

APPLICATION

HemiPy: A Python module for automated estimation of forest biophysical variables and uncertainties from digital hemispherical photographs

Luke A. Brown^{1,2}  | Harry Morris^{2,3}  | Sylvain Leblanc⁴  | Gabriele Bai⁵  |
Christian Lanconelli⁶  | Nadine Gobron⁷  | Courtney Meier⁸  | Jadunandan Dash² 

¹School of Science, Engineering & Environment, University of Salford, Manchester, UK; ²School of Geography and Environmental Science, University of Southampton, Southampton, UK; ³Earth Observation, Climate and Optical Group, National Physical Laboratory, Teddington, UK; ⁴Canada Centre for Remote Sensing, Natural Resources Canada, Ottawa, Ontario, Canada; ⁵ACRI-ST, Sophia-Antipolis, France; ⁶UniSystems, Milan, Italy; ⁷European Commission Joint Research Centre, Ispra, Italy and ⁸National Ecological Observatory Network, Batelle, Boulder, Colorado, USA

Correspondence

Luke A. Brown

Email: l.a.brown4@salford.ac.uk

Funding information

European Space Agency, Grant/Award Number: 4000139939; Joint Research Centre, Grant/Award Number: FWC 932059; National Science Foundation

Handling Editor: Hooman Latifi

Abstract

1. Digital hemispherical photography (DHP) is widely used to derive forest biophysical variables including leaf, plant, and green area index (LAI, PAI and GAI), the fraction of intercepted photosynthetically active radiation (FIPAR), and the fraction of vegetation cover (FCOVER). However, the majority of software packages for processing DHP data are based on a graphical user interface, making programmatic analysis difficult. Meanwhile, few natively support analysis of RAW image formats, while none incorporate the propagation or provision of uncertainties.
2. To address these limitations, we present HemiPy, an open-source Python module for deriving forest biophysical variables and uncertainties from DHP images in an automated manner. We assess HemiPy using simulated hemispherical images, in addition to multiannual time-series and litterfall data from several forested National Ecological Observatory Network (NEON) sites, as well as comparison against the CAN-EYE software package.
3. Multiannual time-series of PAI, FIPAR and FCOVER demonstrate HemiPy's outputs realistically represent expected temporal patterns. Comparison against litterfall data reveals reasonable accuracies are achievable, with RMSE values close to the error of ~1 unit typically attributed to optical LAI measurement approaches. HemiPy's PAI, FIPAR and FCOVER outputs demonstrate good agreement with CAN-EYE. Consistent with previous studies, when compared to simulated hemispherical images, better agreement is observed for PAI derived using gap fraction near the hinge angle of 57.5° only, as opposed to values derived using gap fraction over a wider range of zenith angles.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

4. HemiPy should prove a useful tool for processing DHP images, and its open-source nature means that it can be adopted, extended and further refined by the user community.

KEYWORDS

automation, digital hemispherical photography, fraction of absorbed photosynthetically active radiation, fraction of intercepted photosynthetically active radiation, fraction of vegetation cover, green area index, leaf area index, plant area index

1 | INTRODUCTION

Canopy biophysical variables, including leaf, plant and green area index (LAI, PAI and GAI), the fraction of intercepted photosynthetically active radiation (FIPAR), and the fraction of vegetation cover (FCOVER), are key descriptors of vegetation condition, and are useful in applications such as weather and climate modelling, agricultural and forest monitoring, and understanding biogeochemical fluxes between the biosphere and atmosphere. Because of its versatility, digital hemispherical photography (DHP) is a popular technique for in situ estimation of these variables. Using images acquired by a digital camera equipped with a fisheye lens facilitating zenithal and azimuthal sampling, multi-angular estimates of gap fraction can be derived, from which canopy biophysical variables can be estimated. Advantages of DHP include (a) the possibility of simultaneously deriving multiple canopy biophysical variables (including LAI/PAI/GAI, FIPAR and FCOVER; Li et al., 2015, 2021; Weiss et al., 2014), (b) the ability to use upward- and downward-facing images to sample both tall and short (or understory) canopies (Demarez et al., 2008; Garrigues, Shabanov, et al., 2008), (c) lower cost and reduced sensitivity to illumination conditions than other passive optical techniques such as ceptometry and the LI-COR LAI-2000 series of instruments (Garrigues, Shabanov, et al., 2008) and (d) provision of a permanent visual record of the canopy, which may be reanalysed with advances in data processing methods (Chianucci & Cutini, 2012).

A range of software has been developed for deriving canopy biophysical variables from DHP, including commercial and freely available solutions, proving a valuable resource to the community over the last 20 years (Supporting Information A). However, the majority are based on a graphical user interface, precluding automated, reproducible, and efficient end-to-end analysis within a single programming environment, even in the case of those supporting batch processing. Fewer still support the analysis of RAW image formats, necessitating pre-analysis conversion. Because formats such as JPEG result in a loss of information, it is argued RAW is 'the only scientifically justifiable file format' (Verhoeven, 2010). Macfarlane et al. (2014) demonstrate use of JPEG images increases sensitivity of gap fraction to photographic exposure, recommending use of the RAW format, which retains original sensor outputs and has a wider dynamic range, enabling detail to be recovered even for poorly exposed images. Finally, to our knowledge, no existing packages support propagation or provision of uncertainties (e.g. due to variability in gap fraction) in the derived variables. Uncertainties are an

increasingly important requirement in a range of applications, enabling observations to be weighted based on their quality (Brown, Camacho, et al., 2021; Raupach et al., 2005; Richardson et al., 2011). For example, the Global Climate Observing System (GCOS, 2019) requires FAPAR and LAI uncertainties to be within 10% and 15%, respectively, for climate modelling and adaptation studies, making uncertainty quantification essential for assessing measurement suitability and fitness-for-purpose.

In this paper, we present HemiPy v0.1.2 (<https://github.com/luke-a-brown/hemipy>), an open-source Python module for deriving forest biophysical variables and uncertainties from DHP images in an automated manner. HemiPy is well-suited to batch processing and supports a wide range of image formats (Supporting Information A). We then evaluate HemiPy using simulated hemispherical images (Section 3), multiannual time-series and litterfall data collected at several forested sites within the National Ecological Observatory Network (NEON) (Sections 4 and 5) and comparison against the CAN-EYE software package (Supporting Information K).

2 | DESCRIPTION AND FEATURES

2.1 | Overview

HemiPy (Figure 1) processes all images within a directory together to provide a single value (and uncertainty) for each canopy biophysical variable (Supporting Information B). Therefore, each directory should correspond to a single measurement plot, which typically contains between five and 15 images (Baret et al., 2005; Campbell et al., 1999; Demarez et al., 2008; Fernandes et al., 2014; Weiss et al., 2004). Though it is technically possible to process a directory containing a single image, this is discouraged, as more images provide a better estimate of the biophysical variable and its uncertainty. The maximum number of images it is possible to process, as well as processing time, will depend on image resolution and the computing environment's available memory (Supporting Information C).

2.2 | RAW image pre-processing

In the case of RAW images, pre-processing is carried out by default, as recommended by Macfarlane et al. (2014). This involves gamma correction and contrast stretching, such that 1% of pixels at the high

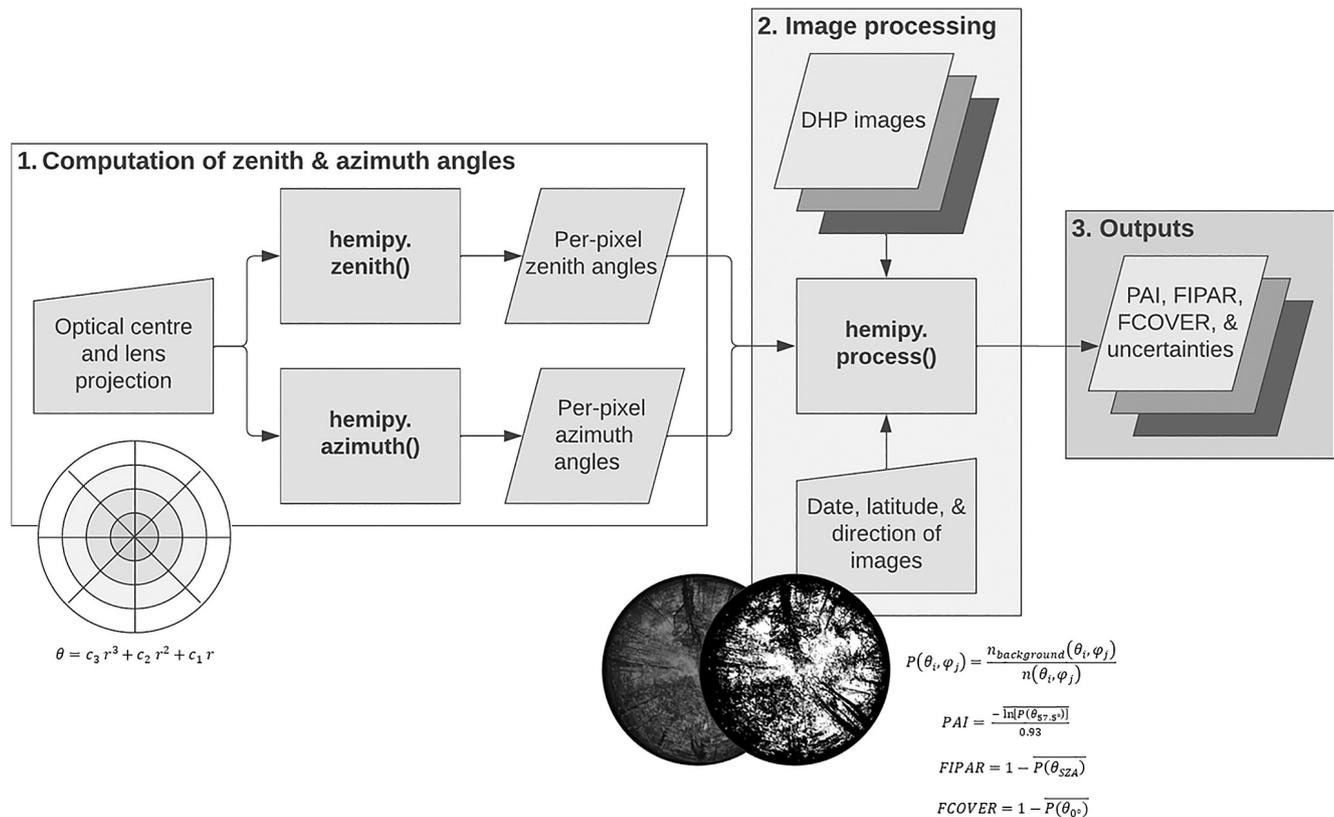


FIGURE 1 Overview of HemiPy's core functionality.

and low ends of the histogram are saturated. Pre-processed images are then stored in 8-bit form to speed up subsequent processing steps. For 8-bit formats such as JPEG, pre-processing is not carried out. Processing of RAW images in full bit depth is also possible (Supporting Information B).

2.3 | Derivation of multi-angular gap fraction

Gap fraction is determined using an automated binary image classification that depends on the image direction. For upward-facing images, Ridler and Calvard's (1978) clustering algorithm is used to separate sky and canopy pixels, as it was shown to be the most robust of 35 algorithms tested by Jonckheere et al. (2005). In this case, only the blue band is used to maximise contrast and minimise chromatic aberration and within-canopy multiple scattering (Leblanc et al., 2005; Macfarlane et al., 2007, 2014; Zhang et al., 2005). For downward-facing images, Meyer and Neto's (2008) approach is adopted to separate green vegetation from the underlying soil background (Supporting Information D).

Once classified, each image is divided into zenith rings and azimuth cells, the number of which is dependent on the specified angular resolution (Supporting Information E). Gap fraction is then computed as

$$P(\theta_i, \varphi_j) = \frac{n_{\text{background}}(\theta_i, \varphi_j)}{n(\theta_i, \varphi_j)}, \quad (1)$$

where θ_i and φ_j are the central angles of zenith ring i and azimuth cell j , $n_{\text{background}}$ is the number of pixels classified as background, and n is the total number of pixels, respectively.

2.4 | Computation of PAI

Due to limitations inherent to the image classification, HemiPy computes PAI or GAI rather than LAI. Upward-facing images represent PAI, as the image classification is sensitive to all canopy elements (Bréda, 2003), whereas downward-facing images represent GAI, as the image classification is sensitive to green elements (which may include stems as well as foliage; Baret et al., 2010). For brevity, the terms PAI and GAI are used interchangeably hereafter.

Two approaches are implemented to estimate effective PAI (PAI_e), in which a random distribution of plant material is assumed: a method derived from Warren-Wilson's (1963) approach, which considers gap fraction at the hinge region surrounding 57.5° only (where gap fraction is nearly independent of leaf angle distribution) (Baret et al., 2010), and a generalised version of Miller's (1967) integral, which uses a fuller range of multi-angular observations. In both cases, to account for foliage clumping and derive PAI as opposed to PAI_e , Lang and Yueqin's (1986) method is adopted. Thus, PAI_e and PAI according to the hinge approach are computed as

$$\text{Hinge } PAI_e = \frac{-\ln[P(\theta_{57.5})]}{0.5/\cos(57.5^\circ)} = \frac{-\ln[P(\theta_{57.5})]}{0.93}, \quad (2)$$

$$\text{Hinge PAI} = \frac{-\overline{\ln[P(\theta_{57.5})]}}{0.5/\cos(57.5^\circ)} = \frac{-\overline{\ln[P(\theta_{57.5})]}}{0.93}, \quad (3)$$

where $\overline{P(\theta_{57.5})}$ is the mean gap fraction in a zenith ring centred on 57.5° over all azimuth cells and images, while $\overline{\ln[P(\theta_{57.5})]}$ is the mean of the natural logarithm of gap fraction values in a zenith ring centred on 57.5° over all azimuth cells and images. Likewise, PAI_e and PAI according to the generalised Miller approach are calculated as

$$\text{Miller PAI}_e = 2 \sum_{i=1}^n -\ln[\overline{P(\theta_i)}] \cos(\theta_i) w_i, \quad (4)$$

$$\text{Miller PAI} = 2 \sum_{i=1}^n -\overline{\ln[P(\theta_i)]} \cos(\theta_i) w_i, \quad (5)$$

where $\overline{P(\theta_i)}$ is the mean gap fraction in a zenith ring centred at θ_i over all azimuth cells and images, $\overline{\ln[P(\theta_i)]}$ is the mean of the natural logarithm of gap fraction values in a zenith ring centred at θ_i over all azimuth cells and images, w_i is the weight associated to the zenith ring, and n is the number of rings. Weights are computed to sum to one, accounting for the range of sampled zenith angles (which may be less than 90°) in the same way as the LI-COR LAI-2200 instrument (LI-COR, 2013), such that

$$w_i = \frac{\sin(\theta_i) d\theta_i}{\sum_{i=1}^n \sin(\theta_i) d\theta_i}, \quad (6)$$

where $d\theta_i$ is the range of zenith angles covered by the ring. The last ring is not weighted as if it extends down to the horizon (as is the case for the LAI-2200's predecessor the LAI-2000; Welles & Norman, 1991), since Leblanc and Chen (2001) demonstrate this leads to systematic underestimation of PAI.

For Equations 3 and 5, a problem arises in dense canopies when an azimuth cell contains no gaps, as the natural logarithm of zero is undefined. Discarding these cells from the computation would underestimate PAI, while adding a gap of one pixel is undesirable due to dependency on image resolution (Leblanc et al., 2005; Yan et al., 2019). In HemiPy, such cells are assigned a gap fraction according to an arbitrary user-defined 'saturated' PAI value (Supporting Information F).

2.5 | Computation of FIPAR

In HemiPy, FIPAR is computed according to the instantaneous black-sky (i.e. direct illumination) definition, such that

$$\text{FIPAR} = 1 - \overline{P(\theta_{\text{SZA}})}, \quad (7)$$

where $\overline{P(\theta_{\text{SZA}})}$ is the mean gap fraction in a zenith ring centred at the solar zenith angle at the local solar time defined as an input to HemiPy, over all azimuth cells and images. A default value of 10:00 local solar time is adopted, because in addition to providing consistency with satellite-derived fraction of absorbed photosynthetically active radiation (FAPAR) products, previous work has shown instantaneous FIPAR

at 10:00 local solar time is a good approximation of daily integrated FIPAR (Baret et al., 2006, 2007; Li et al., 2015; Weiss & Baret, 2016).

2.6 | Computation of FCOVER

Corresponding to the fraction of ground covered by vegetation viewed from nadir, FCOVER is determined as

$$\text{FCOVER} = 1 - \overline{P(\theta_{0^\circ})}, \quad (8)$$

where $\overline{P(\theta_{0^\circ})}$ is the mean gap fraction in a zenith ring centred at nadir and extending to the maximum zenith angle defined as an input to HemiPy, over all azimuth cells and images (Li et al., 2015; Weiss & Baret, 2017).

2.7 | Uncertainty propagation

The propagation of uncertainties is handled in accordance with the International Standards Organisation (ISO) Guide to the Expression of Uncertainty in Measurement (GUM; Working Group 1 of the Joint Committee for Guides in Metrology, 2008), and in line with the recommendations of the Fiducial Reference Measurements for Vegetation (FRM4VEG) project (Brown, Camacho, et al., 2021). Uncertainties due to variability in the input quantity (gap fraction at a given zenith angle) are automatically and analytically propagated through the measurement equations described in Sections 2.4–2.6 using the uncertainties module (Lebigot, 2017). Since multiple gap fraction observations are available for a given zenith ring by virtue of (a) multiple azimuth cells, and (b) multiple images, variability in gap fraction is considered at within- and between-image scales, such that

$$u[\overline{P(\theta_i)}] = \sqrt{\left(\frac{1}{n} \sqrt{\sum_{j=1}^n u[\overline{P(\theta_i)_{\text{within}}}]_j^2}\right)^2 + u[\overline{P(\theta_i)_{\text{between}}}]^2}, \quad (9)$$

where $u[\overline{P(\theta_i)_{\text{within}}}]_j$ is the standard error of the mean gap fraction (or natural logarithm of mean gap fraction) over all azimuth cells in zenith ring i of image j , n is the number of images, and $u[\overline{P(\theta_i)_{\text{between}}}]$ is the standard error of the mean gap fraction (or natural logarithm of mean gap fraction) in zenith ring i , over all images (Brown, Camacho, et al., 2021). Note that HemiPy quantifies uncertainties due to random, but not systematic, effects.

3 | VERIFICATION USING SIMULATED HEMISPHERICAL IMAGES

To verify our implementation of the algorithms discussed in Section 2.4, a simulation-based approach was adopted. To assess estimates of PAI_e , we used CANOPIX, as described by Schleppi et al. (2007), which enables simulation of hemispherical images for

a two-dimensional canopy of a specified PAI_e and mean leaf angle, based on Campbell's (1986) ellipsoidal leaf angle distribution. We simulated images for canopies with a mean leaf angle of 25°, 57° and 65°, representing planophile, random and erectophile leaf angle distributions, respectively (Leblanc & Fournier, 2014). In each case, we varied PAI_e from 0.3 to 7.5. Images simulated with CANOPIX were processed using HemiPy's default settings (Supporting Information B) but with no downsampling.

CANOPIX represents the canopy as a two-dimensional turbid medium (i.e. with a random distribution of leaves), so does not simulate foliage clumping (Gonsamo et al., 2018). To assess HemiPy's estimates of PAI as opposed to PAI_e , a three-dimensional simulation approach was required. We simulated 10 hemispherical images for 75 virtual forest scenes (PAI_e ranging from 0.3 to 7.5), rendered using the Persistence of Vision Raytracer (POV-Ray). Each scene could represent a Neyman, plantation, or random stem distribution (Supporting Information G). As in the CANOPIX simulations, all 75 scenes were simulated with planophile, random, and erectophile leaf angle distributions. Parameters used in the generation of the virtual forest scenes and POV-Ray simulations are described by Leblanc and Fournier (2014), with the exception that simulations in this study were carried out at increased resolution (4000×4000 compared to 4000×3000 pixels). Images simulated by POV-Ray were processed using HemiPy's default settings (Supporting Information B). Since FIPAR and FCOVER were not input parameters into CANOPIX or POV-Ray, these variables were not investigated.

Assessed against CANOPIX simulations, PAI_e values derived according to the hinge approach provided estimates closest to those used to simulate the images (Figure 2a), with low bias (−0.01 to 0.06), RMSE (0.03 to 0.06), and normalised RMSE (NRMSE, i.e. the RMSE divided by the mean of the reference values; 0.75% to 1.67%; Table 1). As expected, the approach was insensitive to leaf angle distribution. In contrast, PAI_e values derived according to the generalised Miller approach were subject to increased bias (−0.32 to 1.29), RMSE (0.02 to 1.50) and NRMSE (0.48% to 38.41%), with differences depending strongly on leaf angle distribution (Figure 2b).

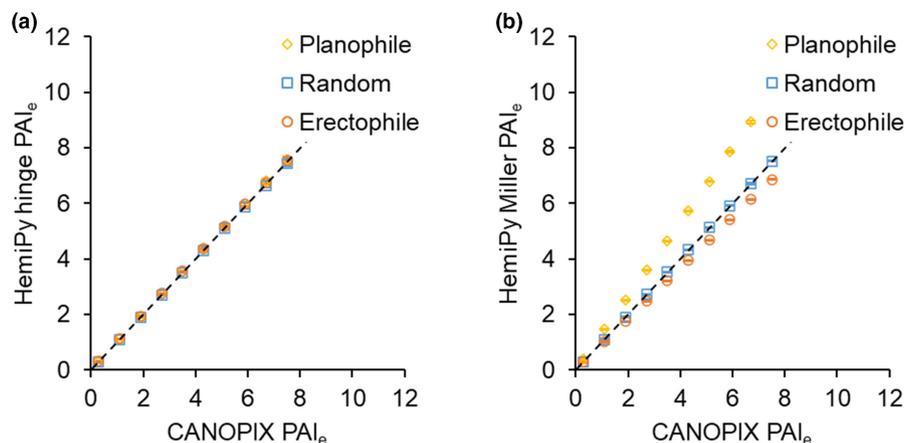


FIGURE 2 Comparison of HemiPy's hinge (a) and generalised Miller (b) effective plant area index (PAI_e) outputs for simulated hemispherical images of a given PAI_e and leaf angle distribution generated with CANOPIX.

The smallest bias occurred in the case of the random leaf angle distribution, while the planophile leaf angle distribution was characterised by the largest bias (Table 1). While not a verification criterion in and of itself, these results corroborate the findings of Leblanc and Fournier (2014), who suggest the hinge approach provides more stable estimates for canopies with different leaf angle distributions and reflect the more recent results of Liu et al. (2021).

When assessed against the three-dimensional POV-Ray simulations, similar results were obtained, with the best agreement for PAI_e outputs derived according to the hinge approach (Figure 3a and Table 2; bias = −0.50 to −0.78, RMSE = 0.97 to 1.16, NRMSE = 34.35% to 45.86%). As with the CANOPIX simulations, greater biases were observed when PAI_e was derived according to the generalised Miller approach (−0.40 to −1.05), leading to increased RMSE and NRMSE (1.09 to 1.38, 38.84% to 49.71%; Figure 3b and Table 2). For the same POV-Ray simulations (performed at a resolution of 4000×3000 pixels), Leblanc and Fournier (2014) achieved similar results for DHP-TRACWin's implementation of the hinge and generalised Miller approaches (RMSE = 1.00 to 1.10).

4 | DEMONSTRATION OF HEMIPY'S OUTPUTS

To demonstrate the use of HemiPy, we present time-series for four deciduous broadleaf and two evergreen needleleaf forest sites within NEON (Supporting Information H), for which RAW DHP images have been processed under the Copernicus Ground Based Observations for Validation (GBOV) service using HemiPy's default settings (Brown et al., 2020; Brown, Fernandes, et al., 2021; data available at <https://land.copernicus.eu/global/gbov>). At each site, DHP images are collected in at least three 20m×20m measurement plots every 2 weeks throughout the growing season, using a Nikon D750, D800 or D810 digital single lens reflex (DSLR) camera equipped with an AF Fisheye-Nikkor 16 mm f/2.8D lens (Meier et al., 2018; NEON, 2019a). At forested sites, each measurement

| Method | Leaf angle distribution | r | RMSE | NRMSE (%) | Bias |
|--------------------|-------------------------|------|------|-----------|-------|
| Hinge | Planophile | 1.00 | 0.06 | 1.67 | 0.06 |
| | Random | 1.00 | 0.03 | 0.75 | -0.01 |
| | Erectophile | 1.00 | 0.05 | 1.36 | 0.05 |
| | Overall | 1.00 | 0.05 | 1.31 | 0.03 |
| Generalised Miller | Planophile | 1.00 | 1.50 | 38.41 | 1.29 |
| | Random | 1.00 | 0.02 | 0.48 | 0.02 |
| | Erectophile | 1.00 | 0.37 | 9.55 | -0.32 |
| | Overall | 0.95 | 0.89 | 22.85 | 0.33 |

TABLE 1 Performance statistics associated with HemiPy's hinge and generalised Miller effective plant area index (PAI_e) outputs for simulated hemispherical images of a specified PAI_e and leaf angle distribution generated with CANOPIX.

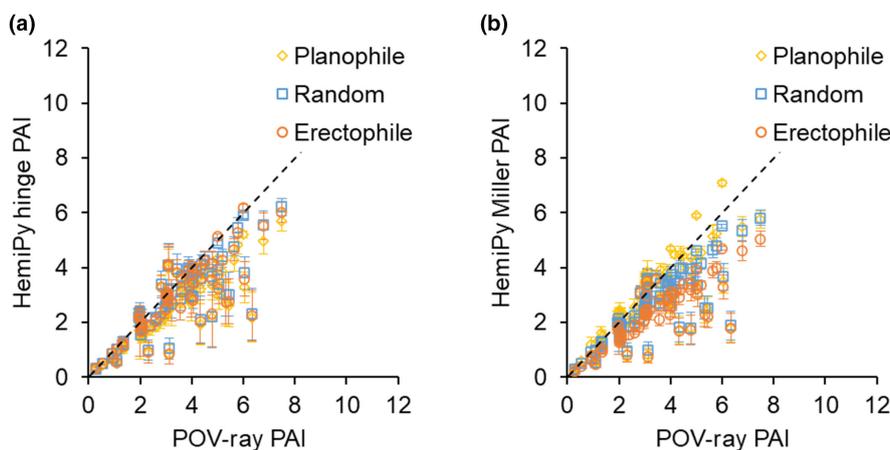


FIGURE 3 Comparison of HemiPy's hinge (a) and generalised Miller (b) plant area index (PAI) outputs for simulated hemispherical images of a given PAI and leaf angle distribution generated with POV-Ray. Error bars represent expanded uncertainties at the $k=3$ coverage interval.

| Method | Leaf angle distribution | r | RMSE | NRMSE (%) | Bias |
|--------------------|-------------------------|------|------|-----------|-------|
| Hinge | Planophile | 0.84 | 1.16 | 45.86 | -0.78 |
| | Random | 0.85 | 0.97 | 34.35 | -0.50 |
| | Erectophile | 0.83 | 1.03 | 37.04 | -0.55 |
| | Overall | 0.84 | 1.06 | 31.81 | -0.61 |
| Generalised Miller | Planophile | 0.76 | 1.10 | 43.43 | -0.40 |
| | Random | 0.84 | 1.09 | 38.84 | -0.67 |
| | Erectophile | 0.83 | 1.38 | 49.71 | -1.05 |
| | Overall | 0.79 | 1.20 | 36.12 | -0.71 |

TABLE 2 Performance statistics associated with HemiPy's hinge and generalised Miller plant area index (PAI) outputs for simulated hemispherical images of a specified PAI and leaf angle distribution generated with POV-Ray.

plot contains 12 sampling locations configured in a 'cross' at which upward- and downward-facing images are acquired. PAI, GAI, FIPAR and FCOVER values were combined to provide total values (Supporting Information I).

Time-series of PAI, FIPAR and FCOVER derived from the combination of upward- and downward-facing images using HemiPy revealed realistic temporal patterns. At the deciduous sites, annual phenological events including green-up, peak of the growing

season, and senescence could be clearly identified (Figures 4–6), highlighting the information provided by the processed images and the consistency of HemiPy's outputs. As expected, phenological events were less apparent at the evergreen sites (Figures 4–6). Each time-series exhibited a distinct shape and timing, demonstrating the complementary information the different variables can provide on canopy structure and interception of radiation (Figures 4–6).

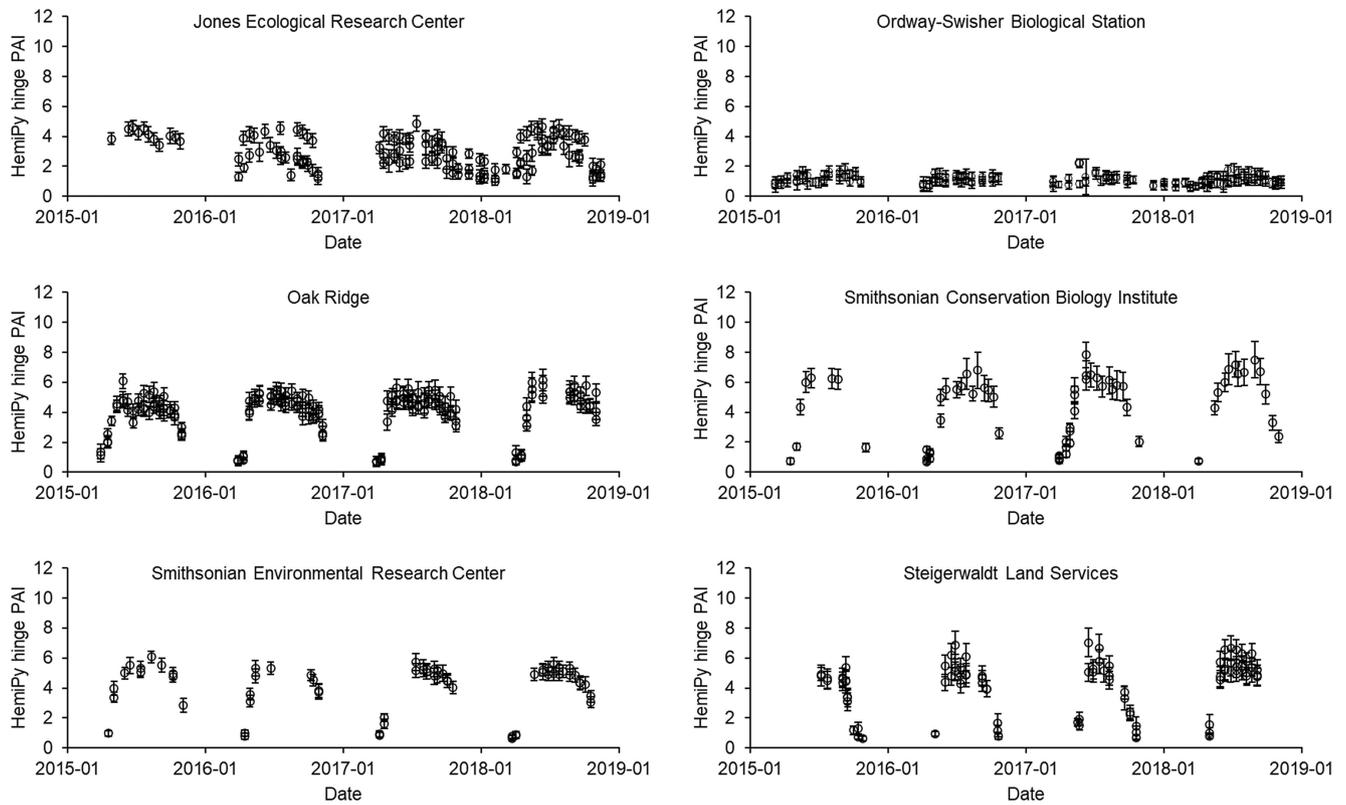


FIGURE 4 Time-series of plant area index values at six forested National Ecological Observatory Network sites derived from the combination of upward- and downward-facing digital hemispherical photographs using HemiPy's implementation of the hinge approach. Error bars represent expanded uncertainties at the $k=3$ coverage interval.

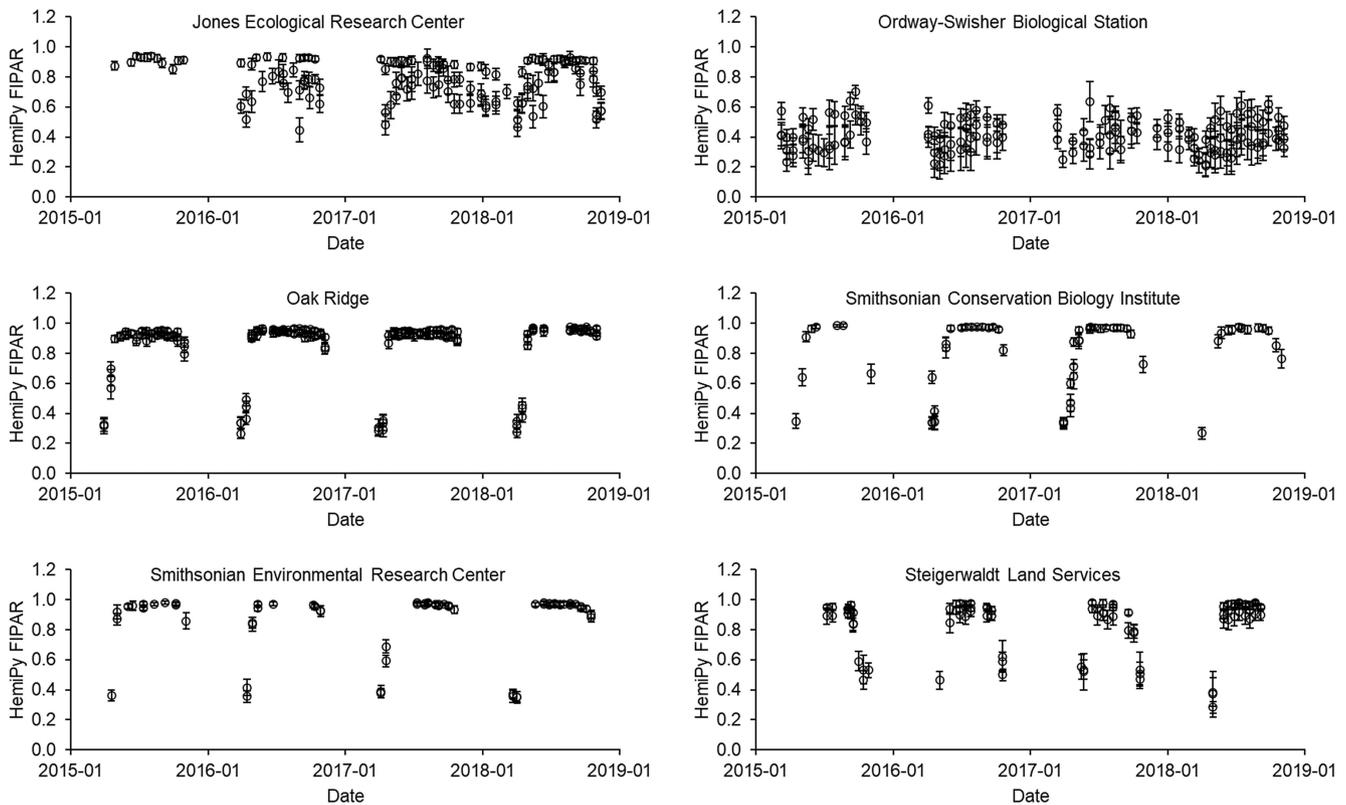


FIGURE 5 Time-series of fraction of intercepted photosynthetically active radiation values at six forested National Ecological Observatory Network sites derived from the combination of upward- and downward-facing digital hemispherical photographs using HemiPy. Error bars represent expanded uncertainties at the $k=3$ coverage interval.

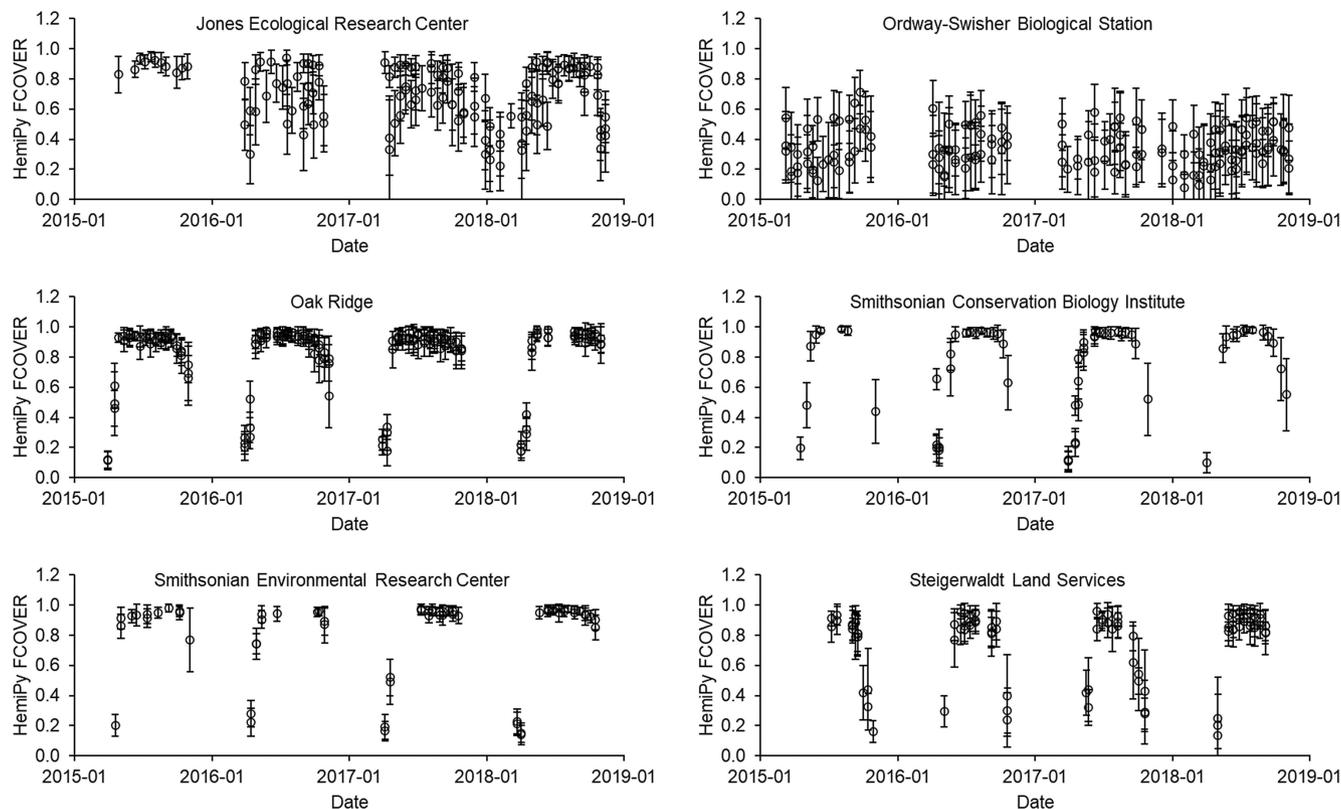


FIGURE 6 Time-series of fraction of vegetation cover values at six forested National Ecological Observatory Network sites derived from the combination of upward- and downward-facing digital hemispherical photographs using HemiPy. Error bars represent expanded uncertainties at the $k=3$ coverage interval.

5 | EVALUATION AGAINST LITTERFALL DATA

At each of the deciduous sites described in Section 4, NEON acquire litterfall observations in the same plots as DHP images, reporting the dry mass of leaves collected from littertraps throughout the senescent period (NEON, 2019b). For each site and year, total litterfall (i.e. peak) LAI could, therefore, be determined by dividing the dry mass of leaves collected during senescence (defined here as September, October, November and December) by leaf mass per area (LMA) and correcting for the size of the littertraps. At each site, LMA was determined as the mean of all LMA measurements reported in ‘tower’ plots, which are in close proximity to the littertraps (NEON, 2019c). Total litterfall LAI was compared with the maximum PAI value from upward-facing DHP images for the year in question. To account for the fact that PAI outputs incorporate the contribution of woody material, we also computed LAI by subtracting wood area index (WAI), which was determined as the minimum PAI value observed before leaf emergence at each site (Supporting Information J).

The best agreement between litterfall observations and HemiPy-derived PAI was for values derived according to the hinge approach, which slightly overestimated litterfall LAI (bias=0.20, $r=0.50$, RMSE=1.14, NRMSE=43.45%; Figure 7a). Despite a higher correlation ($r=0.60$), PAI derived according to the generalised Miller

approach was subject to substantially increased overestimation (bias=0.83, RMSE=1.42, NRMSE=54.26%; Figure 7b). For both methods, some degree of overestimation is to be expected, as the HemiPy-derived PAI values incorporate the contribution of woody material, whereas the litterfall data do not. The degree of overestimation observed is within the range of WAI values reported for various deciduous species (Bréda, 2003; Gower et al., 1999).

When DHP-derived PAI was converted to LAI through subtraction of WAI, the hinge approach demonstrated slightly reduced correlation ($r=0.43$), and increased RMSE (1.22) and NRMSE (46.58%) compared to the generalised Miller estimate ($r=0.54$, RMSE=1.18 and NRMSE=45.12%), however bias was lower for the former (−0.37) than the latter (0.47) (Figure 7c,d). Previous work has attributed errors of ~1 unit to optical LAI measurement approaches (Camacho et al., 2013; Fernandes et al., 2003; Garrigues, Lacaze, et al., 2008), and the RMSE values obtained in this study are close to this value. It is also worth noting that the National Land Cover Database (NLCD) definition of ‘deciduous forest’ used to characterise NEON measurement plots allows up to 25% evergreen trees (which would lead to further discrepancies between litterfall and HemiPy-derived LAI estimates; Homer et al., 2020). Nevertheless, LAI derived according to the hinge approach again provided the lowest bias, corroborating our comparison against simulated images (Section 3) and the results of Leblanc and Fournier (2014) and Liu et al. (2021).

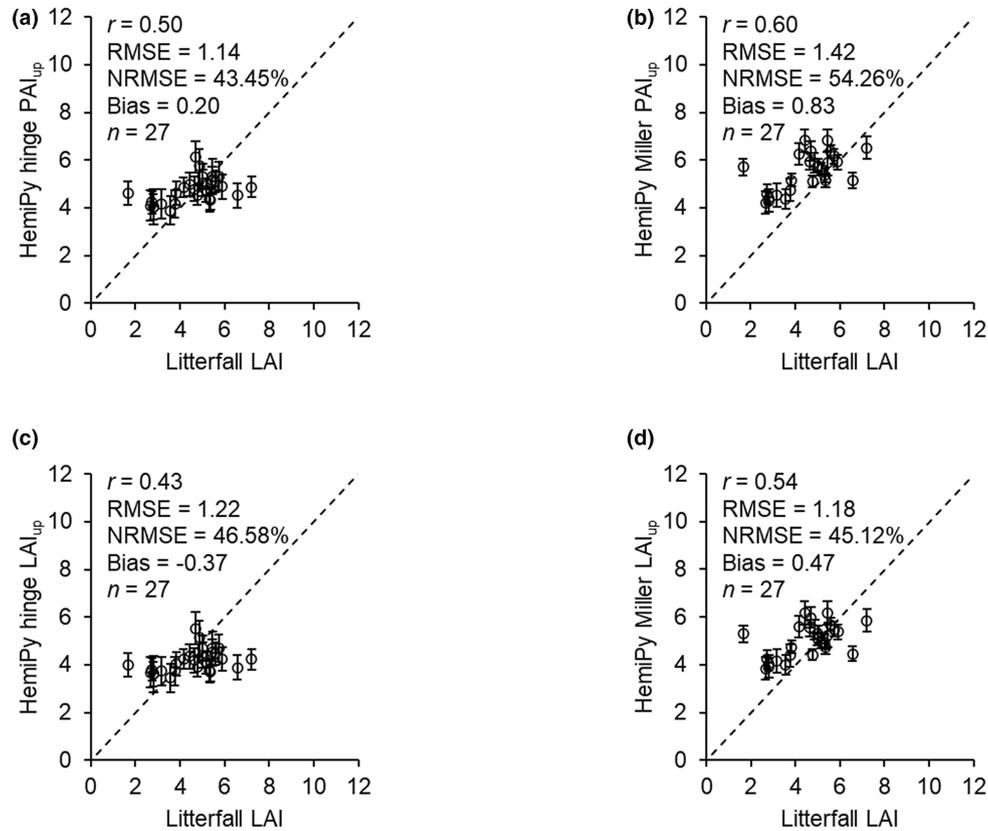


FIGURE 7 Comparison of HemiPy's maximum hinge (a, c) and generalised Miller (b, d) upward-facing plant area index and leaf area index (LAI) outputs in each year with respect to the corresponding total litterfall LAI. Error bars represent expanded uncertainties at the $k=3$ coverage interval.

6 | SUMMARY

In this paper, the HemiPy Python module is presented for automated derivation of forest biophysical variables from DHP images. Features include the ability to process images programmatically, support for RAW images, and propagation and provision of uncertainties. We used simulated hemispherical images to verify PAI_e and PAI estimates derived via HemiPy algorithms. Consistent with previous studies, better agreement was observed for values derived using gap fraction near the hinge angle of 57.5° only, as opposed to values derived using gap fraction over a wider range of zenith angles. Multiannual time-series of PAI, FIPAR and FCOVER were presented, revealing realistic temporal patterns. Comparison of HemiPy-derived PAI and LAI values against litterfall data at four deciduous NEON sites revealed reasonable accuracies, while good agreement was observed between HemiPy and CAN-EYE PAI, FIPAR and FCOVER (Supporting Information K). Based on our findings, HemiPy should prove a useful tool for DHP processing. Its open-source nature means that it can be adopted, extended and further refined by the user community.

AUTHOR CONTRIBUTIONS

Conceptualisation: Luke A. Brown. Methodology: Luke A. Brown and Sylvain Leblanc. Software: Luke A. Brown, Harry Morris, Sylvain

Leblanc and Gabriele Bai. Validation: Luke A. Brown, Harry Morris, Sylvain Leblanc and Gabriele Bai. Formal analysis: Luke A. Brown and Sylvain Leblanc. Investigation: Luke A. Brown and Sylvain Leblanc. Resources: Luke A. Brown, Sylvain Leblanc, Gabriele Bai, Courtney Meier and Jadunandan Dash. Data curation: Luke A. Brown, Harry Morris, Sylvain Leblanc, Gabriele Bai and Courtney Meier. Writing—original draft preparation: Luke A. Brown. Writing—review and editing: Luke A. Brown, Harry Morris, Sylvain Leblanc, Gabriele Bai, Christian Lanconelli, Nadine Gobron, Courtney Meier and Jadunandan Dash. Visualisation: Luke A. Brown and Sylvain Leblanc. Supervision: Luke A. Brown, Nadine Gobron and Jadunandan Dash. Project administration: Luke A. Brown, Nadine Gobron and Jadunandan Dash. Funding acquisition: Luke A. Brown and Jadunandan Dash.

ACKNOWLEDGEMENTS

The authors are grateful to Fernando Camacho and Beatriz Fuster of EOLAB for collecting the DHP data at Wytham Woods and processing them with CAN-EYE, and to Nigel Fisher and the University of Oxford for facilitating access to Wytham Woods. The authors also thank the Associate Editor and the two anonymous reviewers for their constructive comments, which helped to substantially improve the manuscript. This activity was carried out under the Living Planet Fellowship, a programme of and funded by the European Space Agency. The view expressed in this publication can in no

way be taken to reflect the official opinion of the European Space Agency. This study has been undertaken using data from GBOV 'Ground Based Observations for Validation' (<https://land.copernicus.eu/global/gbov>) funded by the European Commission Joint Research Centre FWC 932059, part of the Global Component of the European Union's Copernicus Land Monitoring Service. The National Ecological Observatory Network is a Programme sponsored by the National Science Foundation and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the National Science Foundation through the NEON Programme. The study has been undertaken using data from "Fiducial Reference Measurements for Vegetation—Phase 2" (FRM4VEG—Phase 2), which was funded by the European Space Agency.

CONFLICT OF INTEREST STATEMENT

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.

DATA AVAILABILITY STATEMENT

Raw DHP, litterfall and LMA data used in this study are available through the NEON data portal (<https://data.neonscience.org>), while processed DHP data are available through the GBOV portal (<https://land.copernicus.eu/global/gbov>) and FRM4VEG project site (<https://frm4veg.org>). The HemiPy module is available via GitHub (<https://github.com/luke-a-brown/hemipy>), and the version described in this study is archived on Zenodo (<https://doi.org/10.5281/zenodo.8187645>; Brown, 2023).

ORCID

Luke A. Brown  <https://orcid.org/0000-0003-4807-9056>

Harry Morris  <https://orcid.org/0000-0001-9049-2819>

Sylvain Leblanc  <https://orcid.org/0000-0003-2456-7119>

Gabriele Bai  <https://orcid.org/0000-0003-2965-4131>

Christian Lanconelli  <https://orcid.org/0000-0002-9545-1255>

Nadine Gobron  <https://orcid.org/0000-0002-0584-4195>

Courtney Meier  <https://orcid.org/0000-0003-3576-883X>

Jadunandan Dash  <https://orcid.org/0000-0002-5444-2109>

REFERENCES

- Baret, F., de Solan, B., Lopez-Lozano, R., Ma, K., & Weiss, M. (2010). GAI estimates of row crops from downward looking digital photos taken perpendicular to rows at 57.5° zenith angle: Theoretical considerations based on 3D architecture models and application to wheat crops. *Agricultural and Forest Meteorology*, 150, 1393–1401. <https://doi.org/10.1016/j.agrformet.2010.04.011>
- Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., Roujean, J. L., & Leroy, M. (2007). LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION. *Remote Sensing of Environment*, 110, 275–286. <https://doi.org/10.1016/j.rse.2007.02.018>
- Baret, F., Morisette, J. T., Fernandes, R. A., Champeaux, J. L., Myneni, R. B., Chen, J., Plummer, S., Weiss, M., Bacour, C., Garrigues, S., & Nickeson, J. E. (2006). Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 1794–1803. <https://doi.org/10.1109/TGRS.2006.876030>
- Baret, F., Weiss, M., Allard, D., Garrigues, S., Leroy, M., Jeanjean, H., Fernandes, R., Myneni, R., Privette, J., Morisette, J., & Bohbot, H. (2005). VALERI: A network of sites and a methodology for the validation of medium spatial resolution land satellite products. Institut National de la Recherche Agronomique.
- Bréda, N. J. J. (2003). Ground-based measurements of leaf area index: A review of methods, instruments and current controversies. *Journal of Experimental Botany*, 54, 2403–2417. <https://doi.org/10.1093/jxb/erg263>
- Brown, L. A. (2023). luke-a-brown/hemipy: v0.1.2 (v0.1.2). Zenodo. <https://doi.org/10.5281/zenodo.8187645>
- Brown, L. A., Camacho, F., García-Santos, V., Origo, N., Fuster, B., Morris, H., Pastor-Guzman, J., Sánchez-Zapero, J., Morrone, R., Ryder, J., Nightingale, J., Boccia, V., & Dash, J. (2021). Fiducial reference measurements for vegetation biogeophysical variables: An end-to-end uncertainty evaluation framework. *Remote Sensing*, 13, 3194. <https://doi.org/10.3390/rs13163194>
- Brown, L. A., Fernandes, R., Djamaï, N., Meier, C., Gobron, N., Morris, H., Canisius, F., Bai, G., Lerebourg, C., Lanconelli, C., Clerici, M., & Dash, J. (2021). Validation of baseline and modified Sentinel-2 Level 2 Prototype Processor leaf area index retrievals over the United States. *ISPRS Journal of Photogrammetry and Remote Sensing*, 175, 71–87. <https://doi.org/10.1016/j.isprsjprs.2021.02.020>
- Brown, L. A., Meier, C., Morris, H., Pastor-Guzman, J., Bai, G., Lerebourg, C., Gobron, N., Lanconelli, C., Clerici, M., & Dash, J. (2020). Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over North America using Copernicus Ground Based Observations for Validation data. *Remote Sensing of Environment*, 247, 111935. <https://doi.org/10.1016/j.rse.2020.111935>
- Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., & Weiss, M. (2013). GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sensing of Environment*, 137, 310–329. <https://doi.org/10.1016/j.rse.2013.02.030>
- Campbell, G. S. (1986). Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. *Agricultural and Forest Meteorology*, 36, 317–321. [https://doi.org/10.1016/0168-1923\(86\)90010-9](https://doi.org/10.1016/0168-1923(86)90010-9)
- Campbell, J. L., Burrows, S., Gower, S. T., & Cohen, W. B. (1999). *BigFoot: Characterizing land cover, LAI and NPP at the landscape scale for EOS/MODIS validation—Field manual* (2.1 ed.). Oak Ridge National Laboratory.
- Chianucci, F., & Cutini, A. (2012). Digital hemispherical photography for estimating forest canopy properties: Current controversies and opportunities. *iForest-Biogeosciences and Forestry*, 5, 290–295. <https://doi.org/10.3832/ifer0775-005>
- Demarez, V., Duthoit, S., Baret, F., Weiss, M., & Dedieu, G. (2008). Estimation of leaf area and clumping indexes of crops with hemispherical photographs. *Agricultural and Forest Meteorology*, 148, 644–655. <https://doi.org/10.1016/j.agrformet.2007.11.015>
- Fernandes, R., Butson, C., Leblanc, S., & Latifovic, R. (2003). Landsat-5 TM and Landsat-7 ETM+ based accuracy assessment of leaf area index products for Canada derived from SPOT-4 VEGETATION data. *Canadian Journal of Remote Sensing*, 29, 241–258. <https://doi.org/10.5589/m02-092>
- Fernandes, R., Plummer, S., Nightingale, J., Baret, F., Camacho, F., Fang, H., Garrigues, S., Gobron, N., Lang, M., Lacaze, R., Leblanc, S., Meroni, M., Martinez, B., Nilson, T., Pinty, B., Pisek, J., Sonnentag, O., Verger, A., Welles, J., ... Nickeson, J. (2014). Global leaf

- area index product validation good practices. In R. Fernandes, S. Plummer, & J. Nightingale (Eds.), *Best practice for satellite-derived land product validation. Land product validation subgroup*. Committee on Earth Observation Satellites Working Group on Calibration and Validation. <https://doi.org/10.5067/doc/ceoswgcvc/lpv/lai.002>
- Garrigues, S., Lacaze, R., Baret, F., Morisette, J. T., Weiss, M., Nickeson, J. E., Fernandes, R., Plummer, S., Shabanov, N. V., Myneni, R. B., Knyazikhin, Y., & Yang, W. (2008). Validation and intercomparison of global leaf area index products derived from remote sensing data. *Journal of Geophysical Research*, 113, G02028. <https://doi.org/10.1029/2007JG000635>
- Garrigues, S., Shabanov, N. V., Swanson, K., Morisette, J. T., Baret, F., & Myneni, R. B. (2008). Intercomparison and sensitivity analysis of leaf area index retrievals from LAI-2000, AccuPAR, and digital hemispherical photography over croplands. *Agricultural and Forest Meteorology*, 148, 1193–1209. <https://doi.org/10.1016/j.agrformet.2008.02.014>
- GCOS. (2019). Essential climate variables [WWW Document]. <https://public.wmo.int/en/programmes/global-climate-observing-system/essential-climate-variables>
- Gonsamo, A., Walter, J.-M., Chen, J. M., Pellikka, P., & Schleppi, P. (2018). A robust leaf area index algorithm accounting for the expected errors in gap fraction observations. *Agricultural and Forest Meteorology*, 248, 197–204. <https://doi.org/10.1016/j.agrformet.2017.09.024>
- Gower, S. T., Kucharik, C. J., & Norman, J. M. (1999). Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, 70, 29–51. [https://doi.org/10.1016/S0034-4257\(99\)00056-5](https://doi.org/10.1016/S0034-4257(99)00056-5)
- Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., Gass, L., Funk, M., Wickham, J., Stehman, S., Auch, R., & Riitters, K. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, 184–199. <https://doi.org/10.1016/j.isprsjprs.2020.02.019>
- Jonckheere, I., Nackaerts, K., Muys, B., & Coppin, P. (2005). Assessment of automatic gap fraction estimation of forests from digital hemispherical photography. *Agricultural and Forest Meteorology*, 132, 96–114. <https://doi.org/10.1016/j.agrformet.2005.06.003>
- Lang, A. R. G., & Yueqin, X. (1986). Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. *Agricultural and Forest Meteorology*, 37, 229–243. [https://doi.org/10.1016/0168-1923\(86\)90033-X](https://doi.org/10.1016/0168-1923(86)90033-X)
- Lebigot, E. O. (2017). Uncertainties: A Python package for calculations with uncertainties [WWW Document]. <http://pythonhosted.org/uncertainties>
- Leblanc, S. G., & Chen, J. M. (2001). A practical scheme for correcting multiple scattering effects on optical LAI measurements. *Agricultural and Forest Meteorology*, 110, 125–139. [https://doi.org/10.1016/S0168-1923\(01\)00284-2](https://doi.org/10.1016/S0168-1923(01)00284-2)
- Leblanc, S. G., Chen, J. M., Fernandes, R., Deering, D. W., & Conley, A. (2005). Methodology comparison for canopy structure parameters extraction from digital hemispherical photography in boreal forests. *Agricultural and Forest Meteorology*, 129, 187–207. <https://doi.org/10.1016/j.agrformet.2004.09.006>
- Leblanc, S. G., & Fournier, R. A. (2014). Hemispherical photography simulations with an architectural model to assess retrieval of leaf area index. *Agricultural and Forest Meteorology*, 194, 64–76. <https://doi.org/10.1016/j.agrformet.2014.03.016>
- Li, W., Fang, H., Wei, S., Weiss, M., & Baret, F. (2021). Critical analysis of methods to estimate the fraction of absorbed or intercepted photosynthetically active radiation from ground measurements: Application to rice crops. *Agricultural and Forest Meteorology*, 297, 108273. <https://doi.org/10.1016/j.agrformet.2020.108273>
- Li, W., Weiss, M., Waldner, F., Defourny, P., Demarez, V., Morin, D., Hagolle, O., & Baret, F. (2015). A generic algorithm to estimate LAI, FAPAR and FCOVER variables from SPOT4_HRVIR and Landsat sensors: Evaluation of the consistency and comparison with ground measurements. *Remote Sensing*, 7, 15494–15516. <https://doi.org/10.3390/rs71115494>
- LI-COR. (2013). *LAI-2200C Plant Canopy Analyser instruction manual*. LI-COR.
- Liu, J., Li, L., Akerblom, M., Wang, T., Skidmore, A., Zhu, X., & Heurich, M. (2021). Comparative evaluation of algorithms for leaf area index estimation from digital hemispherical photography through virtual forests. *Remote Sensing*, 13, 3325. <https://doi.org/10.3390/rs13163325>
- Macfarlane, C., Grigg, A., & Evangelista, C. (2007). Estimating forest leaf area using cover and fullframe fisheye photography: Thinking inside the circle. *Agricultural and Forest Meteorology*, 146, 1–12. <https://doi.org/10.1016/j.agrformet.2007.05.001>
- Macfarlane, C., Ryu, Y., Ogden, G. N., & Sonnentag, O. (2014). Digital canopy photography: Exposed and in the raw. *Agricultural and Forest Meteorology*, 197, 244–253. <https://doi.org/10.1016/j.agrformet.2014.05.014>
- Meier, C., Everhart, J., & Jones, K. (2018). *TOS protocol and procedure: Measurement of leaf area index* (K. ed.). National Ecological Observatory Network.
- Meyer, G. E., & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, 63, 282–293. <https://doi.org/10.1016/j.compag.2008.03.009>
- Miller, J. (1967). A formula for average foliage density. *Australian Journal of Botany*, 15, 141–144. <https://doi.org/10.1071/BT9670141>
- NEON (National Ecological Observatory Network). (2019a). Digital hemispheric photos of plot vegetation (DP1.10017.001). <https://data.neonscience.org>
- NEON (National Ecological Observatory Network). (2019b). Litterfall and fine woody debris production and chemistry (DP1.10033.001). <https://data.neonscience.org>
- NEON (National Ecological Observatory Network). (2019c). Plant foliar traits (DP1.10026.001). <https://data.neonscience.org>
- Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., & Schimullius, C. C. (2005). Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, 11, 378–397. <https://doi.org/10.1111/j.1365-2486.2005.00917.x>
- Richardson, A. D., Dail, D. B., & Hollinger, D. Y. (2011). Leaf area index uncertainty estimates for model-data fusion applications. *Agricultural and Forest Meteorology*, 151, 1287–1292. <https://doi.org/10.1016/j.agrformet.2011.05.009>
- Ridler, T. W., & Calvard, S. (1978). Picture thresholding using an iterative selection method. *IEEE Transactions on Systems, Man, and Cybernetics*, 8, 630–632. <https://doi.org/10.1109/TSMC.1978.4310039>
- Schleppi, P., Conedera, M., Sedivy, I., & Thimonier, A. (2007). Correcting non-linearity and slope effects in the estimation of the leaf area index of forests from hemispherical photographs. *Agricultural and Forest Meteorology*, 144, 236–242. <https://doi.org/10.1016/j.agrformet.2007.02.004>
- Verhoeven, G. J. J. (2010). It's all about the format—Unleashing the power of RAW aerial photography. *International Journal of Remote Sensing*, 31, 2009–2042. <https://doi.org/10.1080/01431160902929271>
- Warren-Wilson, J. (1963). Estimation of foliage denseness and foliage angle by inclined point quadrats. *Australian Journal of Botany*, 11, 95–105.
- Weiss, M., & Baret, F. (2016). *S2ToolBox level 2 products: LAI, FAPAR, FCOVER (1.1 ed.)*. Institut National de la Recherche Agronomique.
- Weiss, M., & Baret, F. (2017). *CAN-EYE V6.4.91 user manual*. Institut National de la Recherche Agronomique.
- Weiss, M., Baret, F., Block, T., Koetz, B., Burini, A., Scholze, B., Lecharpentier, P., Brockmann, C., Fernandes, R., Plummer, S., Myneni, R., Gobron, N., Nightingale, J., Schaepman-Strub, G.,

- Camacho, F., & Sanchez-Azofeifa, A. (2014). On line validation exercise (OLIVE): A web based service for the validation of medium resolution land products. Application to FAPAR products. *Remote Sensing*, 6, 4190–4216. <https://doi.org/10.3390/rs6054190>
- Weiss, M., Baret, F., Smith, G. J., Jonckheere, I., & Coppin, P. (2004). Review of methods for in situ leaf area index (LAI) determination part II: Estimation of LAI, errors and sampling. *Agricultural and Forest Meteorology*, 121, 37–53. <https://doi.org/10.1016/j.agrformet.2003.08.001>
- Welles, J. M., & Norman, J. M. (1991). Instrument for indirect measurement of canopy architecture. *Agronomy Journal*, 83, 818–825. <https://doi.org/10.2134/agronj1991.00021962008300050009x>
- Working Group 1 of the Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data—Guide to the expression of uncertainty in measurement*. Bureau International des Poids et Mesures.
- Yan, G., Hu, R., Luo, J., Weiss, M., Jiang, H., Mu, X., Xie, D., & Zhang, W. (2019). Review of indirect optical measurements of leaf area index: Recent advances, challenges, and perspectives. *Agricultural and Forest Meteorology*, 265, 390–411. <https://doi.org/10.1016/j.agrformet.2018.11.033>
- Zhang, Y., Chen, J. M., & Miller, J. R. (2005). Determining digital hemispherical photograph exposure for leaf area index estimation.

Agricultural and Forest Meteorology, 133, 166–181. <https://doi.org/10.1016/j.agrformet.2005.09.009>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting Information A–K. Additional background information and verification results.

How to cite this article: Brown, L. A., Morris, H., Leblanc, S., Bai, G., Lanconelli, C., Gobron, N., Meier, C., & Dash, J. (2023). HemiPy: A Python module for automated estimation of forest biophysical variables and uncertainties from digital hemispherical photographs. *Methods in Ecology and Evolution*, 00, 1–12. <https://doi.org/10.1111/2041-210X.14199>