f)** 4 from SLZ-3.

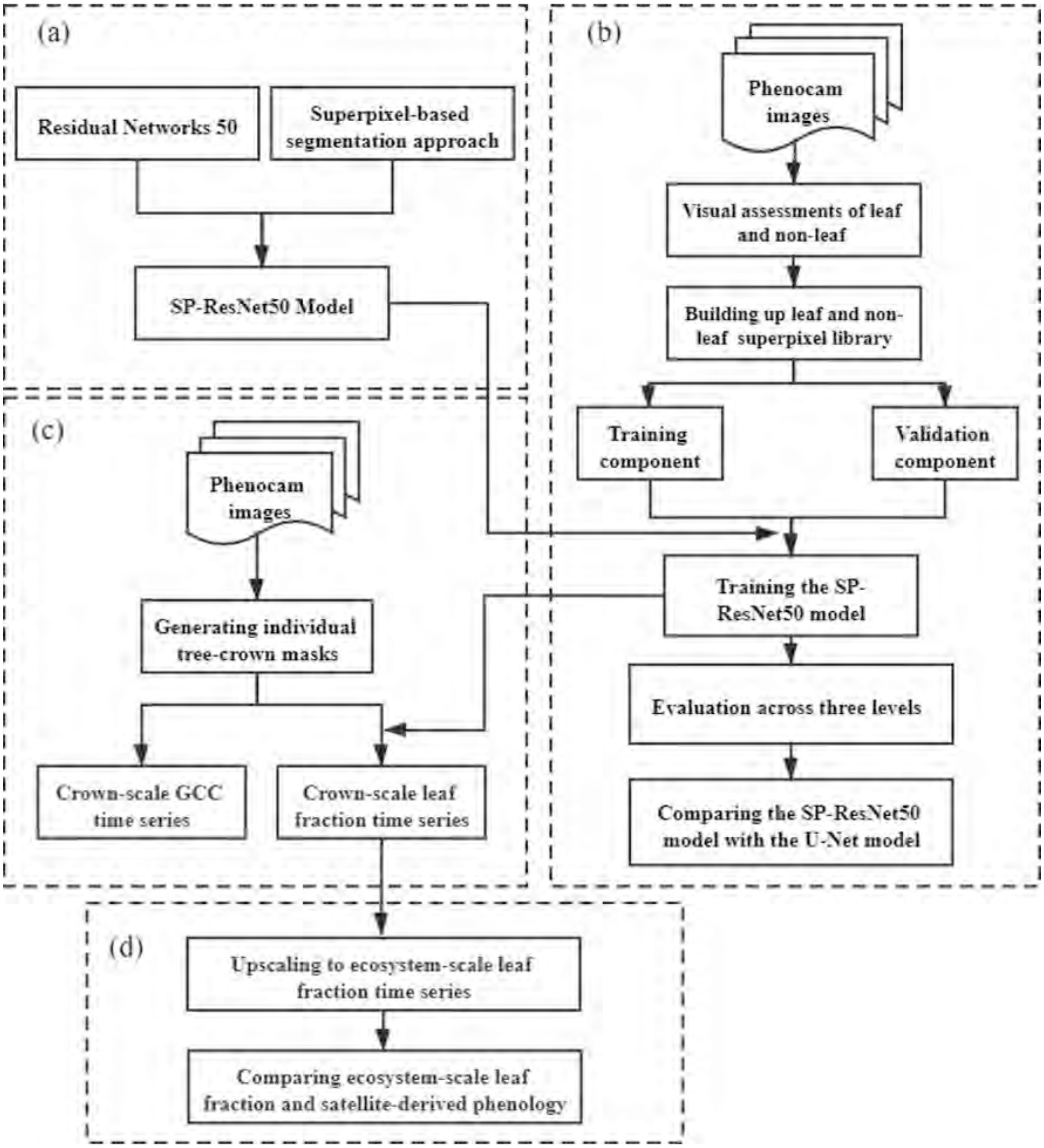

Fig. 3. Flowchart of our method. **(a)** Building up the SP-ResNet50 model by integrating the Residual Networks 50 model with a superpixel-based segmentation approach; **(b)** training and evaluating the SP-ResNet50 model with the leaf and non-leaf superpixel library from phenocam images timeseries across three levels, and cross-comparing the SP-ResNet50 model results with those from the U-Net model; **(c)** cross-comparing the leaf phenology metrics between the SP-ResNet50-derived leaf fraction and GCC; and **(d)** exploring ecosystem-scale phenology of leaf fractions across two sites.

3.1. Introducing the SP-ResNet50 model

We applied a SP-ResNet50 model (He et al., 2016) to differentiate leaves from non-leaves in phenocam images. The SP-ResNet50 model consists of two components. First, we employed a superpixel-based segmentation approach to generate superpixels that represent a group of spectrally similar and spatially contiguous pixels from the phenocam images. This was done because our phenocam images include billions of image pixels (six phenocams taking images containing 3240 × 6080 pixels each over a year), creating large computational challenges if a pixel-based image classification is used. Using the superpixel approach improves computational efficiency as it merges spatially close pixels with spectral similarity into a single representative unit. Additionally, previous studies that used the superpixel-based segmentation approach for remote sensing image classifications demonstrated an increase in classification accuracy along with increased computational efficiency (Chen et al., 2019; Gong et al., 2017; Zhao et al., 2017).

Second, we used the ResNet50 model as the classifier for the superpixel level image classifications as it has been shown to improve both the representation of hierarchical characteristics in remote sensing data and classification accuracy when compared to traditional machine learning methods (Tao et al., 2017).There are three key layers in the ResNet50 model, namely the convolutional layer, pooling layer, and fully-connected layer. The convolutional layer works by extracting image features with convolutional kernels, while the pooling layer down-samples these extracted image features to avoid overfitting. The fully-connected layer serves as the classifier to generate predictive probabilities of the image classification using the image features which are created by the convolutional layer and the pooling layer. In addition, the ResNet50 model uses the residual connections to improve its accuracy with a much deeper network. For details about the ResNet50 model, see He et al. (2016).

3.2. Training and evaluating the SP-ResNet50 model

We separated the model training and evaluation into three sub-steps: 1) building a comprehensive superpixel library; 2) dividing the resulting library into training and testing components; and 3) training the model from scratch using the training component then evaluating the model with the testing component across three levels.

3.2.1. Building the superpixel library

3.2.1.1. Labelling phenocam image pixels as leaf and non-leaf. We adopted a two-class approach to analyze the phenocam images. Since previous studies have shown that the representativeness of the training data directly affects the deep learning model performance (Lv et al., 2019), we selected leaf and non-leaf samples that comprehensively covered various conditions, including leaf pixels of different leaf ages (young, mature, and senescent; Lopes et al., 2016) and exposure conditions (partially shaded, sunlit, and over-exposed), and non-leaf pixels that included branches, sky, canopy gaps, buildings, and others. Collectively, we labelled the leaf and non-leaf pixels across the phenocam images that were acquired at different times of day (representing various sky conditions and sun angle effects) and across the six different phenocams and the two forest sites.

Our labelling of leaves and non-leaves included the following two steps. First, in order to make our library representative both within a day and across the seasons and sites, we selected all high-quality images (i.e., excluding rainy and foggy images) from all six phenocams on the 15th day of each month throughout the year. Second, for each selected image, we manually identified leaf and non-leaf classes for an average of c. 90,000 pixels per image in ENVI (Research Systems, Inc., Boulder, CO) to represent both classes within each image. To minimize the bias of our library, we adopted the following two criteria for our labelling: 1) the chosen samples were randomly and evenly distributed across the whole phenocam image; and 2) the percentage area of selected pixels from each image were around 0.4% of each other. Altogether, we generated around 155 million pixels from a total of 1728 phenocam images across the entire year with labels of leaves or non-leaves.

3.2.1.2. Generating the library of superpixels. To build the superpixel library, we employed a simple linear iterative clustering (SLIC) algorithm (Achanta et al., 2012) to generate superpixel sets from the manually labelled leaf and non-leaf pixels as described in Section 3.2.1.1. SLIC generates relatively uniform and compact superpixel objects using a k-means cluster method (Achanta et al., 2012). The key parameter for this algorithm is the superpixel number (K), which is calculated from the ratio between the total image pixel number (N; $N = 3240 \times 6080$ for our case) and the pre-defined superpixel size (S; number of pixels within one superpixel), i.e. $K = N / S$. The SLIC-derived superpixels were then labelled as leaves or non-leaves based on the class that was dominant (greater than50% of phenocam pixels). To convert the irregularly shaped superpixels into square image patches that can be used by the SP-ResNet50 model, we used an enclosing square approach (Fig. 4; Lv et al., 2019). Here, each superpixel was converted to a square image patch of a specific size (W) with the same geometric center and same assigned label as its contained superpixel. These square image patches formed the basic unit of the superpixel library.

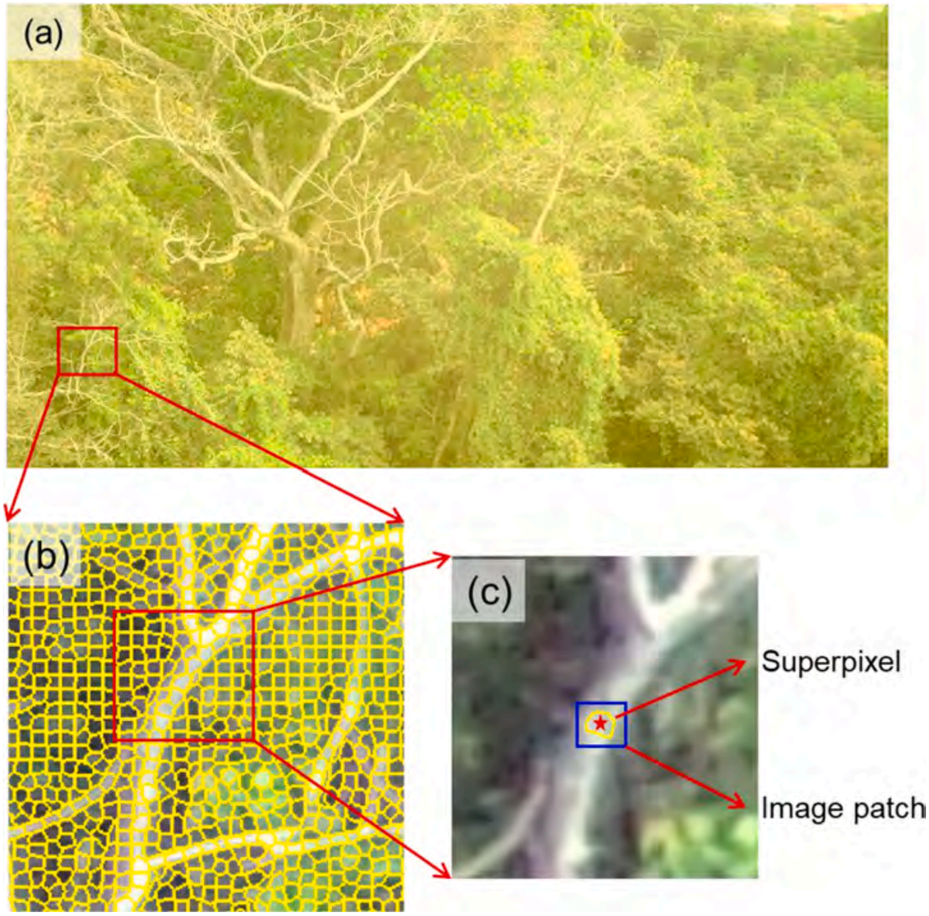

Fig. 4. Example workflow for superpixel library generation. **(a)** The phenocam image is segmented with SLIC algorithm and the superpixels are shown as polygons with yellow borders; **(b)** A closer view of the superpixels in example region of the segmented phenocam image; **(c)** An example of how a single image patch is generated from a superpixel: blue square represents the image patch, yellow irregular polygon is the superpixel, and the red star indicates the geometric center of the superpixel. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.2. *Fine-tuning the optimal parameters for the SP-ResNet50 model*

Two key parameters are needed when generating the superpixel library: S, the number of pixels within one superpixel; and W, the size of the associated square patch. A larger S reduces data redundancy but risks including undesired pixel types within each superpixel, thus affecting the model accuracy. Similarly, a larger W allows more feature information to be extracted (Chen et al., 2019), and thus improves the model classification accuracy, but risks including redundant feature information of non-target pixels and decreases model accuracy and efficiency. To determine the optimal paired values of these two parameters, we conducted a sensitivity analysis. We randomly divided the superpixel library into training (80%) and testing (20%) sets, trained the model using the training set, predicted the classes of the testing set, and evaluated the models using the overall accuracy (OA; Eq. (1)):

$$OA = \frac{Tp + Tn}{Nsp} \times 100\% \tag{1}$$

where Tp and $Tn$are the number of leaf and non-leaf superpixels that are correctly predicted by the model, and Nsp is the total number of superpixels in the testing set.

We conducted the sensitivity analysis in two steps. First, to obtain the optimal image patch size W, we selected a range of patch sizes from 5 × 5 to 61 × 61 pixels at increments of 4 × 4 pixels, with the superpixel size S fixed to 50 pixels based on previous research (Achanta et al., 2012). We found that OA would stabilize when W reached 25 × 25 pixels, and thus we chose W of 25 × 25 pixels as the optimal image patch size (Fig. 5a). Second, with W fixed at 25 × 25 pixels, we tested a range of superpixel sizes from 50 to 100 pixels in increments of 10 pixels. The sensitivity analysis demonstrated that S = 70 pixels resulted in the highest OA (Fig. 5b). Therefore, we used W = 25 × 25 pixels and S = 70 pixels in our subsequent SP-ResNet50 model for phenocam images classification, which consequently generated around 5.32 million image patches, including around 2.64 million patches at the PNM site and 2.68 million patches at the SLZ site.

3.2.3. *Dividing the superpixel library into training and testing components*

Spatial and temporal independence is essential when splitting the library of superpixels into training and testing components. To ensure spatial independence, we randomly separated the superpixel library of each phenocam according to its 2-D image space, with 80% of the superpixels allocated to training sets and the remaining allocated to testing sets. To ensure temporal independence, for each phenocam, we used the superpixels from odd months for model training and even months for model testing.

3.2.4. *Evaluating the SP-ResNet50 model*

We trained the SP-ResNet50 model with the training sets created above, then evaluated the model's performance using the testing sets at three different levels. First, we evaluated the model at the superpixel level with four tests. In test-1, the model was trained using the training sets from four phenocams (PNM-1, PNM-2, SLZ-1, and SLZ-2) and evaluated using the testing sets from the remaining phenocams (PNM-3 and SLZ-3). In test-2, the model was trained using the training sets from all six phenocams and evaluated using the same testing sets as test-1 from two phenocams (PNM-3 and SLZ-3). In test-3, the model was trained using data from odd months across all six phenocams and evaluated using the data from even months across all six phenocams. In test-4, the model was evaluated with five-fold cross-validation, which is a commonly-used validation method in image classifications (Ma et al., 2021; Zhang et al., 2018). Specifically, the full library of superpixels were one-time randomly divided into five equal groups, where any of the four groups were used to train the model while the fifth group was used to evaluate the model, and the process repeated until every group has been used for evaluation. Afterwards, the mean accuracy of all these iterations were calculated for final model performance evaluation. The detailed numbers of the training and testing sets of the above four tests are shown in Table S1. In addition to OA, we also used the producer's accuracy (PA) of both leaves and non-leaves, and the user's accuracy (UA) of both leaves and non-leaves to evaluate the model (Congalton, 1991). PA_{leaf} shown in Eq. (2) refers to the percentage of leaf superpixels that are correctly predicted by the SP-ResNet50 model in the total number of leaf superpixels in the testing sets, and UA_{leaf} shown in Eq. (3) is the percentage of correct classifications of leaf superpixels in all superpixels predicted as leaves by the SP-ResNet50 model in the testing sets.

$$PA_{leaf} = \frac{Tp}{Tp + Fn} \times 100\% \tag{2}$$

$$UA_{leaf} = \frac{Tp}{Tp + Fp} \times 100\% \tag{3}$$

where Tp is the number of leaf superpixels that are correctly classified, Fn is the number of leaf superpixels that are falsely classified as non-leaf superpixels, and Fp is the number of non-leaf superpixels that are falsely classified as leaf superpixels.

Second, we evaluated the model performance at the crown level by comparing the SP-ResNet50 results with those derived from image-specific supervised classifications. This is because image-specific supervised classifications with more training samples should generate more accurate classification results for each selected image, and thus can be used as a benchmark to evaluate the model performance of SP-ResNet50 that applies to all images across different phenocams and sites. With the classification results of leaves and non-leaves and the manually created tree-crown masks (Fig. 2), we further derived the leaf

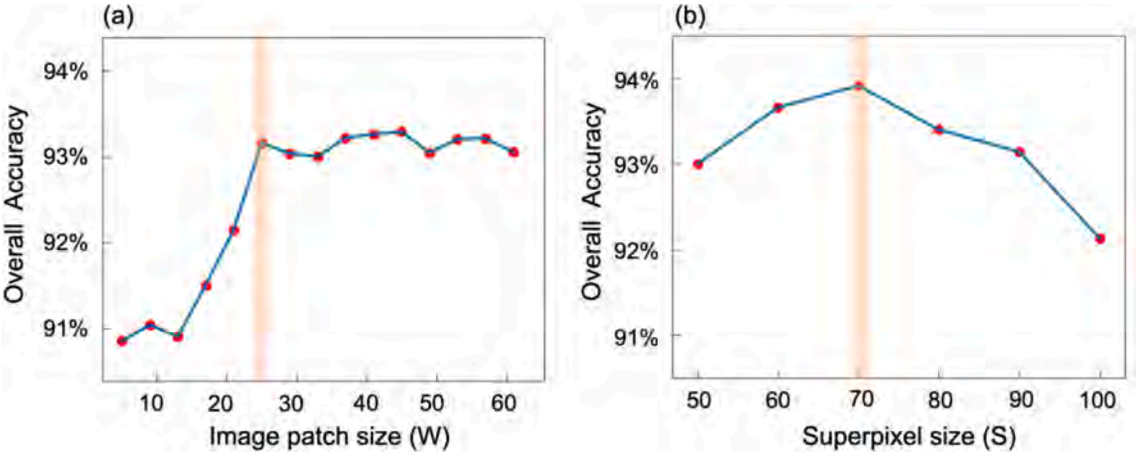

Fig. 5. The influence of **(a)** image patch size (W) and **(b)** superpixel size (S) on the overall accuracy of the SP-ResNet50 model based on the testing set.

fraction of each individual tree-crown, using Eq. (4).

$$leaf\ fraction = \frac{N_{leaf}}{N_{crown}} \tag{4}$$

where N_{leaf} and N_{crown} are leaf pixel number and all pixel number within a tree-crown, respectively.

For this assessment, we selected three images from each phenocam, one in the early leaf falling stage, one in the peak leaf falling stage, and one in the early leaf emergence stage (specific dates shown in Fig. S1) to represent a large gradient in crown-scale leaf fraction. For each image, we manually labelled around 890,000 pixels (roughly 10 times more than the labelled data per image involved in the SP-ResNet50 model) as leaf or non-leaf using the same method presented in Section 3.2.1.1. We then used this expanded training data to train a supervised random forest model (Breiman, 2001) on a per image basis. Next, we overlaid the classification results with the manual tree-crown masks to derive leaf fraction for each tree-crown (Fig. 2). Lastly, we cross-compared these derived crown-scale leaf fractions with those derived from the SP-ResNet50 model, using two common goodness-of-fit metrics, R^2 and root-mean-square-error (RMSE).

Third, we evaluated the model at the diurnal level by assessing the stability of the model-derived crown-scale leaf fractions across the images acquired within a day. For each phenocam, we selected diurnal images from four different phenological stages for the majority of tree-crowns within each phenocam's field of view: one in the early leaf falling stage, one in the peak leaf falling stage, one in the early leaf emergence stage, and one in the peak leaf emergence stage (specific dates shown in Fig. S2). We then applied the SP-ResNet50 model to these images and extracted the leaf fractions of each tree-crown. The leaf fraction for each tree-crown should not change much within a day, so the model-derived crown level leaf fractions in a single day's set of diurnal images are expected to be roughly constant if the SP-ResNet50 model is robust. We used the coefficient of variation (CV), defined as the ratio between the standard deviation and the mean value of the model-derived leaf fractions over a single day, to assess the robustness of the SP-ResNet50 model. For every identified tree-crown, we calculated the crown-specific leaf fraction CV across all four phenological stages.

3.2.5. *Cross-comparing the SP-ResNet50 model results with a semantic segmentation model*

In addition to the SP-ResNet50 model, another possible way of classifying leaf and non-leaf is by using semantic segmentation, which automates annotation of specific objects in the image and has shown great success in the computer vision and remote sensing fields (Kattenborn et al., 2021; Kotaridis and Lazaridou, 2021; Wagner et al., 2019). In order to assess whether the model selection would affect our results, we compared results from the SP-RestNet50 with those from a U-Net model, which is a popular semantic segmentation algorithm (Zhang et al., 2020). Specifically, U-Net adopts an encoder-decoder structure to address the semantic segmentation tasks (Ronneberger et al., 2015). The encoder part includes a standard CNN, which aims to extract the representative feature maps from the input image, while the decoder part uses the deconvolutions to up-sample the feature maps to then recover the original resolution. Additionally, U-Net includes several skip connections between the encoder and decoder, which help improve the decoder's accuracy in semantic segmentation. For more details of the U-Net model refer to Ronneberger et al. (2015). In this study, we used the VGG16 network, a well-known CNN structure in computer vision (Simonyan and Zisserman, 2014) as the encoder, due to its high computational efficiency and accuracy in performing image classification tasks (Ding et al., 2018; Zan et al., 2020).

We used the sparsely and randomly labelled pixels of leaf and non-leaf on the phenocam images generated in Section 3.2.1.1 as the training data for the U-Net model. Firstly, we generated benchmark images, where the pixels labelled as leaf or non-leaf were set as specific categories and the other pixels were set as the null value. Then, we cropped the benchmark and phenocam images simultaneously into patches of 512 × 512 pixels, which can ensure high computational efficiency and achieve a high segmentation accuracy (Du et al., 2021; Hussein et al., 2021; Jiao et al., 2020). Finally, we obtained roughly 130,000 samples to train the U-Net model. We only considered the labelled pixels and ignored the pixels with the null value when training the U-Net model. After training the model, we also evaluated the model accuracy at three levels (pixel, crown, and diurnal) following the similar process as Section 3.2.4, and then compared the performances of the SP-ResNet50 model and the U-Net model by quantitative accuracy assessment and visual inspection.

3.3. *Cross-comparing the SP-RestNet50-derived leaf fraction with GCC*

To examine the potential advantages of using leaf fraction alone and a combination of leaf fraction and GCC for tropical leaf phenology monitoring, we conducted the following three explorations.

First, we analyzed the relationship between leaf fraction and GCC at the tree-crown scale. Specifically, we first extracted the GCC time series of each individual tree-crown for all six phenocams (Fig. 2). For each tree-crown, we then followed Richardson et al. (2018) and Sonnentag et al. (2012), and calculated the 90th percentile values of GCC within a 3-day moving window to represent daily values, minimizing potential confounding effects from poor weather conditions, cloud/aerosol contaminations, and large shifts in solar illumination geometry. Further, we calculated the monthly mean value of GCC for each tree-crown from this GCC time series and then cross-compared it with the monthly mean value of SP-ResNet50-derived leaf fraction.

Second, we cross-compared four key phenological dates derived using leaf fraction with those derived using GCC. As evergreen trees maintain a stable leaf fraction across the year, we only derived the key phenological dates for the semi-deciduous and deciduous trees within our study sites. The phenology curve for deciduous and semi-deciduous trees follows a double logistic curve, with the phenology curve from previous complete leaf fall to the new leaf flushing and then complete leaf maturity following a typical logistic curve (or leaf flushing phenology) and from the complete leaf maturity to the next round complete leaf fall following a second logistic curve but in a reverse patter (or leaf fall phenology) (Ratkowsky, 1983; Zhang et al., 2003). Therefore, for each deciduous or semi-deciduous tree crown, we fitted the corresponding seasonal dynamic of GCC/leaf fraction for leaf flushing phenology and leaf fall phenology, respectively, using a standard logistic curve model (Eq. (5)), following Zhang et al. (2003). We accordingly derived four phenological dates per crown, including the start of leaf emergence (SOLE), end of leaf emergence (EOLE), start of leaf falling (SOLF), and end of leaf falling (EOLF). Also see Fig. 6 for an example.

$$y(t) = \frac{c}{1 + e^{a+b\times t}} + d \tag{5}$$

where t is the time in day of year (DOY) and $y(t)$is either the leaf fraction or the GCC at time t, a and b are fitting coefficients, $c + d$is the maximum value of a given leaf phenology curve, and d is the minimum value of the given leaf phenology curve.

Third, we explored the potential of the combined use of leaf fraction and GCC for tropical leaf quality phenology monitoring. Critical leaf aging processes such as leaf emergence, maturation and senescence are tightly connected with the seasonal variability in leaf biochemical and physiological properties (Chavana-Bryant et al., 2017; Wu et al., 2017), offering important measures of leaf quality (Albert et al., 2018; Wu et al., 2016). Measuring change in the leaf fraction of a crown is insufficient to identify these critical leaf aging processes, while GCC is sensitive to changes in leaf color related to leaf age. Thus, we integrated the time series of both phenology metrics at the tree-crown scale and investigated the potential to infer the critical leaf aging processes with the combined use of these two metrics.

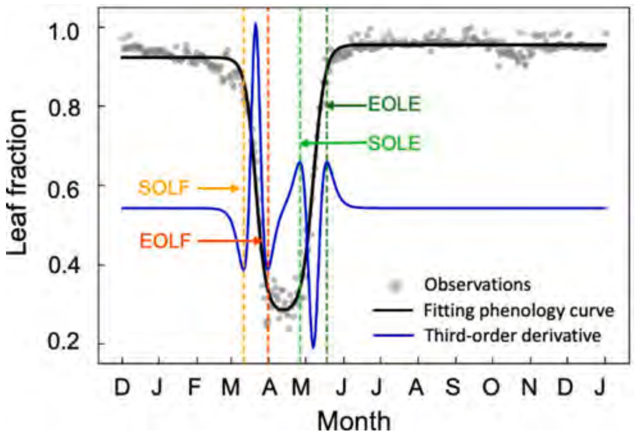

Fig. 6. Example of deriving the four leaf phenological dates from the tree-crown leaf fraction time series (gray dots), using a double logistic model (black line). The third-order derivative (blue curve) of the fitted leaf phenology curve is used to derive the four leaf phenology dates, following Zhang et al. (2003). These phenological dates include the start of leaf emergence (SOLE; Light green), end of leaf emergence (EOLE; Dark green), start of leaf falling (SOLF; Yellow) and end of leaf falling (EOLF; Orange). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. Exploring ecosystem-scale leaf fraction phenology

We used the crown level leaf fraction time series derived from the SP-ResNet50 model to explore ecosystem-scale patterns across the two sites (PNM and SLZ) in Panama. For each site, we first calculated the monthly mean leaf fraction of each tree-crown. We then upscaled the crown-scale leaf fraction phenology to the ecosystem-scale by averaging the monthly values across all the tree-crowns within the same site. We next used the MODIS BRDF/Albedo model parameter product (MCD43A1; Schaaf et al., 2002) and downloaded the daily product with a 500-meter resolution that covers the twenty-years from 2000 to 2019. Following Wang et al (2020), we then extracted the mean seasonality of BRDF-corrected enhanced vegetation index (EVI) of a 1 km × 1 km area, respectively for PNM and SLZ sites, and compared them with the ecosystem-scale seasonality of leaf fractions derived from phenocam images.

4. Results

4.1. Accuracy and robustness assessments of the SP-ResNet50 model

The model performance of SP-ResNet50 was evaluated at three levels: 1) at the superpixel level, testing the spatial and temporal robustness of the SP-ResNet50 classifications (Table 1); 2) at the crown level, testing the accuracy of the SP-ResNet50-derived leaf fraction (Fig. 7 and S3); and 3) at the diurnal level, testing the stability of SP-ResNet50-derived crown-scale leaf fraction within a day (Figs. 8 and 9).

In the first assessment, we showed that the SP-ResNet50 model accurately classified leaf and non-leaf superpixels in all groups of the spatially (different phenocams) and temporally (odd against even months) independent testing sets, as well as the five-fold cross-validation, with OAs ranging from 90.62% to 94.81% (Table 1). Specifically, with the spatially independent testing sets, the SP-ResNet50 model trained by four phenocams (PNM-1, PNM-2, SLZ-1, and SLZ-2) accurately classified leaf and non-leaf superpixels of the other two phenocams (PNM-3 and SLZ-3; OA = 92.89%; PA = 95.61%, UA = 96.50% for leaf fraction; PA = 87.16%, UA = 80.96% for non-leaf fraction; Table 1a). Model accuracy was further improved when data from all phenocams were used (OA = 94.81%, PA = 96.46%, UA = 97.64% for leaf fraction; PA = 90.50%, UA = 86.35% for non-leaf fraction; Table 1b). With the temporally independent assessment, the SP-ResNet50 model trained by the data from odd months showed good performance when applied to the data from the even months (OA = 90.62%; PA = 92.84%, UA = 95.81% for leaf fraction; PA = 86.65%, UA = 78.17% for non-leaf fraction, Table 1c). Moreover, the model also achieved good performance when evaluated with the five-fold cross-validation (OA = 95.28%; PA = 96.81%, UA = 97.08% for leaf fraction; PA = 90.10%, UA = 89.22% for non-leaf fraction, Table 1d).

Table 1
Accuracy assessments of the SP-ResNet50 model based on the independent testing sets, including that **(a)** the model was trained using the training sets from four phenocams (PNM-1, PNM-2, SLZ-1, and SLZ-2) and evaluated using the testing sets from the two remaining phenocams (PNM-3 and SLZ-3), **(b)** the model was trained using the training sets of all six phenocams and evaluated using the same independent testing sets as **(a)**, **(c)** the model was trained using all the data from odd months of all six phenocams and evaluated with all the data from even months of all six phenocams, and **(d)** the model was trained and evaluated using the all the data from six phenocams with the five-fold cross-validation method. The accuracy metrics included producer's accuracy (PA), user's accuracy (UA), and overall accuracy (OA).

	PA (%)	UA (%)	OA (%)
(a) Test-1			
Leaf	**95.61**	**96.50**	**92.89**
Non-leaf	87.16	80.96	
(b) Test-2			
Leaf	**96.45**	**97.64**	**94.81**
Non-leaf	90.50	86.35	
(c) Test-3			
Leaf	**92.84**	**95.81**	**90.62**
Non-leaf	86.65	78.17	
(d) Test-4			
Leaf	**96.81**	**97.08**	**95.28**
Non-leaf	90.10	89.22	

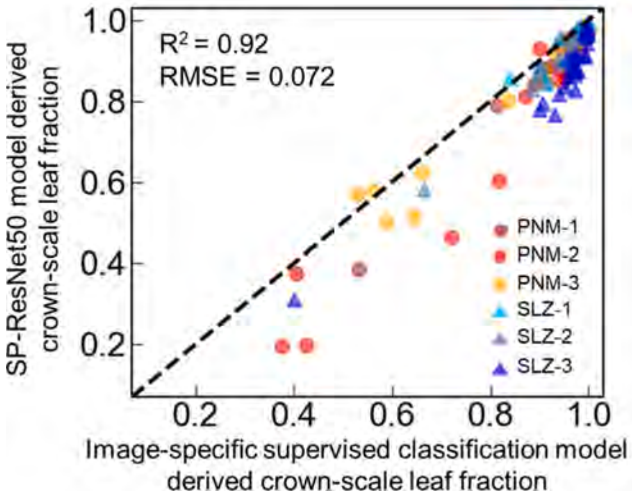

Fig. 7. Leaf fraction values derived from SP-ResNet50 and image-specific supervised classification. The dots represent individual tree-crowns at the PNM site and the triangles represent individual tree-crowns at the SLZ site; different colors indicate the trees from different phenocams; the black dashed line indicates the 1:1 line.

In the second assessment, we evaluated the model performance of SP-ResNet50-derived leaf fraction at the crown scale. Our results showed that leaf fraction derived from SP-ResNet50 covered the same range of values as that derived from the image-specific supervised classifications (R^2 = 0.92; RMSE = 0.072), though we also observed a slight bias (bias = 0.051 with leaf fraction values ranging from 0.2 to 1.0) in the SP-ResNet50 model results (Fig. 7). Since the tree-crowns selected for this cross-comparison spanned all six phenocams and

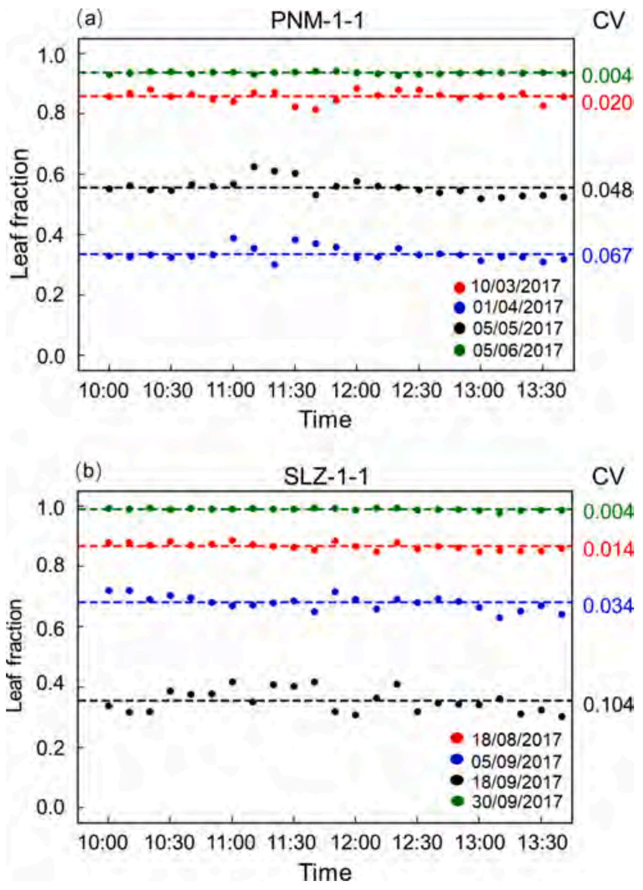

Fig. 8. Variation in leaf fraction throughout the day from two example tree-crowns of **(a)** #PNM-1–1 and **(b)** #SLZ-1–1 across four phenological stages. The dashed lines represent the mean tree-crown leaf fraction within each day with the associated coefficient of variation (CV) next to the line.

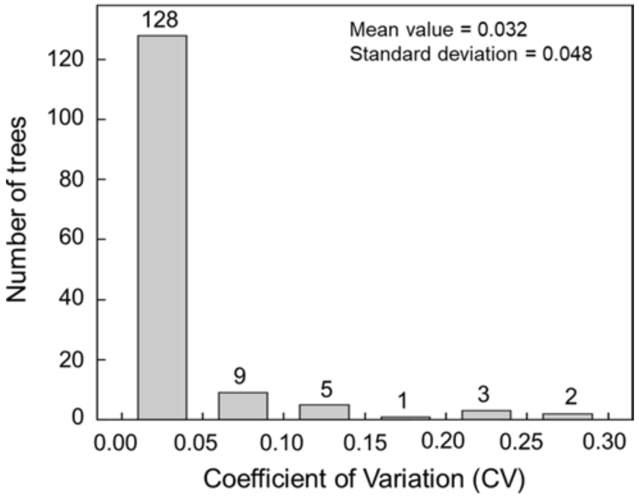

Fig. 9. Histogram of the coefficient of variation (CV) derived using the SP-ResNet50 model for all the identified tree-crowns from all six phenocams across the four phenological stages.

covered three different phenological stages (early and peak leaf falling, and early leaf emergence; Fig. S3), the agreement between the two approaches further suggested that our proposed SP-ResNet50 model accurately captured crown-scale leaf fractions across different phenocams and phenological timing.

In the third assessment, we evaluated the stability of the SP-ResNet50-derived leaf fraction for each tree-crown across a day (at 10-minute intervals from 10:00 AM to 1:50 PM local time) based on the assumption that leaf fraction would remain constant within a day. Our results showed that the SP-ResNet50-derived crown-scale leaf fractions were stable within a day, regardless of phenological stages (Figs. 8 and 9). For example, the SP-ResNet50-derived crown-scale leaf fractions from two deciduous tree-crowns (#PNM-1–1 and #SLZ-1–1; Fig. 2) remained stable within each phenological stage examined here, with diurnal CV values varying from 0.04 to 0.104 (Fig. 8a-b). Moreover, among all the identified tree-crowns from six phenocams and across four phenological stages, 92.6% hold diurnal CV values$<$0.1 (Fig. 9). These results together demonstrated that the SP-ResNet50 model consistently captured crown-scale leaf fractions, and was relatively insensitive to the diurnal and seasonal variation in the sky conditions associated with image acquisition.

To assess whether model selection would affect our classification results, we compared the accuracies between the SP-ResNet50 and U-Net models across three levels, as well as visually inspected the classified images from the two models. At the pixel level, both models effectively classified both leaf and non-leaf pixels with comparable accuracies (Table S2). At the crown level, when comparing with the SP-ResNet50 model (R^2 = 0.92; RMSE = 0.072; Fig. 7), the overall accuracy of crown leaf fraction obtained by the U-Net model was slightly lower, but the RMSE was also lower (R^2 = 0.89; RMSE = 0.055; Fig. 10). At the diurnal level, the U-Net model could also capture the stable crown-scale leaf fractions across the day (Fig. S4) with the mean CV value and associated standard deviation of 0.022 and 0.038, respectively, which was slightly more stable than the SP-ResNet50 model (mean CV = 0.032; standard deviation CV = 0.048; Fig. 9). In addition, from the visual inspection, we found that the SP-ResNet50 model (Fig. 11c; additional examples in Fig. S3) captured more detailed information from the phenocam image (Fig. 11a). The U-Net model captured fewer details and misclassified some canopy gaps within and across leafy tree-crowns (e.g., patches 1–3 in Fig. 11a). U-Net also tended to omit thin branches (e.g., patch 4 in Fig. 11a), while there was more salt and pepper noise within the classified image derived from the SP-ResNet50 model (Fig. 11c-d).

4.2. *Cross-comparison of the SP-ResNet50-derived leaf fraction with GCC*

To explore the potential advantages and disadvantages of the

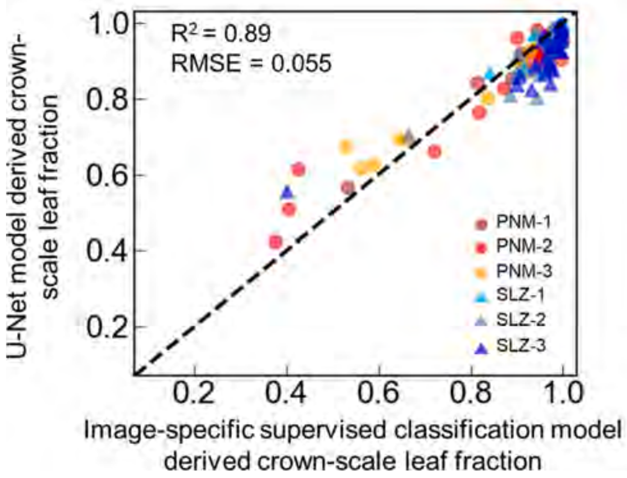

Fig. 10. Comparison between the leaf fraction values derived from the U-Net model and image-specific supervised classification. The dots represent individual tree-crowns at the PNM site; the triangles represent individual tree-crowns at the SLZ site; different colors indicate the trees from different phenocams; the black dashed line indicates the 1:1 line.

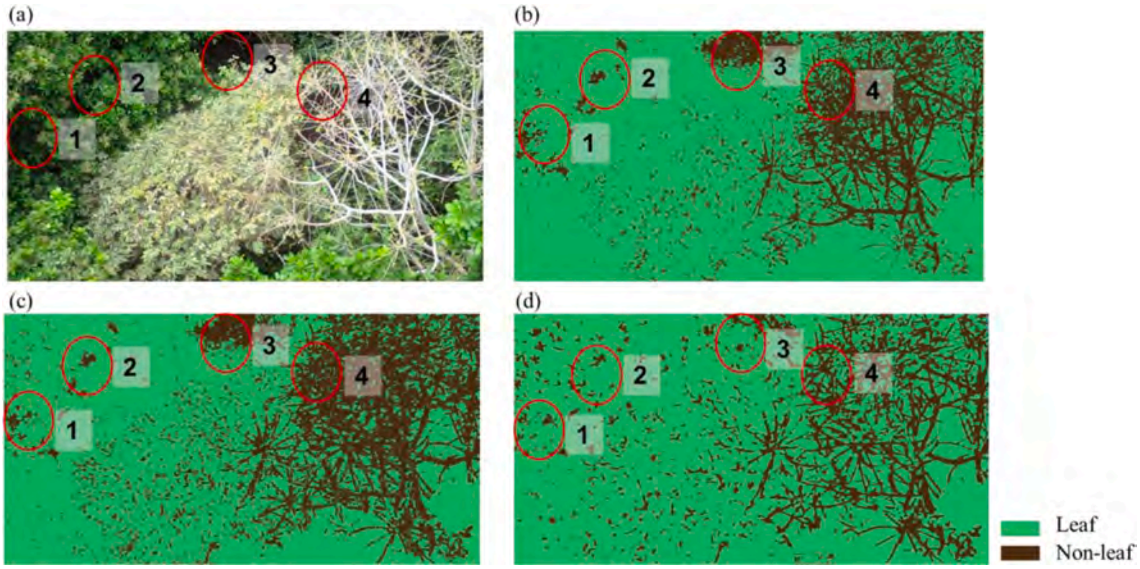

Fig. 11. Cross-method comparisons of leaf and non-leaf classifications. **(a)** The phenocam image from PNM-1 and the corresponding classification results derived from **(b)** the image-specific supervised classification, **(c)** the SP-ResNet50 model, and **(d)** the U-Net model. The red circles indicate the same places for each image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

proposed leaf fraction metric, we compared the monthly mean of SP-ResNet50-derived crown-scale leaf fractions with GCC. Across all the tree-crowns identified from the six phenocams, our results demonstrated that the SP-ResNet50-derived monthly crown-scale leaf fractions had a two-segment piecewise linear relationship with GCC, though the overall correlation was weak ($R^2 = 0.17$; Fig. 12). Specifically, before full canopy closure, with tree crown GCC values ranging from 0.33 to 0.37, we found a tight, linear relationship between the two metrics. However, after full canopy closure when GCC values of crowns were greater than 0.37, the relationship between the two metrics saturated, with the SP-ResNet50-derived leaf fraction stabilizing at around 0.9. These results demonstrate that the relationship between leaf fraction and GCC is complex, and that GCC alone could not be used to directly infer leaf quantity of tropical tree-crowns.

We next compared the four phenological dates (SOLE, ELOE, SOLF, and EOLF) derived from leaf fraction and GCC time series. We found that SP-ResNet50-derived crown-scale leaf fractions detected large variations in these four phenological dates across different tree-crowns (Fig. 13), with SOLE varying from DOY (in days since 01/01/2016) 414 to DOY 628, ELOE from DOY 435 to DOY 640, SOLF from DOY 351 to DOY 604, and EOLF from DOY 378 to 628. We also found that the per-crown phenological dates extracted using the SP-ResNet50-derived leaf fraction were significantly correlated with the analogous phenological dates extracted from the GCC timelines for SOLE ($R^2 = 0.99$, RMSE = 4.1 days, $p < 0.001$; Fig. 13a), ELOE ($R^2 = 0.99$ RMSE = 4.5 days, $p < 0.001$; Fig. 13b), and EOLF ($R^2 = 0.99$, RMSE = 5.6 days, $p < 0.001$; Fig. 13d). The by-crown start of leaf fall (SOLF) extracted using the SP-ResNet50-derived leaf fraction and using GCC appeared to be more discordant ($R^2 = 0.15$, RMSE = 64.7 days, $p < 0.001$; Fig. 13c), primarily associated with a tree crown of SLZ-1–1 (Fig. 14a and S5a) that had an extended leaf senescence duration, resulting in large difference in the SOLF detected using the two different metrics.

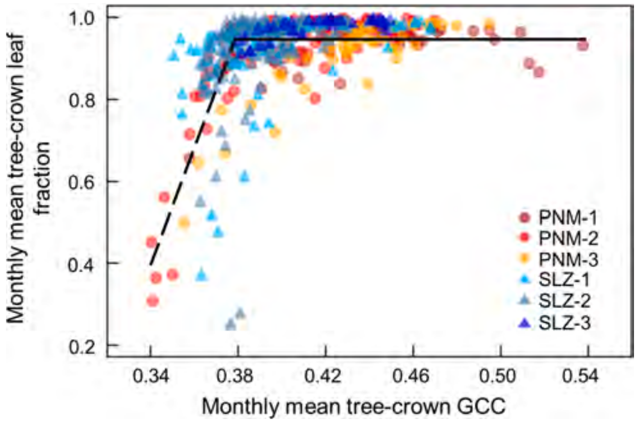

Fig. 12. Monthly mean leaf fraction plotted against GCC at the crown scale for all tree-crowns identified from the six phenocams. The dots represent trees at the PNM site, and the triangles represent trees at the SLZ site; the dashed line represents the piece-wise fit segment when GCC is between 0.33 and 0.37; the solid line represents the fit segment when GCC is greater than 0.37.

Lastly, we explored the separate and combined use of the two metrics to monitor leaf phenology at the tree-crown scale. SP-ResNet50-derived crown-scale leaf fractions showed strong seasonal patterns for tropical deciduous/semi-deciduous trees (Fig. 14a-c and S6a-h) but remained aseasonal for tropical evergreen trees (Fig. 14d and S6i-l). In contrast, the GCC seasonality of evergreen trees was relatively similar to that of the deciduous trees. Moreover, our results showed that the combined seasonal trends of the two metrics helped infer critical leaf age transitions of tropical trees, such as the timing of leaf flushing, development, and senescence (Fig. 14; corresponding image time series in Fig. S5; additional examples in Fig. S6). For tropical evergreen trees, we found that leaf fraction was nearly constant across the year, and thus any seasonality observed by GCC was driven by leaf aging and associated changes in leaf quality, as the initial rapid increase and the peak in GCC would be due to new leaves flushing, and the subsequent slower decrease in GCC would be due to leaf maturation and senescence (Fig. 14d). For deciduous and semi-deciduous trees, senescence and start of leaf fall (SOLF) are often coupled, as shown by the tight relationship between leaf fraction and GCC in Fig. 13c. However, there are instances where this is no longer true, with senescence occurring without leaf fall. In these cases, leaf fraction and GCC become decoupled and both metrics are required to accurately characterize changes in leaf quality (Fig. 14a-b). After leaf fraction returned to its maximum, GCC showed a continuous downward trend, which can be associated with the maturing of recently flushed leaves, similar to the pattern observed in tropical evergreen trees (Fig. 14d).

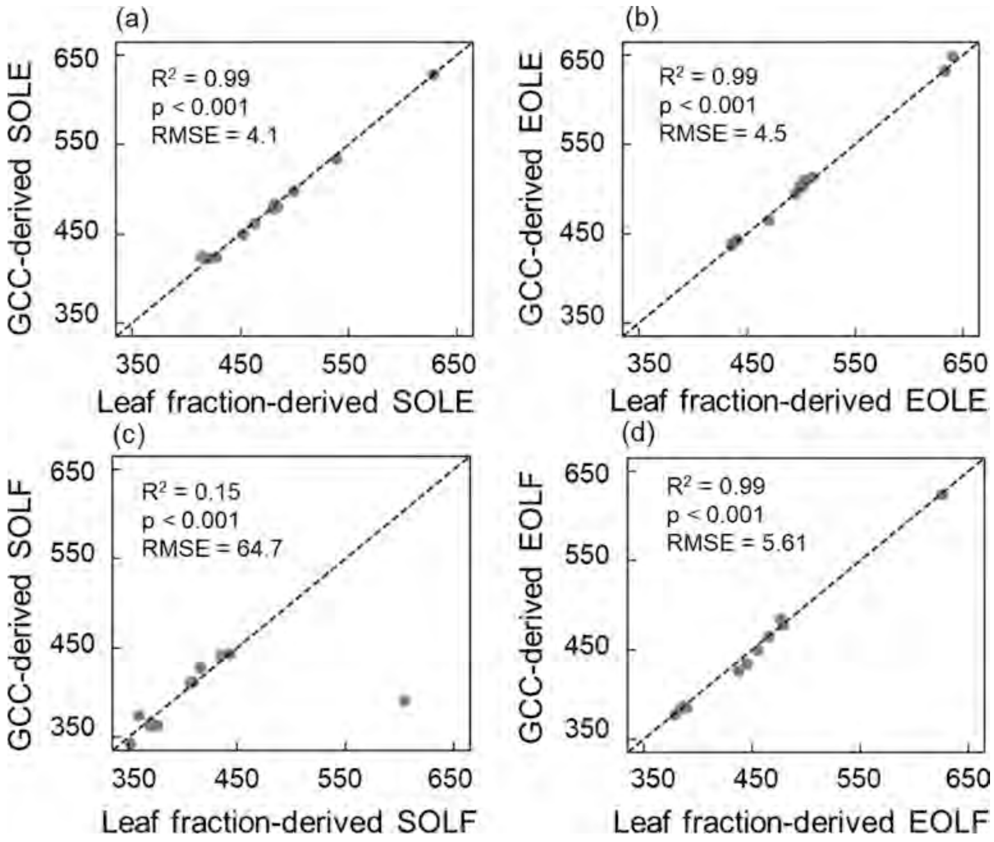

Fig. 13. Comparisons of the four phenological dates derived using leaf fractions with those derived using GCC, including the phenological dates (in days since 01/01/2016) for **(a)** the start of leaf emergence (SOLE), **(b)** the end of leaf emergence (EOLE), **(c)** the start of leaf fall (SOLF), and **(d)** the end of leaf fall (EOLF). The gray dots indicate all deciduous and semi-deciduous tree-crowns; the dashed lines indicate the 1:1 line.

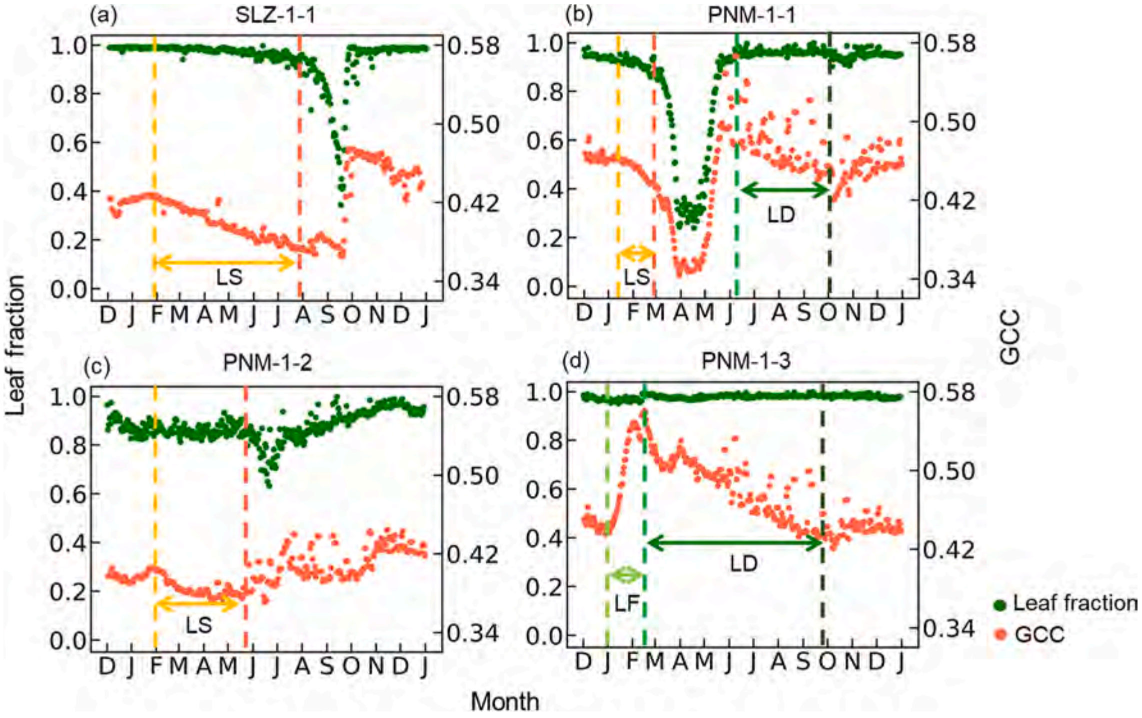

Fig. 14. Examples of crown-scale seasonal trends of leaf fraction and GCC of different tree-crowns, including **(a-b)** deciduous trees, **(c)** a semi-deciduous tree, and **(d)** an evergreen tree. The green and orange dots indicate crown-scale leaf fraction and GCC, respectively; LF represents leaf flushing; LD represents leaf development, and LS represents leaf senescence. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3. Ecosystem-scale phenology patterns of the SP-ResNet50-derived leaf fraction across the two contrasting forests

We explored ecosystem-scale seasonality of the two tropical forests by comparing the monthly mean of crown-scale leaf fractions across all tree-crowns within the same site with the corresponding EVI seasonality extracted from MODIS. We found differences in ecosystem-scale leaf fraction seasonality between the two sites and that the leaf fraction metric shared a similar seasonal pattern with MODIS-derived EVI (Figs. 15 and S7). Specifically, at the wetter evergreen forest site of SLZ, leaf fraction and EVI displayed little seasonality throughout the year (Figs. 15 and S7b). In contrast, at the drier evergreen forest site of PNM, both leaf fraction and EVI displayed stronger and very similar seasonal variation ($R^2 = 0.80$ between leaf fraction and EVI; Figs. 15 and S7a), with an initial decline of both metrics during the dry season from January to April, an increase from April to May, and relative stability in the wet season from May to December. These results demonstrated that our proposed leaf fraction metric effectively captured the phenology pattern of tropical forests, consistent with seasonal patterns observed by MODIS EVI at the ecosystem-scale.

5. Discussion

In this paper we tested the SP-ResNet50 model for extracting tree-crown-scale leaf fraction and its seasonality from tower-based RGB phenocam time-series images. Our results demonstrated that SP-ResNet50 accurately differentiated leaves from non-leaves in time-series phenocam images (Table 1 and Fig. S3) and captured leaf fraction phenology at both tree-crown and ecosystem scales (Figs. 7 and 15). Our results also showed that the combined metrics of leaf fraction and GCC helped infer critical leaf aging processes of crown-scale leaf cohorts such as leaf flush, development and senescence (Fig. 14), which are key to photosynthesis seasonality modeling in moist tropical forests (Manoli et al., 2018; Wu et al., 2016). Collectively, these results suggest that SP-ResNet50 can be an effective and automated means for accurately quantifying leaf fraction, and then monitoring the phenology of leaf quantity in moist tropical forests when applied to phenocam observations.

The accuracy and robustness of our proposed SP-ResNet50 model for leaf fractional cover was rigorously assessed at three levels. First, at the superpixel level, the model was evaluated with the independent testing sets both spatially (within and across phenocams; Table 1a-b) and temporally (model from odd months tested using even months; Table 1c), as well as the five-fold cross-validation (Table 1d). Second, at the crown level, the model-derived leaf fractions (by applying a single SP-ResNet50 model to all phenocam images across all sites and seasons) were cross-compared with those derived using image-specific supervised classification (Fig. 7). Third, at the diurnal level, the stability of the model-derived tree-crown leaf fractions was assessed for all images of tree-crowns acquired within a day and across four different phenological stages (early and peak of leaf falling and leaf emergence) (Figs. 8 and 9). Through these tests, SP-ResNet50 was found to be robust and able to differentiate leaf and non-leaf superpixels with consistently high accuracies (OA = 90.62%-94.81%; Table 1), to accurately characterize tree-crown leaf fractions ($R^2 = 0.92$; Fig. 7), and to generate stable prediction of tree-crown leaf fractions, both within a day and across seasons (mean CV = 0.032; Figs. 8 and 9).

The success and efficiency of SP-ResNet50 for leaf fraction phenology monitoring relies on the three conditions. First, SP-ResNet50 is an advanced deep learning model that automatically learns and extracts representative features (e.g. color and texture; mean and variability) from training samples (Tong et al., 2020; Zhang et al., 2018), and thus is an effective classifier to differentiate leaves from non-leaves in phenocam images. Second, SP-ResNet50 used superpixels as input units, where spatially contiguous pixels sharing similar color and texture features in high spatial resolution images were grouped into one object (Mi and Chen, 2020), improving model accuracy and computational efficiency, particularly when applying to the time-series phenocam images across a full year with large data quantity (c. 2300 images and c. 46 billion image pixels). Third, a comprehensive library of leaf and non-leaf pixels was built to train the SP-ResNet50 model. These were sampled on the 15th day of each month throughout the full year and under various acquisition times, exposure conditions, and specific phenocams. Such various and comprehensive training samples ensure that the model was trained with sufficient and representative features.

The comparison of the SP-ResNet50 model with the U-Net model showed that both the superpixel-based classification and the semantic segmentation approaches worked well across all three levels of model performance assessments (Tables 1 and S2; Figs. 7, 9, 10, S3 and S4), each with its own advantages and disadvantages. SP-ResNet50 models extract local fine detail information of the target object, running superpixel-wise with a sliding window (Gong et al., 2017). On the other hand, U-Net models directly learn global semantic information of the entire input image (Ronneberger et al., 2015). As a result, the U-Net model had limited capacity to extract smaller objects successfully (Fig. 11d and S3), though it also had reduced salt and pepper effect in the output classification image (Fig. 11c-d). Moreover, we showed that sparse but comprehensive training labels for the object of interest can successfully train the semantic segmentation model (Table S2; Figs. 10 and 11 and S3), though most studies have trained semantic segmentation models with training images that are fully covered by the manually created labels (Kemker et al., 2018; Volpi and Tuia, 2018; Zhang et al., 2020), which are often difficult to obtain.

Relative to most previous phenocam-based studies that rely on GCC, to the best of our knowledge, this is the first study to provide a metric with a clear biophysical meaning for tree-crown scale leaf phenology monitoring in the tropics using RGB phenocam observations. This is important for two reasons. First, moist tropical forests exhibit diverse

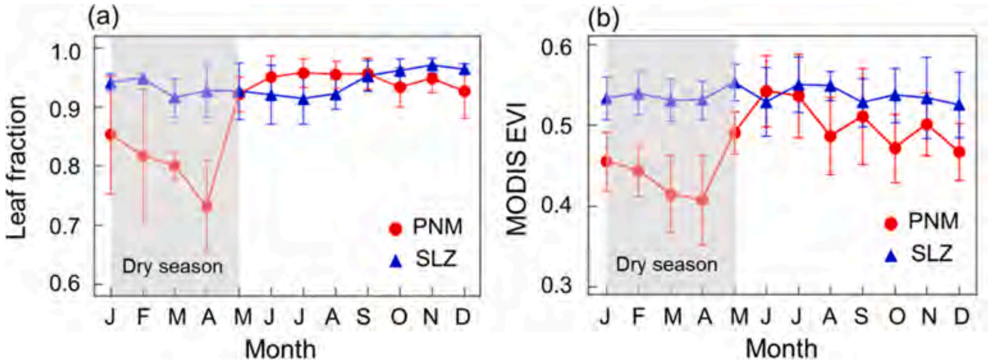

Fig. 15. Ecosystem-scale seasonality of **(a)** leaf fraction and **(b)** MODIS EVI with BRDF correction. Light grey shading indicates the dry season of PNM and SLZ sites; error bars indicate one standard deviation.

leaf phenological habits, ranging from deciduous/semi-deciduous to evergreen seasonal leaf display habits (Fig. 14; also Lopes et al., 2016; Detto et al., 2018). In temperate deciduous forests, changes in leaf quantity are usually accompanied with leaf color change (Keenan et al., 2014; Richardson, 2019; Yang et al., 2014), allowing GCC and leaf fraction to both accurately highlight phenological stages individually. However, in moist tropical forests, due to the various leaf phenological patterns, we found leaf color to be partially decoupled from leaf quantity (Fig. 14a-d and S6i-l), and thus, it is difficult to use the GCC metric alone for such diverse tropical leaf phenology monitoring.

Second, in addition to key phenological dates (Fig. 13), leaf fraction with clear biophysical meaning can help to accurately assess the magnitude of annual leaf exchange in moist tropical forests that have fully deciduous or semi-deciduous crowns, which is another important aspect of tropical leaf phenology (Singh et al., 2020; Williams et al., 2008). The poor relationship between GCC and leaf fraction (Fig. 12) suggests that GCC alone is insufficient to indicate the magnitude of annual leaf exchange (Fig. 14), while the proposed leaf fraction metric is not consistent in deriving leaf phenological dates, particularly for evergreen trees with nearly constant leaf fraction over the annual cycle (Fig. 14d and S6i-l). By combining the two metrics of GCC and leaf fraction, we can monitor both the timing and magnitude of leaf exchange over the annual cycle, as well as the critical leaf phenological transitions, including the timing and duration of leaf emergence, maturation, and senescence (Figs. 14 and S6). Since leaf biochemistry and physiology can both change with leaf aging (Albert et al., 2018; Chavana-Bryant et al., 2017; Wu et al., 2017, 2019), and such age-induced seasonal variability in leaf biochemical and physiological traits importantly regulates ecosystem-scale photosynthesis seasonality (Wu et al., 2016), we recommend the combined use of these two metrics to characterize leaf age and associated changes in leaf quality in the future.

Our approach also brings new insights advancing plant phenology studies. First, it provides important observations to help interpret the phenological scaling process from tree-crowns to forest ecosystems in the tropics (Lopes et al., 2016; Saleska et al., 2016; Wu et al., 2018). The variability in leaf phenology at the tree-crown scale determines the ecosystem-scale phenology pattern (Lopes et al., 2016; Wang et al., 2020), but the necessary crown-scale phenology observations to understand this scaling process remain lacking. With the crown-scale leaf fraction phenology extracted in this study, we observed that both phenocams and MODIS EVI detected aseasonal phenology patterns in a wetter evergreen forest in SLZ with ~ 5% annual variation in canopy leaf fraction, and modest phenology patterns in a drier evergreen forest in PNM with ~ 30% annual variation in canopy leaf fraction (Figs. 15 and S7). Previously, Wu et al (2018) used the similar datasets in an Amazonian moist tropical forest and observed modest seasonal phenology patterns with ~ 10% annual variation in canopy leaf fraction, demonstrating that both leaf fraction and leaf age jointly regulated the ecosystem-scale seasonality detected by MODIS EVI. This previous finding, together with our observations here (Figs. 15 and S7), further suggest that the relative importance of leaf fraction and leaf age would vary largely across moist tropical forests, with the drier site tending to be dominated by the canopy leaf fraction ($R^2 = 0.80$; Figs. 15 and S7a). Meanwhile, it also suggests that more similar observations across large rainfall gradients are still needed to correctly interpret satellite-detected greenness seasonality.

Second, our method could facilitate phenological analyses using phenocam observations because the model-derived leaf fraction is relatively insensitive to ambient conditions for image acquisition, such as diurnal variations in ambient light condition (Figs. 8 and 9). Variable illumination, frequent cloud cover, and mist over tree-crowns have been a common challenge of using phenocam observations for accurate monitoring of leaf phenology in general (Klosterman et al., 2014; Richardson et al., 2018; Sonnentag et al., 2012) and tropical phenology in particular (Lopes et al., 2016; Park et al., 2019), as they can cause coloration artifacts and incorrectly detect leaf color changes within tree-crowns, resulting in inaccurate assessment of leaf phenology patterns. The SP-ResNet50 model's relative insensitivity to ambient light condition is largely because the deep learning model takes advantages of high-level features (e.g. texture and spatial context; Kattenborn et al., 2021; Lecun et al., 2015) information for differentiating leaf and non-leaf pixels. The success of SP-ResNet50 thus represents a reliable alternative to process phenocam images for leaf phenology monitoring (if based on leaf fractional cover) as against previous attempts relying on automatic (e.g. Richardson et al., 2018, Sonnentag et al., 2012) or manual (e.g. Lopes et al., 2016; Wu et al., 2016) approaches for quality control of phenocam images to minimize impacts of variation in ambient light condition on the extracted leaf phenology.

Despite these promising implications, our study also identifies three important steps that need to be considered for further advancement. First, similar to other deep learning algorithms (Maggiori et al., 2017; Scott et al., 2017), the proposed SP-ResNet50 model relies on very large and representative training sample datasets, which require substantial manual effort to build. To resolve this issue, we recommend the future attempt of using other artificial intelligence techniques (e.g. data augmentation, generative adversarial networks) to automatically generate reliable training samples from limited labelled data, thus facilitating dataset construction and lowering labor cost (Goodfellow et al., 2020; Perez and Wang, 2017). Second, the broad applicability of our proposed SP-ResNet50 model in the tropics has not yet been assessed. There are large variations in tree species compositions with diverse phenological patterns across large tropical areas (Reich, 1995; Sakai and Kitajima, 2019) and we only focused on two moist tropical forest sites in this study. Thus, it is important to explore whether the proposed approach can be extended to other tropical forests. Finally, although our results suggest that the combined use of leaf fraction and GCC helps infer leaf aging processes (and thus likely also leaf quality), it remains difficult to directly infer leaf quality information from phenocam observations. Thus, the accompanying field observations of leaf age and age-associated leaf biochemistry and physiological traits are still needed to explore the possibility of leveraging the combined metrics to monitor leaf quality using phenocam observations.

6. Conclusion

Accurate monitoring of tropical leaf phenology from tree-crowns to forest ecosystems is essential to understand the response of tropical forests to climate change (Wang et al., 2020), but remains challenging. Here, we developed a SP-ResNet50 model and applied it to phenocam observations to enable monitoring of leaf fraction phenology in two moist tropical forests in Panama. Our proposed SP-ResNet50 model was shown to be accurate and robust with three levels of model performance assessments (superpixel, crown, and diurnal levels). Compared with the conventional color-based GCC metric, our model-derived leaf fraction metric has a clear biophysical meaning and can help capture the magnitude of annual leaf exchange and differentiate the leaf phenological habits among deciduous/semi-deciduous and evergreen trees. We also demonstrated that the combined use of leaf fraction and GCC helped infer the change of leaf age over the annual cycle, including the timing and duration of leaf emergence, maturation and senescence, which offers critical information to track the seasonal variation in leaf quality and thus ecosystem-scale photosynthetic capacity in moist tropical forests. Collectively, our approach improves the cross-scale leaf phenology monitoring in the tropics ranging from tree-crowns to forest ecosystems using phenocam observations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the editors and the two anonymous reviewers for providing valuable suggestions and comments, which are greatly helpful in improving this work. J.Wu, G.Song, S.Wu, C.Lee, J.Wang, and Z.Lin were supported by National Natural Science Foundation of China (#31922090). J. Wu and G. Song were in part supported by Hong Kong Research Grant Council (GRF #17305321) and M.K. Ng was in part supported by Hong Kong Research Grant Council (GRF #12300218, #12300519, #17201020, #17300021, and NSFC/RGC #N_HKU769/21). A. Rogers, S. Serbin, K. Ely, and camera imagery collection were supported by the Next-Generation Ecosystem Experiments Tropics (NGEE-Tropics) project that is supported by the Office of Biological and Environmental Research in the Department of Energy, Office of Science, and through the United States Department of Energy contract No. DE-SC0012704 to Brookhaven National Laboratory. B.Wolfe was supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, McIntire Stennis project under LAB94493.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2021.10.023.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (11), 2274–2282.

Albert, L.P., Restrepo-Coupe, N., Smith, M.N., Wu, J., Chavana-Bryant, C., Prohaska, N., Taylor, T.C., Martins, G.A., Ciais, P., Mao, J., Arain, M.A., Li, W., Shi, X., Ricciuto, D. M., Huxman, T.E., McMahon, S.M., Saleska, S.R., 2019. Cryptic phenology in plants: Case studies, implications, and recommendations. Global Change Biology 25 (11), 3591–3608.

Albert, L.P., Wu, J., Prohaska, N., Camargo, P.B., Huxman, T.E., Tribuzy, E.S., Ivanov, V. Y., Oliveira, R.S., Garcia, S., Smith, M.N., Oliveira Junior, R.C., Restrepo-Coupe, N., Silva, R., Stark, S.C., Martins, G.A., Penha, D.V., Saleska, S.R., 2018. Age-dependent leaf physiology and consequences for crown-scale carbon uptake during the dry season in an Amazon evergreen forest. New Phytologist 219 (3), 870–884.

Alberton, B., Almeida, J., Helm, R., da S. Torres, R., Menzel, A., Morellato, L.P.C., 2014. Using phenological cameras to track the green up in a cerrado savanna and its on-the-ground validation. Ecological Informatics 19, 62–70.

Alberton, B., Torres, R.d.S., Cancian, L.F., Borges, B.D., Almeida, J., Mariano, G.C., Santos, J.D., Morellato, L.P.C., 2017. Introducing digital cameras to monitor plant phenology in the tropics: applications for conservation. Perspectives in Ecology and Conservation 15 (2), 82–90.

Berra, E.F., Gaulton, R., Barr, S., 2019. Assessing spring phenology of a temperate woodland: A multiscale comparison of ground, unmanned aerial vehicle and Landsat satellite observations. Remote Sensing of Environment 223, 229–242.

Brando, P.M., Goetz, S.J., Baccini, A., Nepstad, D.C., Beck, P.S.A., Christman, M.C., 2010. Seasonal and interannual variability of climate and vegetation indices across the Amazon. Proceedings of the National Academy of Sciences of the United States of America 107 (33), 14685–14690.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.

Bush, E.R., Abernethy, K.A., Jeffery, K., Tutin, C., White, L., Dimoto, E., Dikangadissi, J.-T., Jump, A.S., Bunnefeld, N., Freckleton, R., 2017. Fourier analysis to detect phenological cycles using long-term tropical field data and simulations. Methods in Ecology and Evolution 8 (5), 530–540.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sensing of Environment 210, 35–47.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G.P., Anastasiou, A., Enquist, B.J., Cosio Caravasi, E.G., 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 (3), 1049–1063.

Chen, Y., Ming, D., Lv, X., 2019. Superpixel based land cover classification of VHR satellite image combining multi-scale CNN and scale parameter estimation. Earth Science Informatics 12 (3), 341–363.

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37 (1), 35–46.

de Moura, Y.M., Galvão, L.S., Hilker, T., Wu, J., Saleska, S., do Amaral, C.H., Nelson, B. W., Lopes, A.P., Wiedeman, K.K., Prohaska, N., de Oliveira, R.C., Machado, C.B., Aragão, L.E.O.C., 2017. Spectral analysis of amazon canopy phenology during the dry season using a tower hyperspectral camera and modis observations. ISPRS Journal of Photogrammetry and Remote Sensing 131, 52–64.

Detto, M., Wright, S.J., Calderón, O., Muller-Landau, H.C., 2018. Resource acquisition and reproductive strategies of tropical forest in response to the El Niño-Southern Oscillation. Nature Communications 9, 1–8.

Ding, P., Zhang, Y.e., Deng, W.-J., Jia, P., Kuijper, A., 2018. A light and faster regional convolutional neural network for object detection in optical remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing 141, 208–218.

Du, S., Du, S., Liu, B.o., Zhang, X., 2021. Mapping large-scale and fine-grained urban functional zones from VHR images using a multi-scale semantic segmentation network and object based approach. Remote Sensing of Environment 261, 112480. https://doi.org/10.1016/j.rse.2021.112480.

Eamus, D., Myers, B., Duff, G., Williams, D., 1999. Seasonal changes in photosynthesis of eight savanna tree species. Tree Physiology 19 (10), 665–671.

Gonçalves, N.B., Lopes, A.P., Dalagnol, R., Wu, J., Pinho, D.M., Nelson, B.W., 2020. Both near-surface and satellite remote sensing confirm drought legacy effect on tropical forest leaf phenology after 2015/2016 ENSO drought. Remote Sensing of Environment 237, 111489. https://doi.org/10.1016/j.rse.2019.111489.

Gong, M., Zhan, T., Zhang, P., Miao, Q., 2017. Superpixel-based difference representation learning for change detection in multispectral remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 55 (5), 2658–2673.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. Communications of the ACM 63 (11), 139–144.

Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., Caylor, K.K., Sheffield, J., Wood, E.F., Malhi, Y., Liang, M., Kimball, J.S., Saleska, S., Berry, J., Joiner, J., Lyapustin, A.I., 2015. Photosynthetic seasonality of global tropical forests constrained by hydroclimate. Nature Geoscience 8 (4), 284–289.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Huete, A.R., Didan, K., Shimabukuro, Y.E., Ratana, P., Saleska, S.R., Hutyra, L.R., Yang, W., Nemani, R.R., Myneni, R., 2006. Amazon rainforests green-up with sunlight in dry season. Geophysical Research Letters 33, 2–5.

Hussein, B.R., Malik, O.A., Ong, W.H., Slik, J.W.F., 2021. Automated Extraction of Phenotypic Leaf Traits of Individual Intact Herbarium Leaves from Herbarium Specimen Images Using Deep Learning Based Semantic Segmentation. Sensors 21 (13), 4549.

Jiao, L., Huo, L., Hu, C., Tang, P., 2020. Refined UNet V2: End-to-end patch-wise network for noise-free cloud and shadow segmentation. Remote Sensing 12, 1–21.

Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 173, 24–49.

Keenan, T.F., Darby, B., Felts, E., Sonnentag, O., Friedl, M.A., Hufkens, K., O'Keefe, J., Klosterman, S., Munger, J.W., Toomey, M., Richardson, A.D., 2014. Tracking forest phenology and seasonal physiology using digital repeat photography: A critical assessment. Ecological Applications 24 (6), 1478–1489.

Kemker, R., Salvaggio, C., Kanan, C., 2018. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS Journal of Photogrammetry and Remote Sensing 145, 60–77.

Klosterman, S.T., Hufkens, K., Gray, J.M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell, L., Norman, R., Friedl, M.A., Richardson, A.D., 2014. Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. Biogeosciences 11 (16), 4305–4320.

Kotaridis, I., Lazaridou, M., 2021. Remote sensing image segmentation advances: A meta-analysis. ISPRS Journal of Photogrammetry and Remote Sensing 173, 309–322.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Lopes, A.P., Nelson, B.W., Wu, J., Graça, P.M.L.d.A., Tavares, J.V., Prohaska, N., Martins, G.A., Saleska, S.R., 2016. Leaf flush drives dry season green-up of the Central Amazon. Remote Sensing of Environment 182, 90–98.

Lv, X., Ming, D., Chen, YangYang, Wang, M., 2019. Very high resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification. International Journal of Remote Sensing 40 (2), 506–531.

Ma, A., Wan, Y., Zhong, Y., Wang, J., Zhang, L., 2021. SceneNet: Remote sensing scene classification deep learning network using multi-objective neural evolution architecture search. ISPRS Journal of Photogrammetry and Remote Sensing 172, 171–188.

Maggiori, E., Tarabalka, Y., Charpiat, G., Alliez, P., 2017. Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. IEEE Transactions on Geoscience and Remote Sensing 55 (2), 645–657.

Manoli, G., Ivanov, V.Y., Fatichi, S., 2018. Dry-Season Greening and Water Stress in Amazonia: The Role of Modeling Leaf Phenology. Journal of Geophysical Research: Biogeosciences 123, 1909–1926.

Mi, L.i., Chen, Z., 2020. Superpixel-enhanced deep neural forest for remote sensing image semantic segmentation. ISPRS Journal of Photogrammetry and Remote Sensing 159, 140–152.

Nagai, S., Ichie, T., Yoneyama, A., Kobayashi, H., Inoue, T., Ishii, R., Suzuki, R., Itioka, T., 2016. Usability of time-lapse digital camera images to detect characteristics of tree phenology in a tropical rainforest. Ecological Informatics 32, 91–106.

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 (13), 1534. https://doi.org/10.3390/rs11131534.

Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:*1712.04621.

Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X.u., Shen, M., Zhu, X., 2019. Plant phenology and global climate change: Current progresses and challenges. Global Change Biology 25 (6), 1922–1940.

Ratkowsky, D.A., 1983. Nonlinear regression modeling—A unified practical approach. Marcel Dekker, New York, pp. 61–91.

Reich, P.B., 1995. Phenology of tropical forests: patterns, causes, and consequences. Canadian Journal of Botany 73 (2), 164–174.

Restrepo-Coupe, N., da Rocha, H.R., Hutyra, L.R., da Araujo, A.C., Borma, L.S., Christoffersen, B., Cabral, O.M.R., de Camargo, P.B., Cardoso, F.L., da Costa, A.C.L., Fitzjarrald, D.R., Goulden, M.L., Kruijt, B., Maia, J.M.F., Malhi, Y.S., Manzi, A.O., Miller, S.D., Nobre, A.D., von Randow, C., Sá, L.D.A., Sakai, R.K., Tota, J., Wofsy, S. C., Zanchi, F.B., Saleska, S.R., 2013. What drives the seasonality of photosynthesis across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network. Agricultural and Forest Meteorology 182-183, 128–144.

Restrepo-Coupe, N., Levine, N.M., Christoffersen, B.O., Albert, L.P., Wu, J., Costa, M.H., Galbraith, D., Imbuzeiro, H., Martins, G., da Araujo, A.C., Malhi, Y.S., Zeng, X., Moorcroft, P., Saleska, S.R., 2017. Do dynamic global vegetation models capture the seasonality of carbon fluxes in the Amazon basin? A data-model intercomparison. Global Change Biology 23 (1), 191–208.

Richardson, A.D., 2019. Tracking seasonal rhythms of plants in diverse ecosystems with digital camera imagery. New Phytologist 222 (4), 1742–1750.

Richardson, A.D., Braswell, B.H., Hollinger, D.Y., Jenkins, J.P., Ollinger, S.V., 2009. Near-surface remote sensing of spatial and temporal variation in canopy phenology. Ecological Applications 19 (6), 1417–1428.

Richardson, A.D., Hufkens, K., Milliman, T., Aubrecht, D.M., Chen, M., Gray, J.M., Johnston, M.R., Keenan, T.F., Klosterman, S.T., Kosmala, M., Melaas, E.K., Friedl, M. A., Frolking, S., 2018. Tracking vegetation phenology across diverse North American biomes using PhenoCam imagery. Scientific Data 5, 1–24.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: In International Conference on Medical image computing and computer-assisted intervention, pp. 234–241.

Sakai, S., Kitajima, K., 2019. Tropical phenology: Recent advances and perspectives. Ecological Research 34 (1), 50–54.

Saleska, S.R., Wu, J., Guan, K., Araujo, A.C., Huete, A., Nobre, A.D., Restrepo-Coupe, N., 2016. Dry-season greening of Amazon forests. Nature 531 (7594), E4–E5.

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 (5), n/a–n/a.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R.P., Hu, B., Liang, S., Privette, J. L., Roy, D., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sensing of Environment 83 (1-2), 135–148.

Scott, G.J., England, M.R., Starms, W.A., Marcum, R.A., Davis, C.H., 2017. Training Deep Convolutional Neural Networks for Land-Cover Classification of High-Resolution Imagery. IEEE Geoscience and Remote Sensing Letters 14 (4), 549–553.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arX*iv:1409.1556.

Singh, B., Jeganathan, C., Rathore, V.S., 2020. Improved NDVI based proxy leaf-fall indicator to assess rainfall sensitivity of deciduousness in the central Indian forests through remote sensing. Scientific Reports 10, 1–16.

Singh, K.P., Kushwaha, C.P., 2016. Deciduousness in tropical trees and its potential as indicator of climate change: A review. Ecological Indicators 69, 699–706.

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A.M., Friedl, M., Braswell, B.H., Milliman, T., O'Keefe, J., Richardson, A.D., 2012. Digital repeat photography for phenological research in forest ecosystems. Agricultural and Forest Meteorology 152, 159–177.

Tang, H., Dubayah, R., 2017. Light-driven growth in Amazon evergreen forests explained by seasonal variations of vertical canopy structure. Proceedings of the National Academy of Sciences of the United States of America 114 (10), 2640–2644.

Tao, Y., Xu, M., Zhang, F., Du, B.o., Zhang, L., 2017. Unsupervised-Restricted Deconvolutional Neural Network for Very High Resolution Remote-Sensing Image Classification. IEEE Transactions on Geoscience and Remote Sensing 55 (12), 6805–6823.

Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., Zhang, L., 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. Remote Sensing of Environment 237, 111322. https://doi.org/10.1016/j.rse.2019.111322.

Volpi, M., Tuia, D., 2018. Deep multi-task learning for a geographically-regularized semantic segmentation of aerial images. ISPRS Journal of Photogrammetry and Remote Sensing 144, 48–60.

Wagner, F.H., Sanchez, A., Tarabalka, Y., Lotte, R.G., Ferreira, M.P., Aidar, M.P.M., Gloor, E., Phillips, O.L., Aragão, L.E.O.C., Pettorelli, N., Clerici, N., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. Remote Sensing in Ecology and Conservation 5 (4), 360–375.

Wang, J., Yang, D., Detto, M., Nelson, B.W., Chen, M., Guan, K., Wu, S., Yan, Z., Wu, J., 2020. Multi-scale integration of satellite remote sensing improves characterization of dry-season green-up in an Amazon tropical evergreen forest. Remote Sensing of Environment 246, 111865.

Williams, L.J., Bunyavejchewin, S., Baker, P.J., 2008. Deciduousness in a seasonal tropical forest in western Thailand: Interannual and intraspecific variation in timing, duration and environmental cues. Oecologia 155 (3), 571–582.

Wright, S.J., Horlyck, V., Basset, Y., Barrios, H., Bethancourt, A., Bohlman, S.A., Gilbert, G.S., Goldstein, G., Graham, E.A., Kitajima, K., Lerdau, M.T., Meinzer, F.C., Ødegaard, F., Reynolds, D.R., Roubik, D.W., Sakai, S., Samaniego, M., Sparks, J.P., Van Bael, S., Winter, K., Zotz, G., 2003. Tropical canopy biology program, Republic of Panama. In: Basset, Y., Horlyck, V., Wright, S.J. (Eds.), Studying forest canopies from above: The international canopy crane network. Smithsonian Tropical Research Institute and the United Nations Environmental Programme, Panama City, Panama, pp. 137–155.

Wu, J., Albert, L.P., Lopes, A.P., Restrepo-Coupe, N., Hayek, M., Wiedemann, K.T., Guan, K., Stark, S.C., Christoffersen, B., Prohaska, N., Tavares, J.V., Marostica, S., Kobayashi, H., Ferreira, M.L., Campos, K.S., da Silva, R., Brando, P.M., Dye, D.G., Huxman, T.E., Huete, A.R., Nelson, B.W., Saleska, S.R., 2016. Leaf development and demography explain photosynthetic seasonality in Amazon evergreen forests. Science 351 (6276), 972–976.

Wu, J., Chavana-Bryant, C., Prohaska, N., Serbin, S.P., Guan, K., Albert, L.P., Yang, X.i., Leeuwen, W.J.D., Garnello, A.J., Martins, G., Malhi, Y., Gerard, F., Oliviera, R.C., Saleska, S.R., 2017. Convergence in relationships between leaf traits, spectra and age across diverse canopy environments and two contrasting tropical forests. New Phytologist 214 (3), 1033–1048.

Wu, J., Kobayashi, H., Stark, S.C., Meng, R., Guan, K., Tran, N.N., Gao, S., Yang, W., Restrepo-Coupe, N., Miura, T., Oliviera, R.C., Rogers, A., Dye, D.G., Nelson, B.W., Serbin, S.P., Huete, A.R., Saleska, S.R., 2018. Biological processes dominate seasonality of remotely sensed canopy greenness in an Amazon evergreen forest. New Phytologist 217 (4), 1507–1520.

Wu, J., Rogers, A., Albert, L.P., Ely, K., Prohaska, N., Wolfe, B.T., Oliveira, R.C., Saleska, S.R., Serbin, S.P., 2019. Leaf reflectance spectroscopy captures variation in carboxylation capacity across species, canopy environment and leaf age in lowland moist tropical forests. New Phytologist 224 (2), 663–674.

Wu, J., Wolfe, B., Serbin, S., Ely, K., & Rogers, A. (2021). Panama Tropical forest digital camera imagery for vegetation phenology, Dec2016-May2019, PA-SLZ, PA-PNM, PA-BCI. NGEE Tropics Data Collection. Accessed at https://doi.org/10.15486/ngt/1770776.

Yang, X., Tang, J., Mustard, J.F., 2014. Beyond leafcolor: Comparing camera-based phenological metrics with leaf biochemical, biophysical and spectral properties throughout the growing season of a temperate deciduous forest. Journal of Geophysical Research: Biogeosciences 119 (3), 181–191.

Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., Dangal, S., 2018. Amazon drought and forest response: Largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016. Global Change Biology 24 (5), 1919–1934.

Zan, X., Zhang, Xinlu, Xing, Z., Liu, W., Zhang, Xiaodong, Su, W., Liu, Z., Zhao, Y., Li, S., 2020. Automatic detection of maize tassels from UAV images by combining random forest classifier and VGG16. Remote Sensing 12, 1–17.

Zhang, Ce, Atkinson, Peter M., George, Charles, Wen, Zhaofei, Diazgranados, Mauricio, Gerard, France, 2020. Identifying and mapping individual plants in a highly diverse high-elevation ecosystem using UAV imagery and deep learning. ISPRS Journal of Photogrammetry and Remote Sensing 169, 280–291.

Zhang, Ce, Sargent, Isabel, Pan, Xin, Li, Huapeng, Gardiner, Andy, Hare, Jonathon, Atkinson, Peter M., 2018. An object-based convolutional neural network (OCNN) for urban land use classification. Remote Sensing of Environment 216, 57–70.

Zhang, Xiaoyang, Friedl, Mark A., Schaaf, Crystal B., Strahler, Alan H., Hodges, John C. F., Gao, Feng, Reed, Bradley C., Huete, Alfredo, 2003. Monitoring vegetation phenology using MODIS. Remote Sensing of Environment 84 (3), 471–475.

Zhao, Wei, Jiao, Licheng, Ma, Wenping, Zhao, Jiaqi, Zhao, Jin, Liu, Hongying, Cao, Xianghai, Yang, Shuyuan, 2017. Superpixel-based multiple local CNN for panchromatic and multispectral image classification. IEEE Transactions on Geoscience and Remote Sensing 55 (7), 4141–4156.