



## Evidence of a bias-variance trade off when correcting for bias in Sentinel 2 forest LAI retrievals using radiative transfer models

Richard Fernandes<sup>\*</sup>, Najib Djamai, Kate Harvey, Gang Hong, Camryn MacDougall, Hemit Shah, Lixin Sun

*Canada Centre for Remote Sensing, Natural Resources Canada, 580 Booth Street, Ottawa, Ontario K1A 0E4, Canada*



### ARTICLE INFO

Edited by Jing M. Chen

**Keywords:**  
Sentinel 2  
Leaf area index  
fAPAR  
PROSAIL  
Validation  
Model inversion

### ABSTRACT

Forest canopies exhibit spatial heterogeneity that impacts the relationship between essential climate variables such as leaf area index (LAI) or the fraction of absorbed photosynthetically active radiation (fAPAR) and bi-directional surface reflectance, and subsequently the estimation of these variables from satellite measurements. The Simplified Level 2 Prototype Processor (SL2P) allows global LAI and fAPAR mapping at 20 m resolution using Sentinel 2 imagery. Previous validation studies over forests indicate SL2P underestimates LAI by up to 50% in comparison to in-situ reference measurements. Our study tests the hypothesis that the SL2P LAI and fAPAR bias can be reduced by replacing the spatially homogenous SAILH canopy radiative transfer model used to calibrate SL2P with the heterogeneous 4SAIL2 model, together with a shoot clumping parameterization. We also hypothesized that the additional parameters involved in this new version of SL2P (SL2P-CCRS) would lead to an increase in precision error and subsequently a bias-variance trade off.

SL2P-CCRS reduced LAI bias by 65%, in comparison to SL2P, during direct validation with 1107 in-situ measurements. The LAI absolute bias reduced by ~0.5 at LAI 3 and by ~1 at LAI 6. SL2P-CCRS reduced fAPAR bias by 31% compared to SL2P but <0.05 on an absolute basis. Bias reduction was accompanied by an increase in precision error so that overall uncertainty, quantified by the root mean square difference in comparison to in-situ measurements, reduced by only 6% for LAI and 5% for fAPAR. These findings support the hypothesis that updating SL2P with a spatially heterogeneous RTM can reduce LAI and fAPAR bias over forests. The results also support the hypothesis that there is a bias-variance trade-off for LAI, and to a lesser extent for fAPAR, when increasing the complexity of SL2P by using a radiative transfer model that accounts for spatial heterogeneity. Nevertheless, SL2P-CCRS increased the agreement rate with Global Climate Observing System uncertainty requirements from 52% to 58% for LAI and 32% to 40% for fAPAR, suggesting that the trade-off is worthwhile, and that algorithms such as SL2P-CCRS, that use a spatially heterogeneous radiative transfer model, should be applied for mapping fAPAR and LAI from Sentinel-2 measurements.

### 1. Introduction

Leaf area index (LAI), defined as half the total foliage surface area per unit horizontal ground area, is an essential climate variable required by regional and global land surface models as well as for local adaptation and agricultural monitoring applications (WMO, 2022; Table 1). There are many global LAI products from low resolution ( $\geq 250$  m) satellite imagers (Fang et al., 2019). However, these products do not satisfy the 100 m Global Climate Observing System (GCOS) baseline spatial resolution requirement. Algorithms suitable for 10 m to 100 m resolution global LAI mapping have been developed for multi-spectral satellite

imagers (Ganguly et al., 2012; Li et al., 2015; Verrelst et al., 2015; Kang et al., 2021; Weiss and Baret, 2016, hereafter WB2016; Weiss and Baret, 2020, hereafter WB2020; Estevez et al., 2022; Laura Martínez-Ferrer et al., 2022). The Simplified Level 2 Prototype Processor (SL2P) (WB2016) is currently used for systematic LAI mapping over North America at 20 m resolution using data from Sentinel 2 (S2) satellite imagers (<https://datacube.services.geo.ca/en/viewer/eo4ce/vegetation/lai.html> accessed on January 23, 2023) and to produce reference LAI and fAPAR maps for validating coarser resolution products (Brown et al., 2020). SL2P uses neural networks, calibrated with canopy parameters and S2 surface reflectance ( $\rho$ ) simulated by the SAILH

\* Corresponding author.

E-mail address: [richard.fernandes@canada.ca](mailto:richard.fernandes@canada.ca) (R. Fernandes).

spatially homogenous radiative transfer model (RTM) (Verhoef, 1985), to estimate LAI and fAPAR given  $\rho$  and acquisition geometry. Validation studies indicate SL2P S2 LAI satisfies GCOS uncertainty requirements over ~92% of non-forested sites but only ~50% of forests sites, ranging from 20% for mixed forests to 85% for evergreen broadleaf forests (Brown et al., 2019; Djamai et al., 2019; Hu et al., 2020; Brown et al., 2021a; Fernandes et al., 2023). The increased SL2P uncertainty in forests corresponds to a LAI underestimation of ~30% for LAI > 2, in comparison to both in-situ reference measurements (RM) and coarse resolution products of known uncertainty (Hu et al., 2020; Brown et al., 2021a; Fernandes et al., 2023). Empirical bias corrections (Fernandes et al., 2023) may address this underestimation but are not guaranteed to be locally representative. Moreover, empirical bias correction will not necessarily preserve the consistency between SL2P estimates of LAI and fAPAR provided by RTM constraints. Rather, SL2P forest validation studies (Hu et al., 2020; Brown et al., 2021a; Fernandes et al., 2023) suggest that SL2P should be updated with a spatially heterogeneous RTM.

In fact, SL2P LAI underestimation over forests could be due to either the RTM used or to incorrect priors for the parameters used to produce calibration data with the RTM. It is unlikely that the SL2P LAI bias is due to the incorrect priors since Brown et al. (2021a) showed that LAI underestimation persists even when using a uniform LAI prior with range 0 to 15. Rather, Brown et al. (2019) showed that a spatially heterogeneous RTM reduced plant area index (PAI, defined as half the total surface area of all vegetation per unit horizontal ground area) underestimation, in comparison to SL2P, from -1.38 to -0.13 at a temperate deciduous broadleaf site. Shabanov et al. (2005) hypothesized that using a RTM that includes spatial heterogeneity caused by crown clumping (Shabanov and Gastellu-Etchegorry, 2018) would reduce the bias of the MODIS MOD15 LAI and fAPAR algorithm over forests. Indeed, the most recent collection 6 MOD15 LAI and fAPAR products are generally unbiased with respect to PAI over deciduous broadleaf and evergreen needleleaf forests (Fang et al., 2019; Brown et al., 2020). However, the 500 m resolution MOD15 LAI products pixels include heterogeneity in canopy architecture and optical properties of soil and vegetation at scales coarser than the 20 m resolution SL2P products. As such, validation results for MOD15 products are not directly applicable at higher spatial resolution. It remains to be seen if using a heterogeneous RTM will reduce SL2P LAI bias when validated over a representative sample of forest sites. Additionally, it is possible that LAI precision will decrease when using a heterogeneous RTM versus a homogenous RTM due to the former model's need for additional parameters. Finally, it is unclear how modifying the SL2P RTM to reduce LAI bias will impact the uncertainty of SL2P estimates of the fraction of absorbed photosynthetically active radiation (fAPAR), also required for monitoring vegetation status and trends.

The goal of our study was to quantify the impact of modifying the SL2P RTM to include canopy spatial heterogeneity on the uncertainty, bias, and precision of LAI and fAPAR estimates over a representative sample of North American forests. To do so, we modified the SAILH canopy RTM by: i. using the model of Smolander and Stenberg (2003) to account for the impact of shoot clumping on shoot single scattering albedo, ii. replacing SAILH with 4SAIL2 (Verhoef and Bach, 2007) to account for the impact of crown clumping on canopy bi-directional

reflectance ( $\rho$ ) and, iii. Using land cover specific priors representative of North American forests when generating RTM simulations. These modifications introduced three new parameters into SL2P: crown cover (CC), crown height to width ratio ( $\chi$ ) and the needle-to-shoot area ratio ( $\gamma$ ). We hypothesized that the modified SL2P (hereafter SL2P-CCRS) would reduce LAI bias observed with SL2P over forests, but not to the extent of that observed in Brown et al. (2019) at one site, due to the broader land cover specific priors required for global application. We further hypothesized that the introduction of three new RTM parameters would reduce the precision of both LAI and fAPAR estimates due to the inversion problem having additional degrees of freedom; indicating the presence of a bias-variance trade-off when using the modified SL2P to map North American forest LAI with S2 imagery.

Algorithms using heterogeneous RTMs have been used to map forest fAPAR and LAI using multispectral imagery between 30 m and 500 m (Shabanov et al., 2005; Huang et al., 2011; Ganguly et al., 2012; Gonsamo and Chen, 2014). However, our study is novel in that it quantifies the magnitude of LAI and fAPAR bias reduction resulting by using a heterogeneous RTM versus a homogeneous RTM, such as in SL2P, in a controlled experiment over a large (>1000) temporal and spatial sample of in-situ sites. Of greater novelty is the evidence, based on our extensive validation dataset, that the bias reduction is accompanied by a decrease in precision for LAI, and to a lesser extent for fAPAR. This bias-variance trade-off has not been previously observed, due to the lack of controlled large sample validation studies, and suggests there may be a limit to the thematic uncertainty of single pixel retrievals of LAI over heterogeneous forests using multispectral imagers without local information on canopy structure. Finally, our study provides new validation results for SL2P-CCRS for North American forests to satisfy the requirements of the Committee of Earth Observing System Land Parameter Validation Sub-Group.

## 2. Materials and methods

### 2.1. Materials

#### 2.1.1. Calibration data

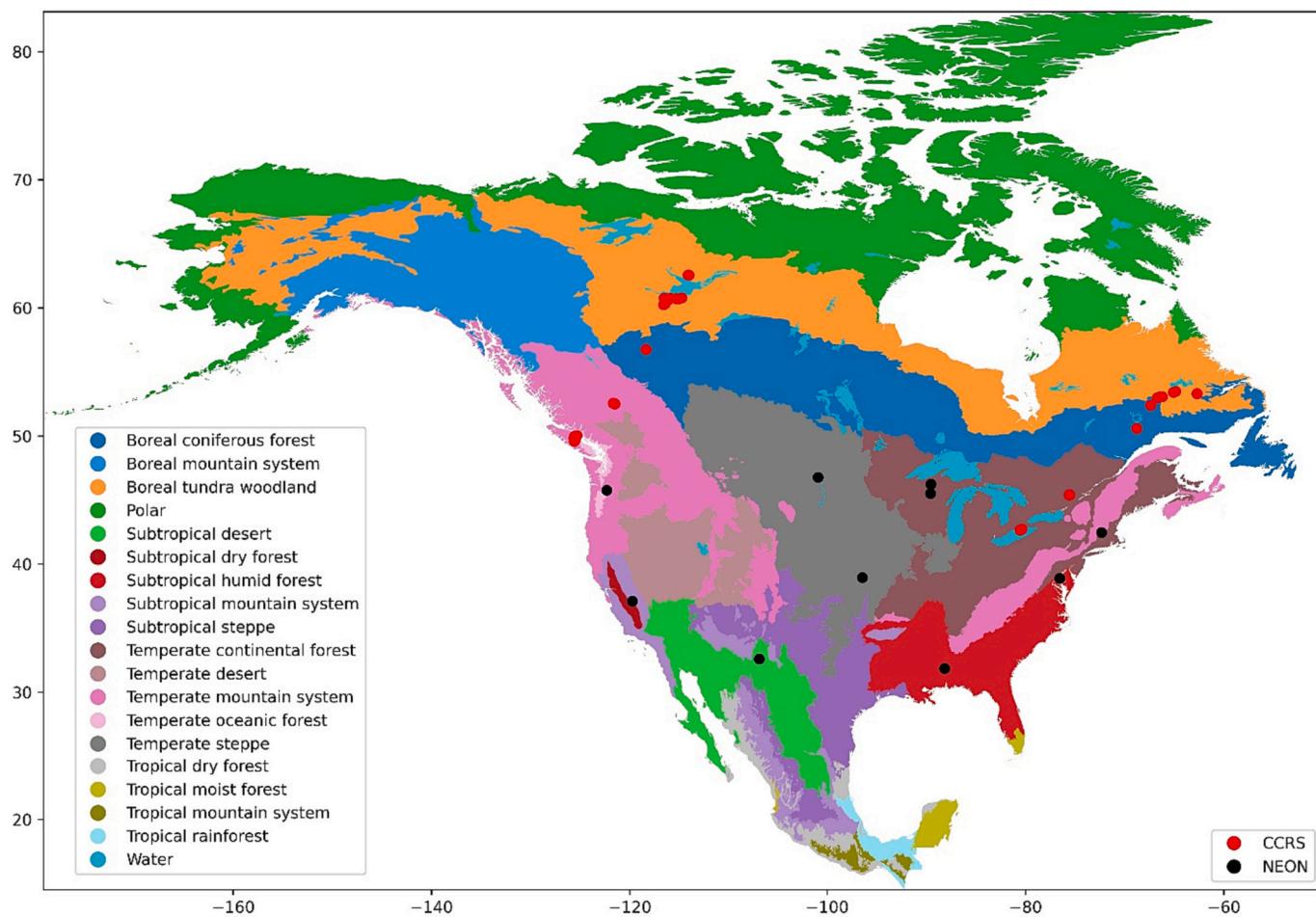
Measurements were acquired from available databases (Appendix B) to calibrate prior probability distributions of parameters used for RTM simulations. The measurements were also qualitatively assessed for representativeness based on having at least 50% of the species sampled occurring in North American forested ecozones (Fig. 1).

#### 2.1.2. In-situ reference measurements

The in-situ RM from Fernandes et al. (2023) were used here for validation without modifications and in their entirety. This was done to facilitate comparison with previous SL2P validation and future studies in the event other researchers wished to test alternate solutions. Briefly, the data correspond to 1107 coincident fAPAR and LAI estimates from digital hemispherical photographs (DHP) acquired during the 2019 or 2020 growing seasons at 281–1 ha elementary sampling units (ESU) across 14 of 17 North American forest ecozones (Fig. 1). Sites in Canada were measured once by Canada Centre for Remote Sensing (CCRS) while sites in the United States of America and Mexico were measured bi-weekly by the National Ecological Observatory Network (NEON; Kao

**Table 1**  
LAI and fAPAR goal (G), baseline (B) and threshold (T) GCOS requirements.

Variable	Definition	Horizontal Resolution			Temporal Resolution			Uncertainty		
		m			d			(%abs), $1\sigma$		
		G	B	T	G	B	T	G	B	T
LAI	Half total green foliage area per horizontal ground area	10	100	250	1	–	10	10%, 0.5	20%, 0.5	
fAPAR	Part of PAR that is effectively absorbed by plants	10	–	250	1	–	10	5%, 0.025	–	10%, 0.05



**Fig. 1.** Location of in-situ reference measurements sites within North American forest ecozones (Commission for Environmental Cooperation, 2022).

et al., 2012). For each RM, between 12 and 14 understory and overstory DHPs were acquired and processed to estimate fraction of intercepted PAR (fIPAR), fraction of plant cover (fPCOVER) and PAI using HemiPy (Brown et al., 2023) for NEON or CANEYEV4.65 (<https://www6.paca.inrae.fr/can-eye/Download/> accessed on September 1, 2023) for CCRS. Estimates of PAI include corrections for canopy clumping as described in Fernandes et al., 2023. fIPAR, fPCOVER and PAI were converted to fAPAR, fCOVER and LAI, respectively, using land cover specific factors (Brown et al., 2021a). The uncertainty of measurement errors and conversion factors was propagated into RM uncertainty following the Fiducial Reference Measurement for Vegetation guidelines (Brown et al., 2021b). RM were geolocated within ~1 m for NEON and ~ 5 m for CCRS at 90% circular error probable.

### 2.1.3. Sentinel 2 input data

Cloud free Copernicus Sentinel S2 Level 2 A Collection 1 surface reflectance products, generated by the European Space Agency using the Sen2Cor algorithm (Müller-Wilm, 2018) and archived within Google Earth Engine (GEE, [https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS\\_S2\\_SR](https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR) accessed on January 23, 2024), were extracted within a +/-7d interval for 3 × 3 20 m S2 pixels centered on the nominal centre of each RM ESU.

## 2.2. Methods

### 2.2.1. SL2P and SL2P-CCRS LAI and fAPAR estimation

For both SL2P and SL2P-CCRS, LAI and fAPAR were estimated using separate non-linear regressions, with regressors corresponding to S2 p and acquisition geometry (Table 2). For a given output variable, SL2P

**Table 2**

SL2P and SL2P-CCRS inputs together with additive and multiplicative noise given in WB2016.

Input	Centre Wavelength nm	Additive Noise	Multiplicative Noise
		%refl	%
B03	559.8	2	4
B04	664.6	2	4
B05	704.1	2	4
B06	740.5	2	4
B07	782.8	2	4
B8A	832.8	2	4
B11	864.7	2	4
B12	945.1	2	4
Solar Zenith Angle (SZA)	–	0	0
View Zenith Angle (VZA)	–	0	0
Relative Azimuth Angle (RAA)	–	0	0

uses a single regression irrespective of land cover, while SL2P-CCRS uses land cover specific regressions. Each regression corresponded to a fully connected single hidden layer neural network, with 5 tangent-sigmoid hidden nodes and a linear output node, calibrated with a database of RTM LAI or fAPAR and corresponding S2 input produced by convolving the 1 nm resolution  $\rho$  simulated by the RTM with Sentinel 2A (S2A) spectral response functions ([https://sentinels.copernicus.eu/documents/247904/685211/S2-SRF\\_COPE-GSEG-EOPG-TN-15-0007\\_3.1.xlsx](https://sentinels.copernicus.eu/documents/247904/685211/S2-SRF_COPE-GSEG-EOPG-TN-15-0007_3.1.xlsx)

accessed on January 23, 2023). Hereafter DBF refers to the SL2P-CCRS FAPAR and LAI regression predictors for deciduous broadleaf forests and woody land cover and ENF refer to the SL2P-CCRS FAPAR and LAI regression predictors for evergreen needleleaf forests.

The ENF and DBF versions of SL2P-CCRS were applied to ESUs with only needleleaf or broadleaf land cover, respectively. The dominant forest type within the mixed forest class for the RM can range from 50% to 75% cover (Fernandes et al., 2023). 4SAIL2 cannot represent mixed forest canopies in a spatially explicit manner. More complex RTMs can simulate mixed forests given local information as to the location and type of mixed vegetation elements (Widlowski et al., 2007) but representative information is generally not available either a priori or during algorithm application for entire biomes. Here, for mixed-forest ESUs, the SL2P-CCRS estimate corresponded to the average of the ENF network and SL2P-CCRS DBF network estimates due to the lack of information to constrain local priors.

SL2P uses the PROSPECT-D leaf optical model (Féret et al., 2017) coupled with SAILH to simulate  $\rho$  at 1 nm intervals between 400 nm and 2500 nm. In contrast, SL2P-CCRS uses PROSPECT-D coupled with 4SAIL2, after modifying the latter to account for foliage clumping within shoots. 4SAIL2 differs from SAILH by modelling uncollided fluxes and single scattering using geometric optics, assuming foliage occurs in identical randomly located ellipsoidal crowns, defined by the crown cover (CC) and crown height to width ratio ( $\chi$ ), but retains the SAILH

$$x_p(LAI) = \max\left(x_{lb}, \min\left\{x_{ub}, \frac{x_{max}(LAI_{ub}) - x_{min}(LAI_{ub})}{x_{max}(LAI_{ub}) - x_{min}(LAI_{lb})} [x_c(LAI) - x_{min}(LAI_{ub})] + x_{min}(LAI_{ub})\right\}\right) \quad (2)$$

multiple scattering model. As a result, for heterogeneous canopies, 4SAIL2 canopy transmittance and the canopy hot spot angular width will increase in comparison to SAILH for the same input parameters. 4SAIL2 also allows for a non-photosynthetic vegetation (NPV) layer between the soil background and the rest of the canopy. This layer was not used here as the SL2P-CCRS soil spectral inputs accounted for NPV understory vegetation. To account for shoot clumping, leaf reflectance and transmittance from PROSPECT-D was scaled by ratio ( $r$ ) of leaf single scattering albedo to the shoot single scattering albedo ( $\alpha_s$ ) as a function of  $\gamma$  (Smolander and Stenberg, 2003):

$$r = 1 - \frac{1}{\gamma} \quad (1)$$

LAI was also scaled by  $\frac{1}{\gamma}$  before use in 4SAIL2 to model the impact of shoot clumping on the foliage interaction cross-section.

SL2P RTM simulations were produced for a single globally applicable calibration database. SL2P-CCRS RTM simulations were produced for

**Table 3**

SL2P (regular font only) and SL2P-CCRS (both regular font and bold font) RTM canopy parameters. Dimensionless units indicated by a dash.

Symbol	Name	Units
LAI	Leaf Area Index	–
CC	<b>Crown Cover</b>	–
ALA	Average leaf angle	° from horizontal
HsD	Foliage hot spot parameter	–
$\chi$	<b>Crown height to width ratio</b>	–
N	PROSPECT N parameter	–
Cab	Foliage chlorophyll A + B content	$\mu\text{g.m}^{-2}$
Cdm	Foliage dry matter content	$\text{g.m}^{-2}$
Cw_rel	Foliage relative water content	–
Cbp	Foliage brown pigment content	–
$\gamma$	<b>Needle to shoot area ratio</b>	–
Bs	Soil brightness factor	–

separate calibration databases corresponding to North American evergreen needleleaf forest (for ENF) and deciduous broadleaf forests and woody land cover (for DBF) international geosphere biosphere (IGBP) land cover classes (Lambin and Geist, 2006). Each RTM simulation required inputs for acquisition geometry (SZA, VZA and RAA) and 9 (for SL2P) or 12 (for SL2P-CCRS) canopy parameters (Table 3). Acquisition geometry was uniformly sampled from the nominal S2A orbital geometry between  $-56^{\circ}\text{S}$  and  $83^{\circ}\text{N}$  latitudes as in WB2016. Canopy parameters were sampled from the joint distribution of input parameters using orthogonal sampling for SL2P (WB2016) and Sobol sampling as implemented by the MATLAB sobolset function (Mathworks, <https://www.mathworks.com/help/stats/sobolset.html#References> accessed on January 23, 2023), as described in Brown et al., 2021a, for SL2P-CCRS. Sobol sampling was used in preference to orthogonal sampling since it offers similar coverage of the prior distribution with sample size increments constrained to powers of 2 rather than the dimensionality of the sampled space while having negligible impact on the cross-validation thematic uncertainty of the calibrated algorithm (Appendix C).

As in WB2016, the joint prior distribution of canopy parameters was specified using conditional distributions of each parameter ( $x$ ) given LAI ( $x_p(LAI)$ ). Each  $x_p(LAI)$  was defined by linear scaling of the conditional distribution  $x$  given LAI assuming independent priors ( $x_c(LAI)$ ) (Table 4 for SL2P, Table 5 for the DBF, and Table 6 for ENF):

where  $x_{lb}$  and  $x_{ub}$  are the lower and upper bounds for the parameter and  $x_{min}(LAI)$  and  $x_{max}(LAI)$  specifies the linear scaling range for  $x$  at a given LAI. Scaling functions for CC were calibrated using data from Fernandes et al. (2023) (Appendix B) while  $\chi$  and  $\gamma$  were assumed independent of LAI.

For SL2P, WB2016 calibrated the univariate prior distributions required for  $x_c(LAI)$  using qualitatively selected probability density functions fit to measurements, available before 2016, irrespective of land cover class (Table 7). For SL2P-CCRS, recognizing that a fit to all available measurements may not be representative of regional land cover conditions, a less qualitative strategy was used resulting in land cover specific priors (Table 8 and Table 9):

- i. For parameters with  $n > 100$  representative samples, an objective distribution fitter (<https://github.com/cokelaer/fitter> accessed on August 21, 2023) was used to select the distribution type that minimizes the Kullback-Leiber Divergence and then the MATLAB fitdist function (<https://www.mathworks.com/help/stats/fitdist.html> accessed on August 21, 2023) was used to fit a truncated version of this distribution.
- ii. For  $20 \leq n \leq 100$  representative samples, a truncated normal distribution was fit using fitdist. The normal distribution was chosen as it is the maximum entropy distribution given the population mean and standard deviation. The standard deviation of the fitted distribution was then doubled to reflect the likelihood that samples were not diverse.
- iii. For  $n < 20$  samples or for non-representative sampling, a uniform distribution was fit to the range of the data. The uniform distribution was chosen since it is the maximum entropy distribution given only the population range. The standard deviation was not further inflated since there were insufficient samples to allow for

**Table 4**

SL2P RTM parameter bounds and scaling ranges.

Symbol	$x_{lb}$	$x_{ub}$	$x_{min}(LAI_{lb})$	$x_{max}(LAI_{lb})$	$x_{min}(LAI_{ub})$	$x_{max}(LAI_{ub})$
LAI	0	15	–	–	–	–
ALA	30	80	30	80	55	65
HsD	0.1	0.5	0.1	0.5	0.1	0.5
N	1.2	2.2	1.2	2.2	1.3	1.8
Cab	20	90	20	90	45	90
Cdm	0.003	0.011	0.003	0.011	0.005	0.011
Cw_rel	0.6	0.85	0.6	0.85	0.7	0.80
Cbp	0	2.0	0	2.0	0	0.20
Bs	0.5	3.5	0.5	3.5	0.5	1.2

**Table 5**

SL2P-CCRS RTM parameter bounds and scaling ranges for DBF class.

Symbol	$x_{lb}$	$x_{ub}$	$x_{min}(LAI_{lb})$	$x_{max}(LAI_{lb})$	$x_{min}(LAI_{ub})$	$x_{max}(LAI_{ub})$
LAI	0	10	–	–	–	–
CC	0.01	0.95	0.01	0.7	1.35	1.35
ALA	20	60	20	60	20	60
HsD	0.1	0.5	0.1	0.5	0.1	0.5
$\chi$	1	4	1	4	1	4
N	1.1	2.2	1.1	2.2	1.1	2.2
Cab	15	65	15	65	45	60
Cdm	0.005	0.01	0.005	0.01	0.005	0.01
Cw_rel	0.5	0.9	0.5	0.9	0.75	0.90
Cbp	0	0.2	0	0.2	0	0.2
$\gamma$	1	1	1	1	1	1
Bs	0.5	3.5	0.5	3.5	0.5	1.2

**Table 6**

SL2P-CCRS RTM parameter bounds and scaling ranges for ENF class.

Symbol	$x_{lb}$	$x_{ub}$	$x_{min}(LAI_{lb})$	$x_{max}(LAI_{lb})$	$x_{min}(LAI_{ub})$	$x_{max}(LAI_{ub})$
LAI	0	10	–	–	–	–
CC	0.01	0.95	0.01	0.52	1.2	1.35
ALA	20	70	20	70	20	70
HsD	0.1	0.5	0.1	0.5	0.1	0.5
$\chi$	1	4	1	4	1	4
N	1.1	2.2	1.1	2.2	1.1	2.2
Cab	15	65	15	65	45	60
Cdm	0.005	0.01	0.005	0.01	0.005	0.01
Cw_rel	0.7	0.9	0.7	0.9	0.75	0.90
Cbp	0	0.2	0	0.2	0	0.2
$\gamma$	1	2	1	2	1	2
Bs	0.5	3.5	0.5	3.5	0.5	1.2

**Table 7**

SL2P RTM parameter univariate prior distributions.

Symbol	Distribution	$x_{lb}$	$x_{ub}$	$\mu$	$\sigma$
LAI	Normal	0	15	2.0	3.0
ALA	Normal	30	80	60	30
HsD	Normal	0.1	0.5	0.2	0.5
N	Normal	1.2	2.2	1.5	0.30
Cab	Normal	20	90	45	30
Cdm	Normal	0.003	0.011	0.005	0.005
Cw_rel	Uniform	0.6	0.85	0.72	0.07
Cbp	Normal	0	2	0	0.3
Bs	Uniform	0.5	3.5	1.2	2.0

a precise estimate of the standard deviation of the sampling distribution underlying the available samples.

SL2P includes a library of 7 soil spectral shapes from which the soil reflectance shape was randomly sampled and then scaled by Bs. This library is also used here for SL2P-CCRS DBF as Fernandes et al. (2023) showed that SL2P estimates of DBF LAI were relatively unbiased for low (<1) LAI. For ENF, this library was supplemented with understory measurements of lichen, moss and litter covered Boreal Forest soils (Miller et al., 1997) (Fig. 2).

**Table 8**

SL2P-CCRS RTM parameter univariate prior distributions for DBF. G.E.V. corresponds to Generalized Extreme Value.

Symbol	Distribution	$x_{lb}$	$x_{ub}$	$\mu$	$\sigma$
LAI	G.E.V. shape = -0.05	0	10	3.2	2.0
CC	Uniform	0.01	0.95	0.46	0.26
ALA	Normal	20	70	45	22
HsD	Normal	0.1	0.5	0.2	0.5
$\chi$	Uniform	1	4	2.5	0.75
N	Uniform	1.1	2.3	1.7	0.12
Cab	Uniform	15	75	40	208
Cdm	Beta	0.005	0.1	0.0074	0.0058
Cw_rel	Beta	0.5	0.9	0.705	0.105
Cbp	Normal	0	0.2	0	0.3
$\gamma$	Uniform	1	1	1	0
Bs	Lognormal	0.5	3.5	0.9	0.6

Calibration databases were produced by applying additive and multiplicative noise (Table 2) to simulated S2  $\rho$ . A database of 41,472 samples was used for SL2P following WB2016. For both ENF and DBF databases, a sample size of 262,144 was selected, corresponding to the size at which the cross-validation root mean square error of the calibrated neural networks for both LAI and fAPAR changed by <1% from the next possible smaller sample size (131,072 samples in this case).

**Table 9**

SL2P-CCRS RTM parameter univariate prior distributions for ENF. G.E.V. corresponds to Generalized Extreme Value.

Symbol	Distribution	$x_{lb}$	$x_{ub}$	$\mu$	$\sigma$
LAI	G.E.V. shape = -0.05	0	10	3.5	2.5
CC	Uniform	0.01	0.95	0.46	0.26
ALA	Normal	20	70	45	22
HsD	Normal	0.1	0.5	0.2	0.5
$\chi$	Uniform	1	4	2.5	0.866
N	Uniform	1.1	2.3	1.7	0.173
Cab	Uniform	15	75	45	17.3
Cdm	Beta	0.005	0.1	0.02	0.0011
Cw_rel	Beta	0.5	0.9	0.67	0.1158
Cbp	Normal	0	0.2	0	0.3
$\gamma$	Uniform	1	2	1.5	0.25
Bs	Lognormal	0.5	3.5	0.9	0.866

For both SL2P and SL2P-CCRS, the calibration database was partitioned into five batches by k-NN clustering of the RTM input parameters. For each of LAI and fAPAR, neural networks were calibrated using batch training for a maximum of 250 epochs, with one batch held out for cross-validation during each epoch. For each epoch, network coefficients were adjusted to minimize the mean square error across the four retained batches using the scaled conjugate gradient algorithm with the MATLAB fitnet package ([https://www.mathworks.com/help/deeplearning/ref/fi\\_tnet.html](https://www.mathworks.com/help/deeplearning/ref/fi_tnet.html) accessed on January 9, 2023). Training was halted if the cross-validation mean square error failed to decrease for six consecutive epochs. The calibrated networks coefficients were uploaded to Google Earth Engine (GEE) and applied on a per-pixel basis to clear sky S2 imagery using the Landscape Evolution and Forecasting (LEAF) Toolbox GEE application (Fernandes et al., 2021).

To quantify the impact of changing priors other than CC,  $\chi$ , and  $\gamma$ , modified versions of SL2P using the SAILH with SL2P-CCRS land cover specific priors were calibrated for the ENF and DBF classes (hereafter SL2P-LandCover, Appendix C). These versions were also validated to quantify the impact of including canopy heterogeneity as in SL2P-CCRS

versus simply updating SL2P priors unrelated to canopy spatial structure as in SL2P-LandCover.

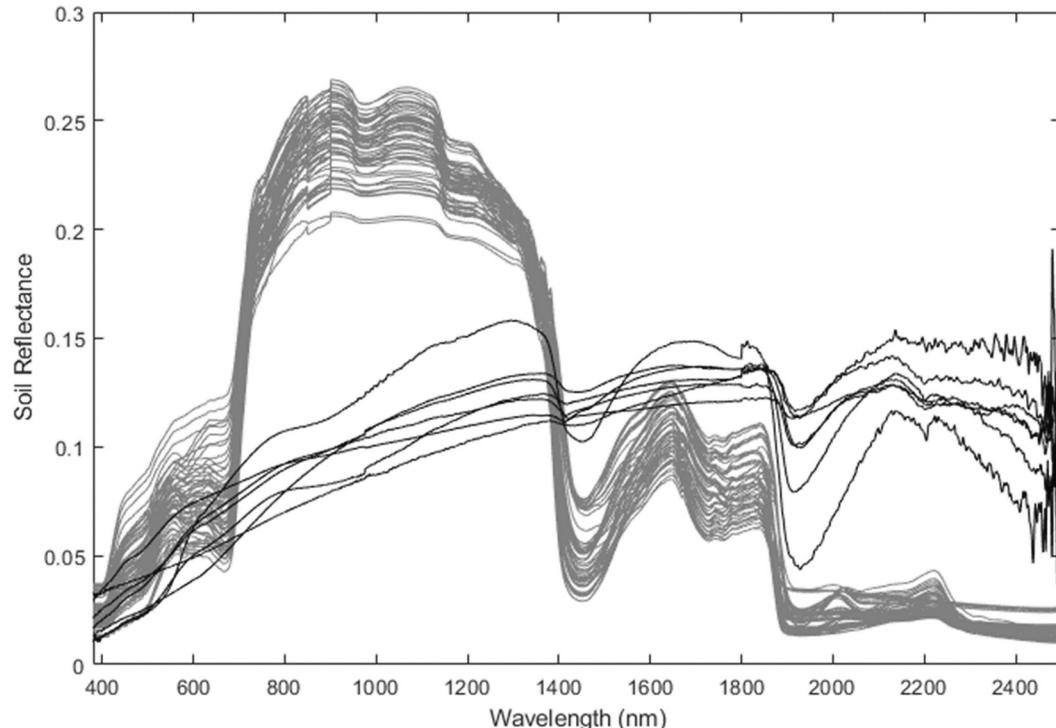
### 2.2.2. Thematic performance assessment

Thematic performance was quantified using direct validation by comparison to in-situ RM from Fernandes et al. (2023) and cross-validation with RTM simulations.

Direct validation was performed by uploading the SL2P, SL2P-CCRS ENF and SL2P-CCRS DBF neural networks to GEE and applying them to L2A products from Sentinel 2A and 2B using the LEAF-Toolbox (Fernandes et al., 2021) to estimate LAI and fAPAR at validation sites. SL2P was applied irrespective of land cover. SL2P-CCRS was applied using the in-situ land cover for the ESU, with the average ENF and DBF estimates used for mixed forest ESUs. As in Fernandes et al. (2023), for each network, residuals between RM and estimates within the  $+/-7d$  window and  $3 \times 3$  S2 pixels, centred on the in-situ RM date and location, were quantified for LAI and fAPAR. To reduce the impact of outliers due to geolocation or atmospheric correction errors, residuals corresponding to dates where the mean absolute residual exceeded the mean absolute residual for all dates were discarded.

Cross-validation was performed by applying the neural networks for each class to a set of independent SL2P-CCRS RTM simulations for ENF and DBF land cover identical in sample size and priors to those used to calibrate the SL2P-CCRS neural networks. SL2P was also applied to both these cross-validation databases to quantify its theoretical uncertainty when applied to heterogeneous canopies. Only RTM simulations with non-negative fAPAR and LAI retrievals for both SL2P and SL2P-CCRS were used for when computing performance metrics for cross-validation to ensure identical validation samples.

Following good practice guidelines (Fernandes et al., 2014; Fernandes et al., 2023) population accuracy (A), precision (P), uncertainty (U) and user agreement ratio (UAR) metrics (Table 10) were reported as well as A,P and U statistics conditional on the reference value. Conditional statistics were estimated by fitting third order robust polynomial regression models using statsmodel version 01.3 (<https://www.statsmodels.org/stable/>)



**Fig. 2.** Soil spectra used for all algorithms (black) and additional soil spectra for ENF algorithm (grey).

**Table 10**  
Thematic performance metrics.

Metric	Acronym	Definition
Uncertainty	U	Square root of the expected value of the squared difference of estimated and product values.
Accuracy	A	Expected value of the estimated value minus the product value.
Precision	P	Square root of the expected value of the square of the total of the estimated value minus both the product value and the accuracy metric.
Uncertainty agreement ratio	UAR	The fraction of validated samples that meet a given uncertainty requirement, in this case, GCOS.
Coefficient of determination	$r^2$	Coefficient of determination.
Slope intercept	–	Slope of ordinary least squares regression line.
	–	Intercept of ordinary least squares regression line.

[smmodels.org/stable/index.html](http://smmodels.org/stable/index.html) accessed January 23, 2024) to predict residuals for a given reference value. The coefficient of determination, slope and intercept of an ordinary least squares regression between estimated and RM values was also included. Code for deriving these metrics is available at [github/rfernand397/LEAF-Toolbox.Papers/SL2PCCRSValidation](https://github/rfernand397/LEAF-Toolbox.Papers/SL2PCCRSValidation) (accessed January 23, 2024).

### 3. Results

#### 3.1. Calibration

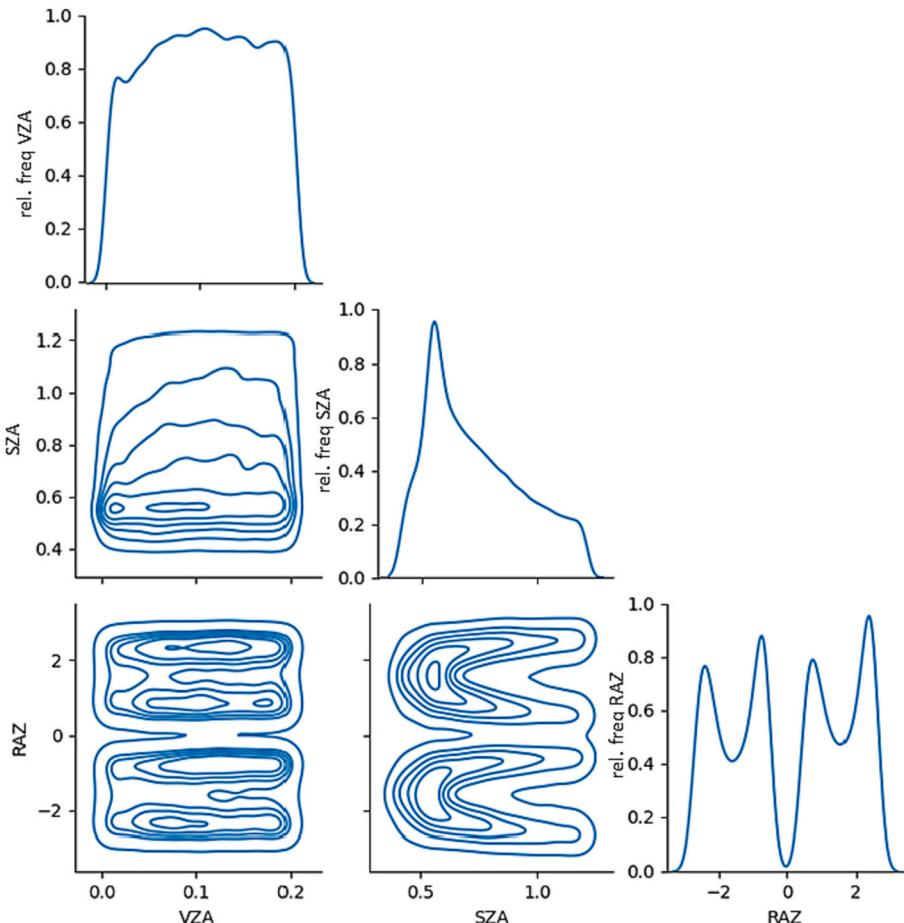
Acquisition geometry posteriors were identical for SL2P and SL2P-CCRS (Fig. 3). VZA was uniform within the nominal S2 swath. SZA showed a mode at  $\sim 0.57$  rad corresponding to the 10:30 am local overpass time for Equatorial regions during the growing season. RAA showed two double modes centred at a  $\sim -2$  rad and  $\sim 2$  rad, corresponding to east and west looking scan angles. The likelihood of acquisitions near the principal plane was low due to the low probability of matching VZA and SZA and the low probability of a RAZ close to zero.

Posterior distributions of RTM parameters differ slightly between ENF and DBF (except for the obvious difference in  $\gamma$ ) but, in many cases, substantially between SL2P-CCRS and SL2P (Fig. 4). SL2P shows greater dispersion than either ENF or DBF for some foliage parameters (Cab, Cdm, Cbp) since it uses global priors that include crops with larger values than forests. Canopy and foliage structure priors such as fAPAR, LAI, CC,  $\gamma$ , ALA and N also show substantial differences in their mode between SL2P and ENF and DBF due to the different underlying distributions.

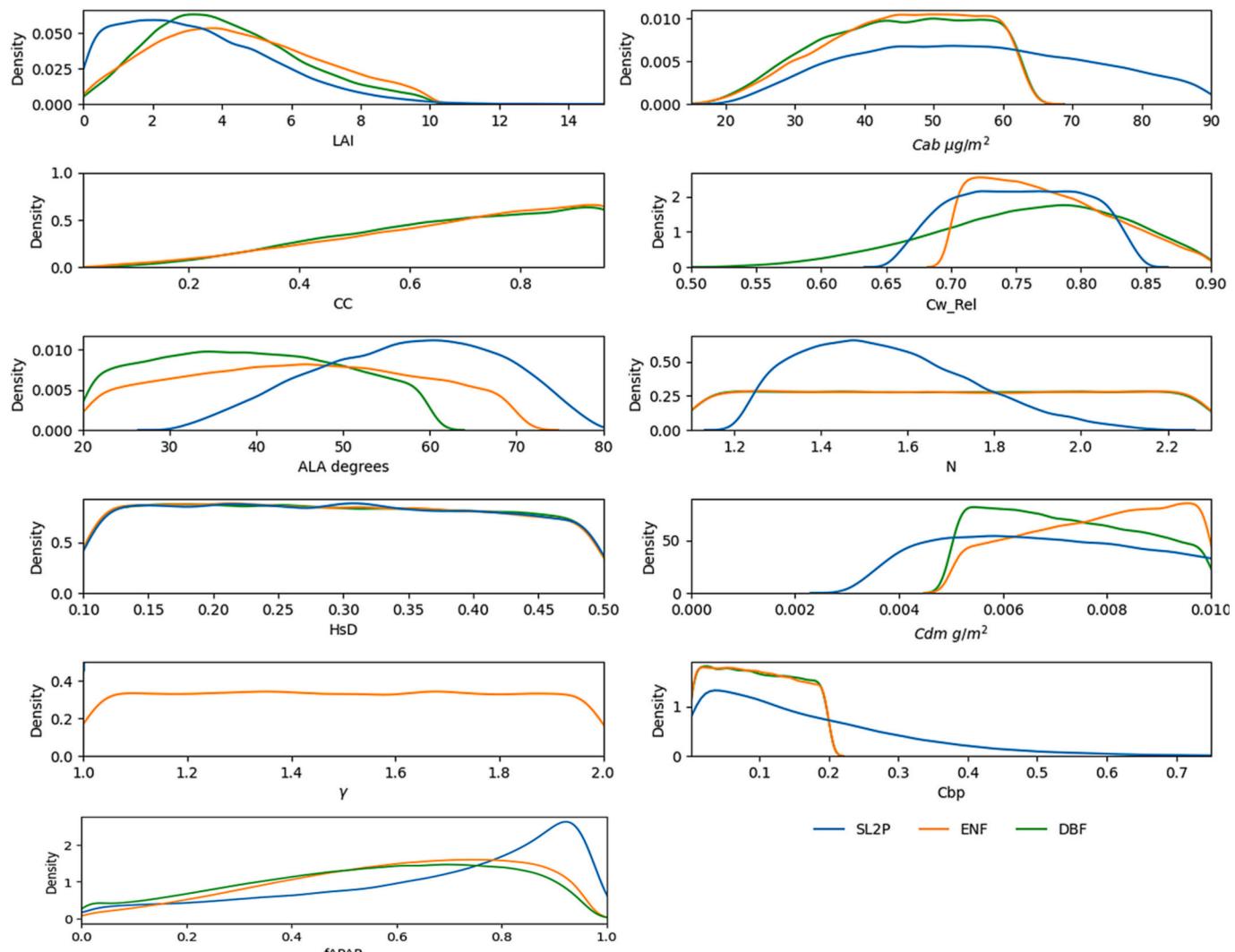
#### 3.2. Thematic performance

##### 3.2.1. Population thematic performance metrics

Validation of SL2P-CCRS LAI against in-situ RM indicated a 65% reduction of accuracy error but only a 6% reduction in uncertainty compared to SL2P (Table 11), resulting in only a modest increase in UAR from 0.51 to 0.58. The improved SL2P-CCRS LAI accuracy compared to



**Fig. 3.** Smoothed histograms (diagonal) and density contour plots for acquisition geometry posterior distributions. Units for variables are radians. Contours range from 0 (outermost) to 1 (innermost) in steps of 0.2. rel. freq. refers to relative frequency of histogram.



**Fig. 4.** Smoothed histograms of posterior distributions of RTM parameters used when calibrating SL2P and DBF and ENF algorithms. DBF and ENF algorithms share the same prior for N. SL2P assumes CC = 1 while SL2P and DBF assume  $\gamma = 1$  (not shown).

**Table 11**  
Population thematic performance metrics for LAI estimation.

Estimate	SL2P			SL2P-CCRS		
Reference	In-Situ	ENF Simulations	DBF Simulations	In-Situ	ENF Simulations	DBF Simulations
<b>Metric</b>						
n	1107	248,832	248,832	1107	248,832	248,832
U	0.910	1.632	1.514	0.855	1.417	1.223
A	-0.375	-0.450	-0.852	-0.132	-0.023	0.01
UAR	0.515	0.362	0.372	0.576	0.445	0.477
r <sup>2</sup>	0.670	0.61	0.67	0.742	0.73	0.66
slope	1.30	1.18	1.00	1.08	1.03	1.03
intercept	-0.41	-0.39	0.87	0.08	-0.48	-0.38

SL2P was also reflected in a slope of 1.08 versus 1.30 and intercept of 0.08 versus -0.41. SL2P-CCRS fAPAR accuracy error was 31% lower than SL2P but uncertainty dropped by only 5% when compared to SL2P (Table 12). Nevertheless, the improved SL2P-CCRS fAPAR accuracy corresponded to an increase in UAR from 0.31 to 0.40 and improvement of slope from 1.07 to 1.00 when compared to SL2P. Validation results for SL2P-LandCover indicated an even greater LAI underestimation than SL2P (-0.91) and decrease in both LAI and fAPAR UAR compared to SL2P, although the bias for fAPAR did reduce by 5% compared to SL2P

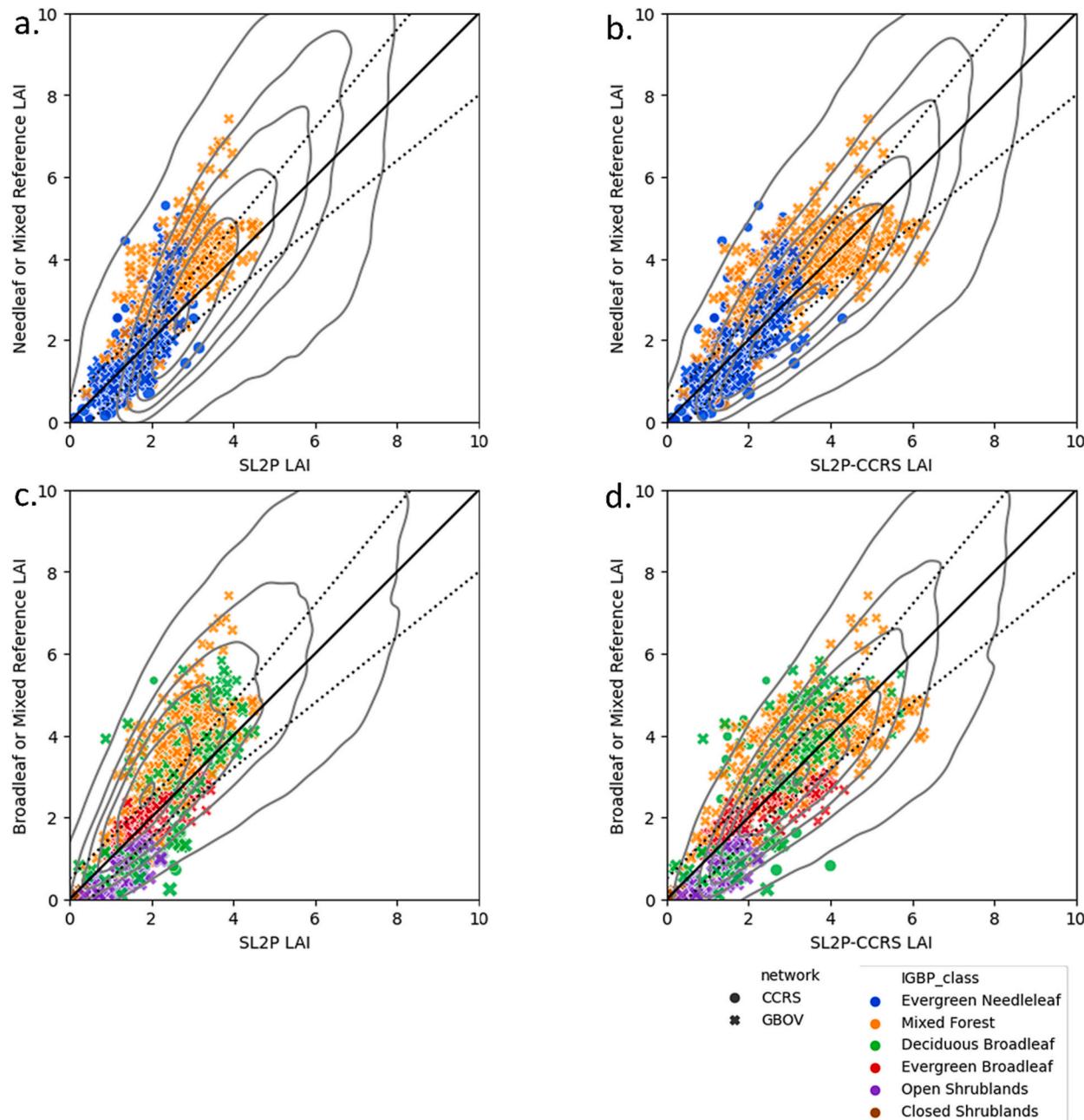
(Appendix D).

The improvement in direct validation accuracy of SL2P-CCRS versus SL2P is encouraging but only applies to conditions representative of the in-situ RM. The representativeness of validation statistics for other forest landscapes in North America was evaluated by comparing the in-situ validation scatter plots with cross-validation density plots using simulations. Fig. 5 (for LAI) and Fig. 6 (for fAPAR) provide scatter plots of SL2P or SL2P-CCRS estimates versus in-situ RM for two land cover groupings (either needleleaf and mixed forests or broadleaf and mixed

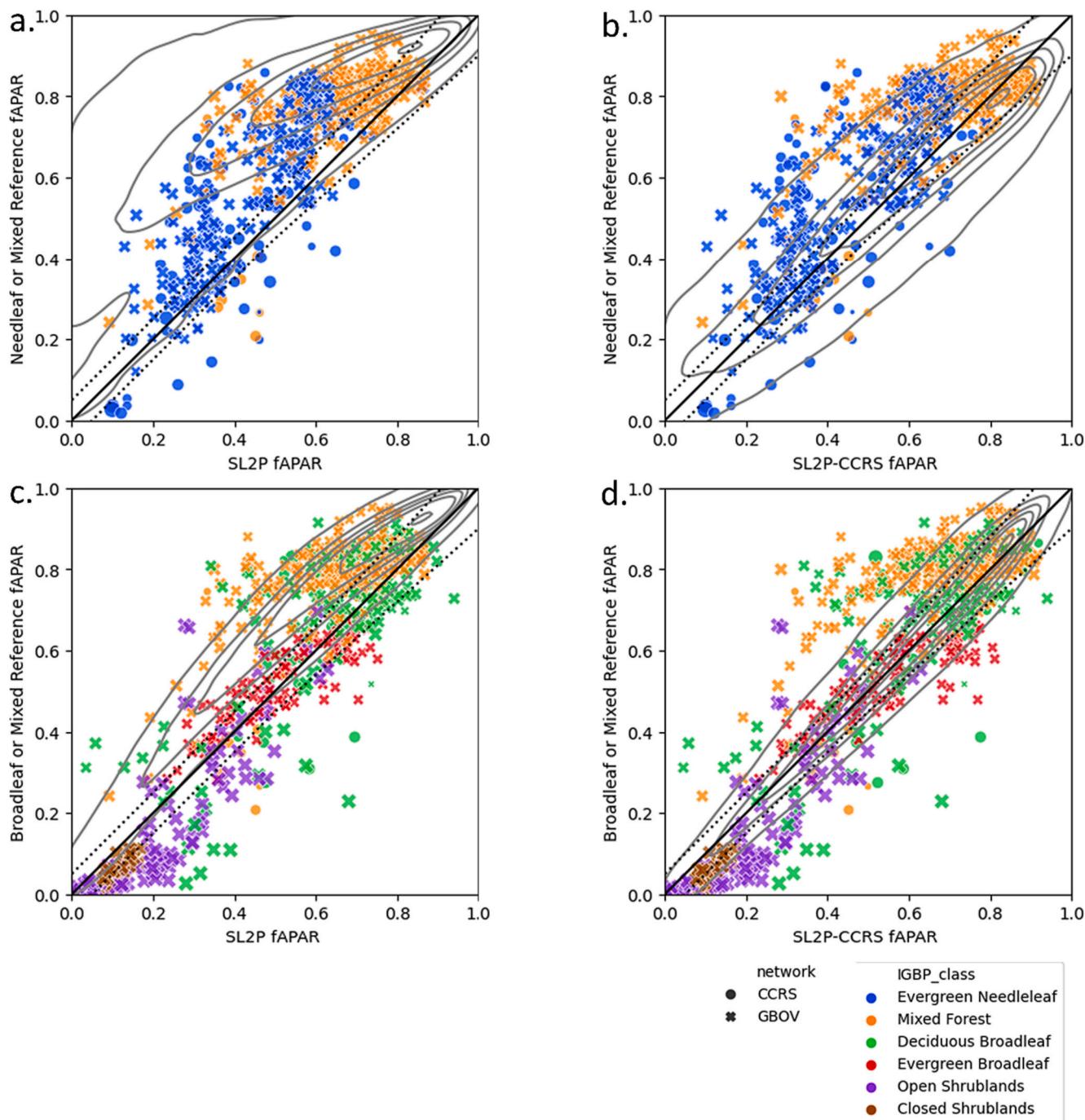
**Table 12**

Population thematic performance metrics for fAPAR estimation.

Estimate	SL2P			SL2P-CCRS			
	Reference	In-Situ	ENF Simulations	DBF Simulations	In-Situ	ENF Simulations	DBF Simulations
Metric							
n		1107	248,832	248,832	1107	248,832	248,832
U		0.145	0.261	0.129	0.136	0.088	0.053
A		-0.065	0.218	0.113	-0.045	0.000	-0.012
UAR		0.313	0.189	0.291	0.398	0.522	0.771
r <sup>2</sup>		0.769	0.71	0.74	0.769	0.77	0.95
slope		1.07	1.46	1.08	1.00	1.03	1.03
intercept		0.07	-0.60	-0.21	0.08	-0.02	0.00



**Fig. 5.** Scatter plots of SL2P (a. and c.) and SL2P-CCRS (b. and d.), LAI estimates versus in-situ RM LAI for needleleaf and mixed land cover (top row) and for broadleaf and mixed land cover (bottom row). Probability density contours correspond to cross-validation of each algorithm using an independent sample of SL2P-CCRS simulations for ENF (upper row) or DBF (lower row) cover. Contours range from 0 (outermost) to 1 (innermost) in steps of 0.2. 1:1 line (solid black) and GCOS user requirements (dotted lines) are included.



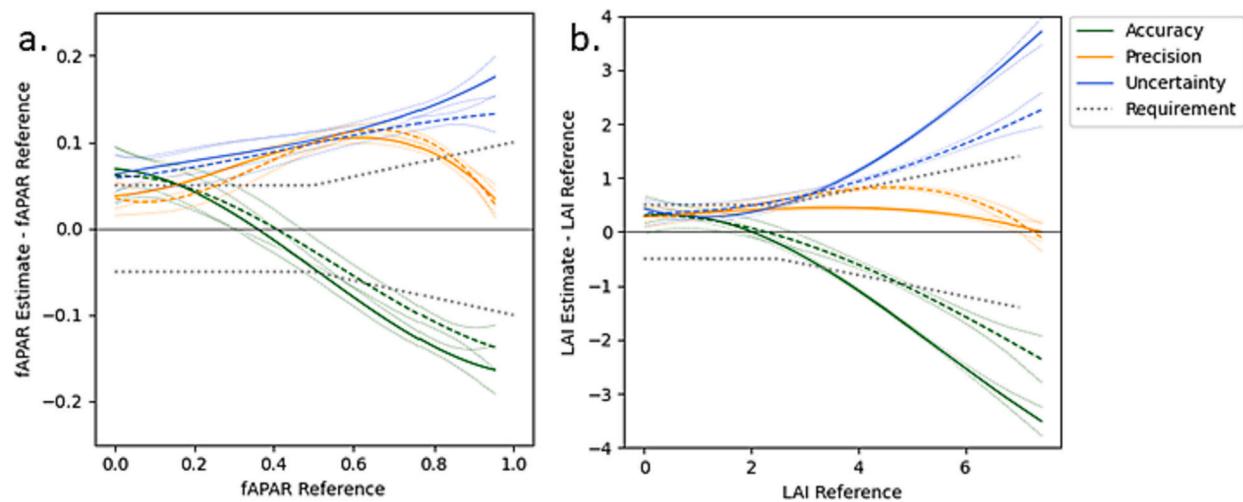
**Fig. 6.** Scatter plots of SL2P (a. and c.) and SL2P-CCRS (b. and d.), fAPAR estimates versus in-situ RM fAPAR for needleleaf and mixed land cover (top row) and for broadleaf and mixed land cover (bottom row). Probability density contours correspond to cross-validation of each algorithm using an independent sample of SL2P-CCRS simulations for ENF (upper row) or DBF (lower row) cover. Contours range from 0 (outermost) to 1 (innermost) in steps of 0.2. 1:1 line (solid black) and GCOS user requirements (dotted lines) are included.

land cover). Superimposed on each scatter plot is a density plot of SL2P or SL2P-CCRS estimates versus reference values from either the ENF or DBF cross-validation databases. Qualitatively, cross-validation density plots captured both the bias for SL2P LAI, and lack thereof for SL2P-CCRS LAI, and the general distribution of in-situ comparisons. The cross-validation and in-situ validation  $r^2$  were within 10% of the  $r^2$  observed during direct validation (Table 11 and Table 12). However, the cross-validation U and UAR was generally worse than for in-situ validation. This is expected as the cross-validation dataset included high LAI values and extreme combinations of foliage optical properties not frequently observed within healthy canopies corresponding to the majority of in-

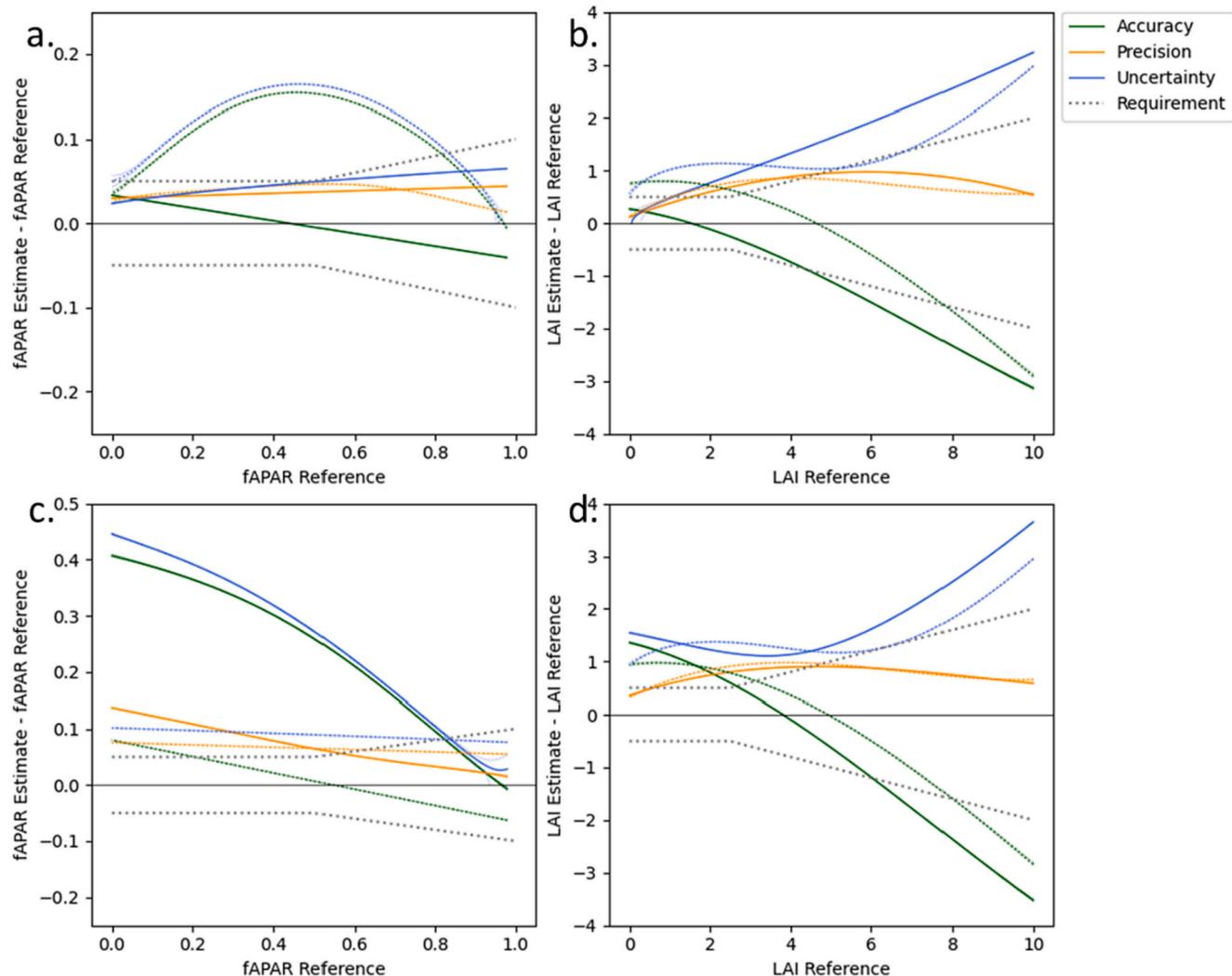
situ ESUs.

### 3.2.2. Conditional metrics

For both fAPAR uncertainty and precision, the  $1\sigma$  confidence intervals for SL2P-CCRS and SL2P conditional models generally overlap, indicating they are not statistically distinguishable given the available RM (Fig. 7a). However, the fAPAR conditional accuracy model only overlaps for fAPAR < 0.65, after which SL2P-CCRS shows an ~10% increase in accuracy (Fig. 7a). For LAI, SL2P-CCRS and SL2P  $1\sigma$  confidence intervals overlap for A, P and U for LAI < 3, but SL2P-CCRS shows an improvement in A and U of between ~10% to ~30% for LAI > 3



**Fig. 7.** Conditional models of direct validation metrics based on third order polynomial fits to residuals between SL2P and in-situ RM (solid-coloured lines) or between SL2P-CCRS and in-situ RM (dashed coloured lines). GCOS user requirements (grey dotted lines) and 1 standard deviation confidence intervals of model fits (coloured dotted lines) are included.



**Fig. 8.** Conditional models of cross-validation metrics based on third order polynomial fits to residuals between SL2P and RTM reference values (solid-coloured lines) or between SL2P-CCRS and RTM reference values (dashed coloured lines) for DBF (a. and b.) and ENF (c. and d.) cross-validation databases. GCOS user requirements (grey dotted lines) and 1 standard deviation confidence intervals of model fits (coloured dotted lines) are included.

(Fig. 7b). However, the  $1\sigma$  lower bound of the SL2P-CCRS LAI precision error also exceeds the  $1\sigma$  upper bound for SL2P for  $3 < \text{LAI} < 6$  (orange curves in Fig. 7b). In terms of GCOS requirements, the net improvement of SL2P-LAI due to improved accuracy results in retrievals meeting GCOS requirements up to  $\sim\text{LAI } 4$  versus  $\sim\text{LAI } 3$  for SL2P; with SL2P-CCRS within 1 LAI unit of the GCOS requirement for  $\text{LAI} < 6$ .

Conditional A, P and U metrics for fAPAR cross-validation are similar in magnitude to those for direct-validation for SL2P-CCRS, but not for SL2P (Fig. 8). SL2P shows very different fAPAR conditional accuracy models for both ENF (Fig. 8a) and DBF (Fig. 8c) cross-validation in comparison to direct validation (Fig. 7), with a large positive bias approaching 0.2 for DBF and 0.4 for ENF. However, SL2P fAPAR precision is similar between cross-validation and direct validation with a magnitude just under 0.1. These results are consistent with the differences in the fAPAR posteriors between SL2P and SL2P-CCRS calibration databases (Fig. 4).

Conditional A, P and U metrics for LAI cross-validation are similar in trend to conditional metrics based on direct validation but are offset in accuracy (Fig. 8b and Fig. 8d). SL2P-CCRS estimates tend to be positively biased during cross-validation for  $\text{LAI} < 3$  rather than being unbiased as observed for direct validation. Additionally, SL2P and SL2P-CCRS show similar LAI precision during cross-validation rather than SL2P having increased precision as observed during direct-validation. While the cross-validation A and P metrics are not representative of direct validation the overall cross-validation uncertainty metrics reflect the potential improvement in uncertainty of SL2P-CCRS versus SL2P for  $\text{LAI} > 3$ .

#### 4. Discussion

Forests exhibit spatial heterogeneity across a range of scales that impacts the relationship between canopy parameters and radiation fluxes quantities such as fAPAR and  $\rho$  (Chen, 1996; Chen and Leblanc, 1997). RTMs can capture this heterogeneity with varying levels of uncertainty if canopy parameters are known (Widlowski et al., 2007; Widlowski et al., 2015). For example, many three-dimensional Monte-Carlo ray tracing models have sufficient accuracy and precision to serve as reference standards for heterogeneous canopies, with uncertainty below that of typical multispectral satellite surface reflectance products (Widlowski et al., 2007). SL2P, as a prototype, used SAILH that provides reasonable accuracy when modelling  $\rho$  for homogenous canopies such as crops and uniform grasslands but not for canopies with crown clumping (Widlowski et al., 2007). Our results indicate that model error, due to using the spatial homogeneous SAILH, is the likely cause of the LAI underestimation of up to 50% previously observed when SL2P was validated over North American forests (Fernandes et al., 2023). Brown et al. (2019) demonstrated at one forest site that the bias could be removed by using a slightly more complicated RTM that used a geometric optics parameterization for uncollided and single scattering. Brown et al. (2020) also showed that the MODIS LAI and fAPAR algorithm, that use heterogeneous RTMs, can provide estimates with low bias, albeit with respect to PAI, over both evergreen needleleaf and deciduous broadleaf forests. However, the MODIS results correspond to spatially heterogeneous 500 m pixels, where errors due to random variations in canopy properties at finer resolution may be averaged.

One may argue that our results are obvious given previous modelling and validation studies (Shabanov and Gastellu-Etchegorry, 2018; Brown et al., 2019; Brown et al., 2020; Fernandes et al., 2023). However, the uncertainty of LAI and fAPAR estimates from algorithms such as SL2P could be due to numerous factors besides model error: input feature selection and measurement error, priors for RTM parameters, inversion algorithm and its calibration, and the representativeness and uncertainty of RM. Our experiment was necessary to quantify the role of model error related to canopy heterogeneity in bias observed with SL2P. On that grounds alone it is novel in that it the first study to provide empirical evidence at biome scale that including canopy heterogeneity

in the SL2P RTM would reduce both LAI and fAPAR bias. Even so, our findings have wider implications in that we have, for the first time, documented that there is a consistent bias-variance trade-off for LAI, and to weaker extent for fAPAR, for a widely used approach for mapping these variables using medium resolution multispectral imagery. It is the strength of support for this hypothesis that motivated the title of our study post hoc.

There are many heterogenous RTMs we could have used with SL2P-CCRS (e.g. Widlowski et al., 2007; Widlowski et al., 2015). 4SAIL2 was selected since its similarity with SAILH allowed us to isolate the effect of including canopy heterogeneity while preserving SAILH input parameters. This also simplified the specification of priors that may not be readily available for more complex RTMs. The parameterization of Smolander and Stenberg (2003) was used to account for the impact of shoot clumping as it is physically rigorous and requires only one parameter for which empirical data are available. These modifications to SAILH required 3 new parameters: crown cover, crown height-to-width ratio and needle-to-shoot area ratio. In this sense, SL2P represents a special case of SL2P-CCRS with only a change in priors corresponding to these three new parameters.

The three additional parameters describing spatial heterogeneity led us to hypothesize that SL2P-CCRS would reduce LAI bias but increase precision error (variance) in comparison to SL2P over forests due to a greater dispersion of priors used to generate the calibration database. It was unclear how this would impact fAPAR since there is empirical evidence that the relationship between fAPAR and simple multispectral vegetation indices are not highly sensitive to forest clumping (e.g. Myneni and Williams, 1994).

The use of representative datasets for SL2P-CCRS priors (Appendix B) reduced the dispersion of calibration database posteriors related to foliage biochemistry (Cab, Cdm, Cbp) in comparison to SL2P (Fig. 4), but also increased the dispersion of priors related to canopy structure other than LAI (CC,  $\chi$ ,  $\Gamma$ , ALA, N). SL2P global priors were retained for certain foliage biochemistry variables not available for forests within global databases (e.g. Kattege et al., 2020). We also noted cases where the SL2P restriction to only truncated normal or truncated uniform priors was not suitable when using global datasets (Bp and Bs, as shown in Appendix B). Nevertheless, the dispersion of the joint posteriors for SL2P was lower than that of SL2P-CCRS simply due to the assumption of constant CC,  $\gamma$ , and  $\chi$  for SL2P.

To ensure that priors unrelated to canopy heterogeneity could not potentially also result in SL2P-CCRS having lower precision than SL2P, we also quantified the sensitivity of SL2P to simply modifying priors other than CC,  $\gamma$ , and  $\chi$  to match those used in SL2P-CCRS ENF and DBF respectively (i.e. SL2P-LandCover). While this experiment is tangential to our research question regarding reducing LAI bias due to clumping, it verified that our comparisons of SL2P and SL2P-CCRS were not sensitive to priors unrelated to canopy spatial heterogeneity (Appendix D).

Direct validation supported our hypothesis that SL2P-CCRS would increase accuracy and decrease precision of LAI and fAPAR estimates, in comparison to SL2P, both at the population level and for conditional metrics for  $3 < \text{LAI} < 6$  (Table 11 and Table 12). The use of fiducial reference measurements was critical to bias quantification as SL2P appears almost unbiased when compared to in-situ LAI measurements without correction for clumping (so called effective LAI) (Fernandes et al., 2023). Other studies have reported almost unbiased LAI estimation using heterogeneous RTM inversion of multi-spectral reflectance over forests (Banksota et al., 2015; Ali et al., 2021; Brown et al., 2020). However, these studies used RM without correction for one or both of clumping and NPV, so their unbiased estimates are potentially biased in reality. In addition, previous studies generally only report population A, P and U statistics so their results could be impacted by cancelling errors or unbalanced sampling.

The difference between population and conditional fAPAR metrics highlights the need to use conditional metrics, as recommended by good practice (Fernandes et al., 2014), for variables that often have a limited

range over healthy mature forests. Additionally, the use of regression models robust to measurement error to estimate conditional metrics, as proposed in [Fernandes et al., 2023](#), simultaneously addresses the issues of measurement errors in retrievals and RM as well as the need to quantify the uncertainty in metrics due to limited sampling.

A new approach of comparing direct validation scatter plots and cross-validation density plots was used to evaluate the representativeness of both the in-situ data and RTM calibrations. The high probability density regions of the cross-validation comparisons qualitatively matched the in-situ sample density ([Fig. 5](#) and [Fig. 6](#)). This test is only a necessary condition for both datasets being representative since they could also both contain similarly biased samples. However, the fact that the RTM priors were calibrated using domain representative distributions for many parameters, and broad priors otherwise, suggests that the two datasets are indeed representative of North American forests. It would be interesting to use ancillary datasets, such as CC from forest databases, high-resolution imagery or LIDAR measurements (e.g. [Weinstein et al., 2021](#)), to further constrain RTM priors to or to compare the posteriors of the direct validation and cross-validation to edit the RTM calibration database.

The fact that there is a bias-variance trade-off for LAI when modifying SL2P with a heterogenous canopy RTM is not entirely surprising considering radiative transfer theory. The probability of photons recolliding ( $p_k$ ) or escaping ( $q_k$ ) from the canopy in each direction can be closely approximated using geometric functions of the scattering order  $k$  and the same probabilities for the infinite scattering order ( $p_\infty$ ,  $q_\infty$ ) ([Lewis et al., 2007](#)).  $p_\infty$  and  $q_\infty$  are, respectively, functions of the ratio of hemispherical gap fraction to LAI and the ratio of directional gap fraction to LAI ([Stenberg and Manninen, 2015](#)). These gap fraction quantities are functions of not just LAI but also canopy structure parameters like CC,  $\gamma$ , and  $\chi$  ([Nilson, 1999](#); [Möttus and Stenberg, 2008](#)). As a result, the relationship between reflectance and LAI is expected to be ill-posed for cases, such as forests, with substantial variation in canopy structure parameters other than LAI. This dimensional analysis may explain why SL2P performs well for LAI estimation over crops and grasslands where vegetation is relatively homogeneous (e.g. [Djamai et al., 2019](#)) and not forests where vegetation is clumped (e.g. [Brown et al., 2021a](#)) and why we observed that SL2P-CCRS has decreased precision over forests in comparison to SL2P. Our observations, and this theory, suggest that further reduction in SL2P-CCRS S2 LAI uncertainty for forests may require local constraints on canopy parameter priors from ancillary databases or empirical bias correction of posteriors as in [Fernandes et al. \(2023\)](#).

[Fernandes et al. \(2023\)](#) recommended replacing SL2P with either land cover specific bias corrections or testing a heterogenous RTM when mapping over forests and woodlands. Both solutions require land cover information to determine when to apply SL2P versus SL2P-CCRS. Such land cover information is available on an annual basis ([Buchhorn et al., 2020](#)) and uncertainty due to land cover error can be modelled by comparing estimates between SL2P and the forested area algorithms. But the question remains as to which strategy is preferable for forested areas. The SL2P-CCRS results from our study ([Table 11](#) and [Table 12](#)) are directly comparable with the bias corrected SL2P results from [Fernandes et al. \(2023\)](#) as they share the same inputs and RM. Accuracy error was  $-0.03$  LAI for bias corrected SL2P compared to  $-0.13$  LAI for SL2P-CCRS and  $-0.03$  fAPAR for bias corrected SL2P compared to  $-0.045$  fAPAR for SL2P-CCRS. However, for both LAI and fAPAR, the uncertainty of both methods differed by  $<10\%$ . While this may suggest that bias correction has less downside than SL2P-CCRS there are two reasons why SL2P-CCRS may be preferable for woody and forest land cover. Firstly, population statistics may not apply to individual retrievals. Indeed, the UAR for SL2P-CCRS LAI was  $0.57$  versus  $0.54$  for SL2P with bias correction. More importantly, bias correction may reduce the physical consistency between variables enforced through the RTM. We suggest that bias correction considered in [Fernandes et al. \(2023\)](#) may be preferable for applications where low bias is critical as long as local RM

are available to verify the calibrated correction is locally representative.

Our study used the identical inversion scheme and a similar RTM, except for canopy heterogeneity parameterization, as SL2P. This was done to conduct a controlled experiment to test the hypothesis that the lack of canopy heterogeneity in SL2P was the reason for the LAI bias observed in [Fernandes et al. \(2023\)](#). However, this control also resulted in three limitations that could have exacerbated the bias-variance trade-off observed for SL2P-CCRS: RTM model error, choice of input S2 features, and the inversion algorithm.

With respect to RTM model error, the uncertainty of 4SAIL2 exceeds that of RAMI reference models ([Widlowski et al., 2007](#)) so it is possible that bias could be further reduced by using a RTM with lower uncertainty. However, the bias-variance trade-off may not improve as all the RAMI reference models require more than double the number of RTM parameters, in comparison to SL2P-CCRS, to specify 3-dimensional canopy structure. In fact, we repeated our validation experiment using the FLIGHT 3-dimensional RAMI reference RTM ([North, 1996](#)) and the CCRS RM. Details regarding the FLIGHT priors are beyond the scope of this study, but the input features and neural network calibration were identical to SL2P-CCRS. LAI validation using the CCRS RM indicated virtually identical A,P and U between SL2P-CCRS and the version using FLIGHT (<https://az659834.vo.msecnd.net/eventsairwesteuprod/production-nikal-public/c25856c1cda8418884bb4a950c730fb6> accessed on January 23, 2023). These results need to be extended to the dataset used in the current study. However, they support our argument, based on canopy recollision probability, that there is a lower limit to the uncertainty of LAI estimation from multispectral reflectance in the absence of additional constraints on priors regarding canopy structure.

With respect to input S2 features, the use of eight S2 bands plus three geometric quantities results in a rather high dimensional feature space with non-negligible measurement error for inputs. Feature reduction, through the use of vegetation indices based on theoretical arguments (e.g. [Myneni and Williams, 1994](#)), RTM modelling (e.g. [Pallarés et al., 2019](#)), or active learning using reference datasets (e.g. [Djamai and Fernandes, 2021](#)) could reduce aleatoric uncertainty by improving the signal to noise of inputs and also reduce epistemic uncertainty by reducing the complexity of the fitted regression predictor. Further experiments in this regard should be performed to quantify if indeed feature reduction can reduce uncertainty of S2 LAI and fAPAR estimates for heterogeneous vegetation.

With respect to the inversion algorithm, our study used a shallow neural network that corresponds to a non-linear regression on tangent-sigmoid basis functions. The output of the network corresponds to the expected value of estimated LAI or fAPAR given inputs. In cases, such as for high LAI, where the input sensitivity to a change in either LAI or fAPAR is low, the distribution of regression residuals will be skewed, increasing the likelihood of observing local biases. Inversion algorithms allowing for skewed residuals, such as look-up-tables, random forests, k-NN regression, and gaussian process regression may serve to reduce bias at high LAI compared to neural networks ([Verrelst et al., 2015](#)). However, the greatest SL2P-CCRS bias was observed at moderate ( $<4$ ) LAI ([Fig. 7](#)) where the relationship between LAI and input reflectance had not yet saturated.

Our study could be improved with additional in-situ sampling, especially over canopies with high horizontal and vertical heterogeneity, and canopies with  $LAI > 4$  or  $fAPAR < 0.5$ . Higher spatial and temporal RM sampling could also be exploited to determine if the precision error of estimated variables decreases with spatial or temporal averaging. Such averaging is facilitated by the 20 m resolution of the S2 MSI inputs compared to the 100 m GCOS baseline requirement and the potential of a revisit interval of  $\leq 2.5d$  when S2 is combined with the planned Landsat NEXT mission that will have spectral bands similar to S2 (<https://landsat.gsfc.nasa.gov/satellites/landsat-next/> accessed on September 27, 2023).

Another limitation of our study was our approach to estimating LAI and fAPAR for mixed forests. 4SAIL2 does not account for spatial

heterogeneity common with mixed forests and prior information regarding the forest mixture for a given S2 pixel does not exist in current continental or global land cover maps. Our solution for estimating mixed forest LAI and fAPAR was to average the ENF and DBF estimates. This was facilitated by the incorporation of mixed forests when calibrating structure priors for both the ENF and DBF algorithms. Our strategy was effective in that it resulted in a reduction in overall bias for both LAI and fAPAR compared to SL2P. However, averaging ENF and DBF estimates over mixed forests may result in physical inconsistencies between retrieved LAI and fAPAR and requires broader priors for each class than if only broadleaf or needleleaf forests were considered. RTMs with 3D structure have been used to retrieve broadleaf forest LAI from multispectral imagery in a physically valid manner (Banskota et al., 2015) but published studies for mixed forests do not offer sufficient sites to provide a reliable estimate of bias. This may be in part due to the need to explicitly resolve within shoot scattering to ensure unbiased estimation of canopy reflectance within 3D models (Rochdi et al., 2006). Mixed forests RTM simulations are likely to require even more parameters than a 3D RM of a single species, potentially further increasing the variance of estimated fAPAR and LAI. Further research is required to quantify the thematic performance of S2 LAI and fAPAR estimates using inversion algorithms based on RTM simulations of mixed forests.

## 5. Conclusions

This study evaluated the potential of reducing bias observed by Fernandes et al. (2023) when using SL2P to estimate LAI over North American forests with S2 imagery. A new version of SL2P, SL2P-CCRS, was calibrated for North American evergreen needleleaf forests and broadleaf forests. SL2P-CCRS replaces the SAILH RTM with 4SAIL2, that accounts for variation in crown cover and crown shape, and by modelling shoot single scattering albedo as a function of the needle-to-shoot area ratio. SL2P-CCRS was calibrated using cover class specific priors for evergreen needleleaf and deciduous broadleaf woody vegetation and validated using the same dataset of in-situ RM as in Fernandes et al. (2023).

SL2P-CCRS reduced LAI bias by 65% and fAPAR bias by 31% in comparison to SL2P during direct validation with 1107 RM. LAI absolute bias reduced by ~0.5 LAI for LAI 3 and by ~1 LAI for LAI 6. The reduction in fAPAR absolute bias was generally <0.05. These findings support the hypothesis that SL2P LAI and fAPAR bias observed over North American forests in Fernandes et al. (2023) was due its use of a spatially homogenous RTM. However, the SL2P-CCRS reduction in bias compared to SL2P was accompanied by an increase in precision error that essentially resulted in only a 6% reduction in uncertainty for LAI and 5% for fAPAR. These findings support the hypothesis that there is a bias variance trade-off for both LAI and fAPAR when increasing the complexity of SL2P by using a RTM that accounts for spatial heterogeneity. Nevertheless, SL2P-CCRS increased the uncertainty agreement rate with GCOS requirements from 52% to 58% for LAI and 31% to 40% for fAPAR, suggesting that the trade-off is worthwhile and that algorithms such as SL2P-CCRS, that use a spatially heterogeneous RTM, should be applied for mapping fAPAR and LAI from S2 p measurements. Moreover, the uncertainty of downstream Level 3 products may be further reduced to the extent precision errors are random in space and time.

Cross-validation using RTM simulations supported the representativeness of the field measurements as well as the bias observed with

SL2P. Cross-validation was also able to qualitatively replicate the bias-variance trade-off for LAI but not for fAPAR. We hypothesize that extreme canopy parameterizations within the RTM simulations resulted in large biases for cross-validation fAPAR for fAPAR<0.9. This artifact may be due to the use of priors with large dispersion for canopy parameters when producing RTM simulations. Ancillary databases should be used to reduce this dispersion for testing purposes.

The observed bias-variance trade-off for LAI and fAPAR leads us to hypothesize that there may be a fundamental limit to the uncertainty of globally applicable algorithms such as SL2P when using medium resolution single view angle multi-spectral reflectance spectra as input. If so, further improvements in the uncertainty of medium resolution fAPAR and LAI products from satellite imagers such as S2 may require integration of local ancillary data for constraining priors related to canopy heterogeneity or the incorporation of temporal or spatial constraints. In the meantime, our results indicate SL2P-CCRS should be used in preference to SL2P for mapping LAI and fAPAR over forests.

## CRediT authorship contribution statement

**Richard Fernandes:** Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Najib Djamel:** Writing – review & editing. **Kate Harvey:** Software, Validation. **Gang Hong:** Validation, Writing – original draft, Writing – review & editing. **Camryn MacDougall:** Software, Validation, Visualization. **Hemit Shah:** Software. **Lixin Sun:** Software, Validation, Writing – original draft, Writing – review & editing.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Richard Fernandes reports financial support was provided by Government of Canada. Richard Fernandes reports a relationship with Government of Canada that includes: employment.

## Data availability

Data will be made available on request.

## Acknowledgements

We acknowledge Dr. Marie Weiss and Dr. Fred Baret for providing the original SL2P code that was used as a basis for the versions of SL2P and SL2P-CCRS used in this study. We acknowledge the use of modified Sentinel 2 data and derived products. This study has been undertaken using data from GBOV “Ground Based Observation for Validation” (<https://land.copernicus.eu/global/gbov>) funded by European Commission Joint Research Centre FWC932059, part of the Global Component of the European Union’s Copernicus Land Monitoring Service. GBOV products are developed and managed by ACRI-ST with the support from University College London, University of Leicester, University of Southampton, University of Valencia and Informus GmbH. We thank the National Ecological Observatory Network for the measurements collected in the field and used to generate GBOV products. We thank excellent reviewers and the editors of this paper. The work was funded by Natural Resources Canada’s Earth Observation for Cumulative Effects Project.

## Appendix A. Acronyms

**Table A1**  
Definition of acronyms.

Parameter	Definition
$p_k$	Canopy recollision probability for scattering order k
$q_k$	Canopy escape probability for scattering order k
A	Accuracy
ALA	Average Leaf Angle
B	Baseline
Bs	Soil brightness
Cab	Foliage chlorophyll A + B content
Cbp	Foliage brown pigment content
CC	Crown Cover
Cdm	Foliage dry matter content
Cw_rel	Foliage relative water content
fAPAR	Fraction of absorbed PAR
fCOVER	Fraction of canopy cover
G	Goal
HsD	Hotspot Dimension parameter
k	Scattering order
LAI	Leaf Area Index
N	PROSPECTD leaf thickness parameter
P	Precision
PAR	Photosynthetically active radiation
RAA	Relative azimuth angle
SZA	Solar zenith angle
T	Threshold
U	Uncertainty
UAR	Uncertainty agreement ratio
VZA	View zenith angle
$\gamma$	Needle to shoot area ratio
$\rho$	Bi-directional reflectance factor
$\chi$	Crown height to width ratio

## Appendix B. Calibration of univariate prior distributions

This appendix documents the calibration of RTM priors for ENF and DBF algorithms. For convenience, the number of samples and sources of data used are summarized in [Table B1](#).

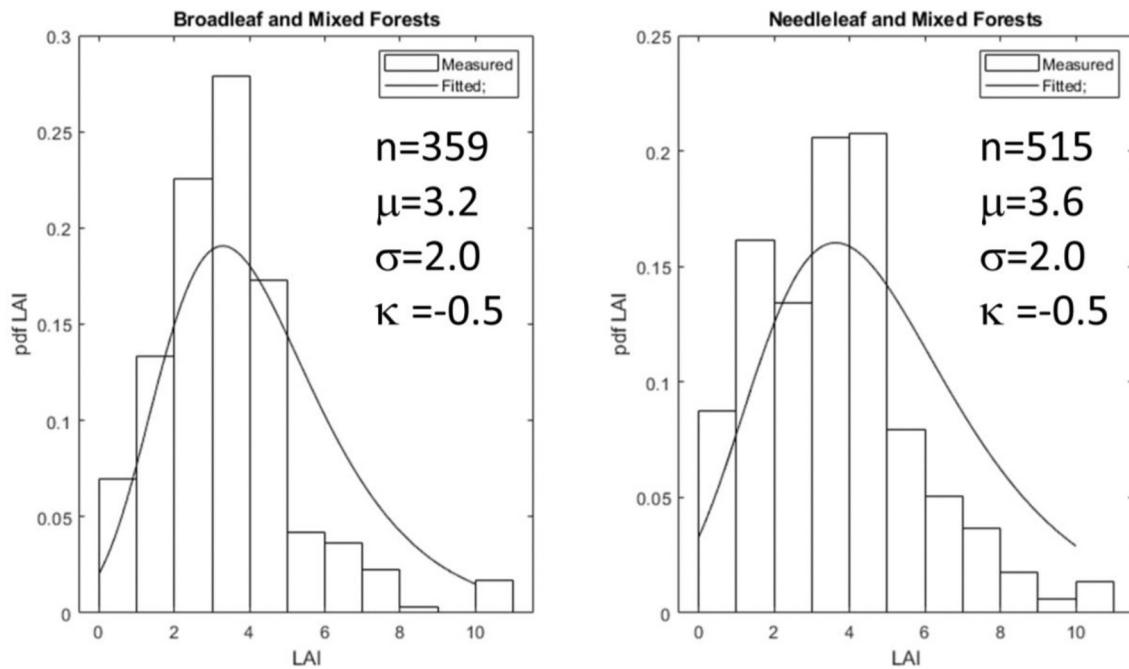
**Table B1**

Summary of measurements used to calibrate prior distributions of RTM parameters. Acronyms are defined in [Appendix A](#). n indicates number of samples.

Parameter	n	Representative	Sources
LAI	874	Yes	<a href="#">Fernandes et al., 2003</a> ; <a href="#">Fernandes et al., 2018</a> ; <a href="#">Fernandes et al., 2023</a> .
CC	1107	No	<a href="#">Fernandes et al., 2023</a> .
ALA	82	No	<a href="#">Barclay, 2001</a> ; <a href="#">Yan et al., 2021</a> ; <a href="#">Pisek et al., 2022</a> .
HsD	22	No	<a href="#">WB2016</a>
$\chi$	250	No	<a href="#">Purves et al., 2007</a> .
N	Unknown	No	<a href="#">WB2016</a>
Cab	1544	No	<a href="#">Hosgood et al., 1994</a> ; <a href="#">Féret et al., 2017</a> ; <a href="#">Croft et al., 2017</a> ; <a href="#">Donnelly et al., 2020</a> .
Cdm	33,338	Yes	<a href="#">Kattge et al., 2020</a>
Cw_rel	19,206	Yes	<a href="#">Kattge et al., 2020</a>
Cbp	Unknown	No	<a href="#">WB2016</a>
$\gamma$	564	No	<a href="#">Stenberg, 1996</a> ; <a href="#">Chen et al., 1997</a> ; <a href="#">Stenberg et al., 2001</a> ; <a href="#">Ishii et al., 2012</a> ; <a href="#">Rautainnen et al., 2012</a> ; <a href="#">Pope and Treitz, 2013</a> ; <a href="#">Qi et al., 2014</a> .
Bs	1500	Yes	<a href="#">WB2016</a>

### Leaf Area Index

LAI distributions were fitted to data from [Fernandes et al. \(2003\)](#), [Fernandes et al. \(2018\)](#) and [Fernandes et al. \(2023\)](#) using truncated Generalized Extreme Value Distributions with mean value constrained to the sample median after qualitative assessment of available MATLAB distributions (<https://www.mathworks.com/help/stats/continuous-distributions.html>) (Fig. B).



**Fig. B1.** Measured LAI for North American forests together with fitted truncated extreme value distributions with limits [0,10].

#### Crown Cover

CC distributions were fitted to data from [Fernandes et al. \(2023\)](#). CC was estimated as:

$$CC = fCOVER_{up} + (1 - fCOVER_{up})fCOVER_{down} \quad (\text{A2.1})$$

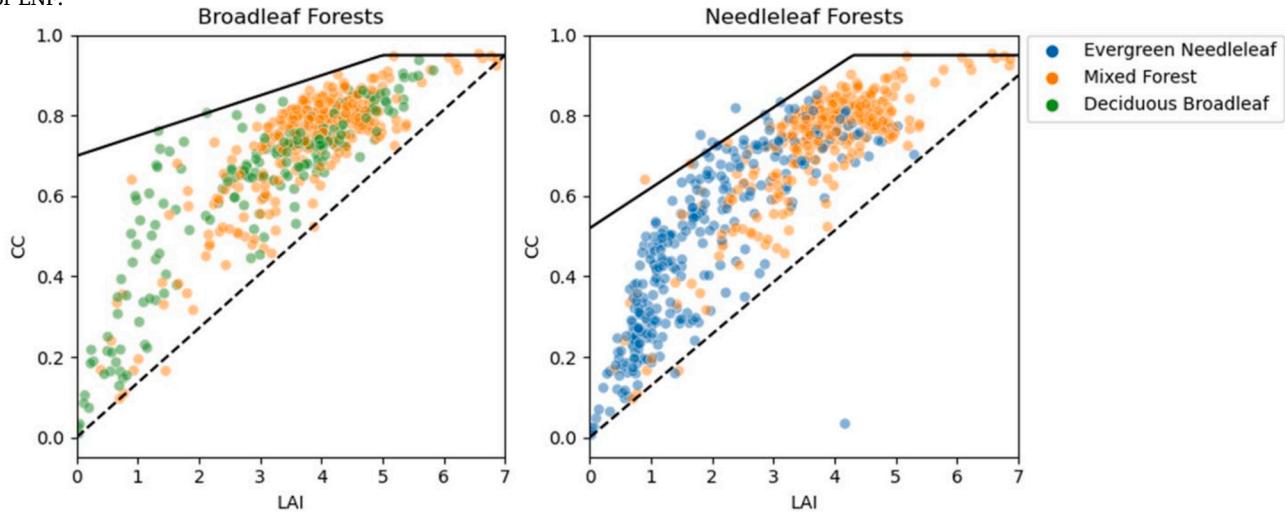
The data likely under-sampled canopies with high CC and low LAI considering measurements focused on healthy undisturbed vegetation during the growing season. Moreover, both the ENF and S DBF regressions had to also incorporate mixed forests. As a result, the CC distribution was assumed uniform with range [0.01,0.95] for both ENF and DBF. Linear scaling parameters for the conditional distribution of CC given LAI were qualitatively fit ([Fig. B1](#)) giving:

$$LAI_{lb} = 0.01, LAI_{ub} = 10, x_{min}(LAI_{lb}) = 0.01, x_{max}(LAI_{lb}) = 0.7, x_{min}(LAI_{ub}) = 1.35, x_{max}(LAI_{ub}) = 1.35 \quad (\text{A2.2})$$

for DBF, and

$$LAI_{lb} = 0.01, LAI_{ub} = 10, x_{min}(LAI_{lb}) = 0.01, x_{max}(LAI_{lb}) = 0.52, x_{min}(LAI_{ub}) = 1.2, x_{max}(LAI_{ub}) = 1.29 \quad (\text{A2.3})$$

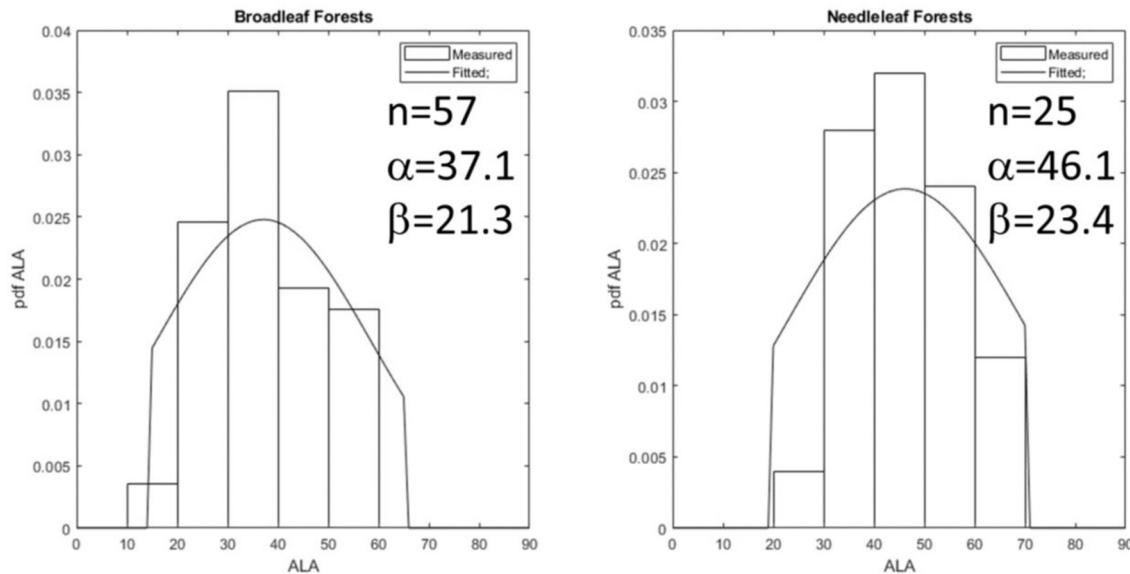
for ENF.



**Fig. B1.** Estimated CC versus observed LAI for assumed values of canopy extinction coefficient ( $k$ ) based on data from [Fernandes et al. \(2023\)](#). Solid (dashed) line corresponds to upper and lower linear bounds as a function of LAI.

#### Average Leaf Angle

ALA data were taken from measurements over temperate and boreal forests ([Barclay, 2001](#); [Yan et al., 2021](#), and [Pisek et al., 2022](#)). Truncated beta distributions were fit using MATLAB fitter rather than Normal distributions due to the asymmetry of the empirical histograms of the data ([Fig. B2](#)).



**Fig. B2.** Measured ALA (degrees from horizontal) together with fitted truncated beta distributions with limits [015,65] for broadleaf and [20,70] for needleleaf forests.

#### Needle to Shoot Area Ratio

Measured values for  $\gamma$  are biased towards Boreal species (Stenberg, 1996; Chen et al., 1997; Stenberg et al., 2001; Rautainen et al., 2012; Pope and Treitz, 2013) with only two studies considering temperate forests (Ishii et al., 2012; Qi et al., 2014). Subsequently a uniform distribution fitted to the range of reported measurement of  $\gamma$  was used.

#### Foliage Hot Spot Dimension

Due to the lack of data the HsD parameter prior from WB2016 was used for all cover classes. This parameter has less impact on 4SAIL2 estimates of  $\rho$  in comparison to SAILH since 4SAIL2 uses CC and  $\zeta$  to model a crown scale hotspot that is likely to be far greater than foliage scale hot-spots for the acquisition geometry of S2 where the view direction is away from the principal plane in North America (Roy et al., 2017).

#### Crown Height to Width Ratio

There are substantial forest mensuration databases for calibrating priors for  $\zeta$  (Purves et al., 2007). However, the distribution of  $\chi$  will depend on both species and site conditions. As such, a uniform distribution was adopted based on typical specific values for  $\chi$  (Purves et al., 2007).

#### PROSPECTD N Parameter

The N parameter was fitted during calibration of PROSPECTD (Féret et al., 2017). Here the range used in WB2016 was used here but with a uniform distribution due to the lack of representativeness of the fitted data. WB2020 was not used as the range there only extends to 1.8 versus 2.2 in WB2016 even though higher N values have been observed for relatively thin leaves (Boren et al., 2019).

#### Foliage Chlorophyll A + B Content

WB2016 use data from Féret et al. (2008) to calibrate the prior for Cab. This data includes both non-tree samples as well as tree samples with extremely low Cab ( $<15 \mu\text{g.cm}^{-2}$ ) not typically seen in healthy green foliage corresponding to the definition of LAI (Zarco-Tejada et al., 2004; Croft et al., 2017; Féret et al., 2017; Donnelly et al., 2020). For example, Cab rarely exceeds  $75 \mu\text{g.cm}^{-2}$  for datasets with only forested samples and the range of Cab for the LOPEX1993 dataset used in Féret et al. (2008) exceeded  $75 \mu\text{g.cm}^{-2}$  for only one forest species (Sycamore maple). Since Cab varies with species, season, and site conditions it was decided that the limited in-situ measurements were not representative. As such, the range from  $15 \mu\text{g.cm}^{-2}$  to  $75 \mu\text{g.cm}^{-2}$  was used for Cab based on the data for forest species from the Lopex93 database used by Féret et al. (2008) as well as data presented in Zarco-Tejada et al. (2004), Croft et al. (2017), Féret et al. (2017), and Donnelly et al. (2020).

#### Foliage Dry Matter Content

The TRY database (Kattge et al., 2020) provided sufficient samples of healthy tree foliage to calibrate fitted truncated beta distributions for Cdm for EBF and ENF cover classes using the MATLAB fitter routine (Fig. B3).

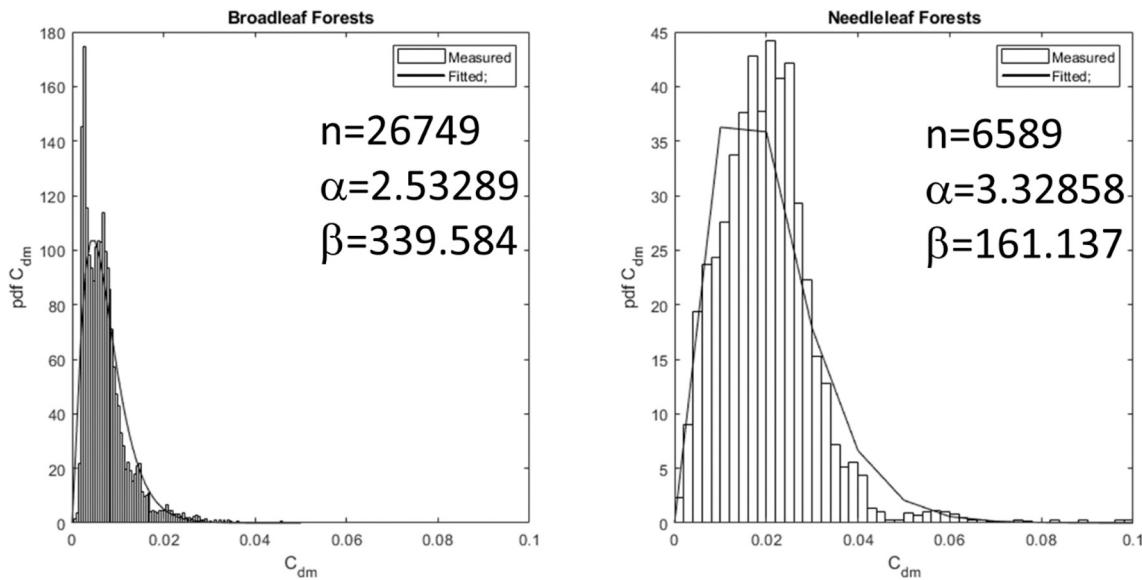


Fig. B3. Measured Cdm for forest foliage from the TRY database together with fitted truncated beta distributions with limits [0,0.1].

#### Foliage Relative Water Content

WB2016 and WB2020 assumed a modal value of 0.75 for Cw\_rel, with the former assigning an arbitrary uniform distribution with range [0.60,0.85] and the latter a normal distribution with mean 0.75 and standard deviation 0.08. Here, the TRY database (Kattge et al., 2020) provided sufficient samples of tree foliage to calibrate fitted truncated beta distributions for Cw\_rel for EBF and ENF cover classes using the MATLAB fitter routine (Fig. B4). The distribution was truncated at Cw\_rel > 0.5 to reduce the likelihood of fitting to foliage sampled during extreme water stress (Barr and Weatherley, 1962).

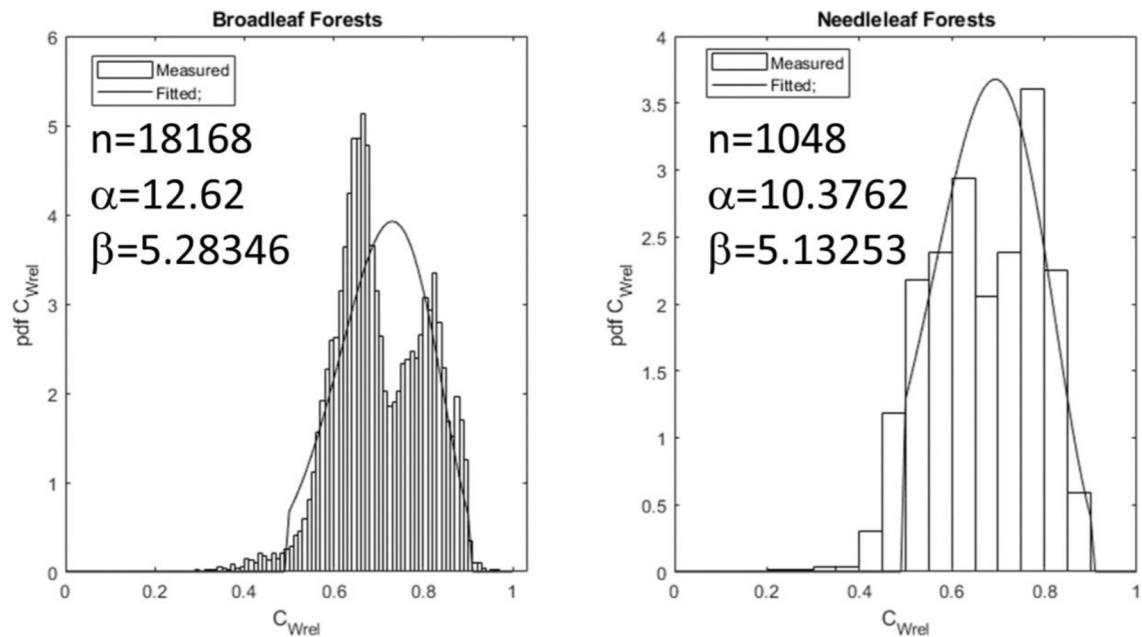
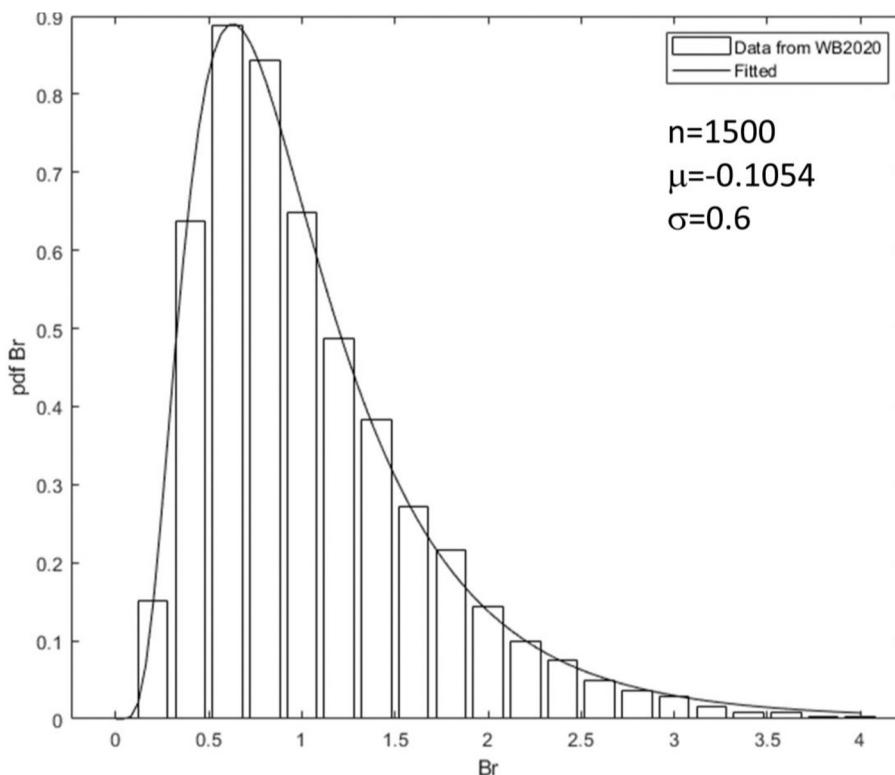


Fig. B4. Measured Cwrel for forest foliage from the TRY database together with fitted truncated beta distributions with limits [0.5,0.9].

#### Soil Brightness Parameter

WB2016 fit Bs to a large database of 1500 soil spectra (Liu et al., 2002). As the original measurements were not available, we digitized the histogram of Bs given in Fig. 9 of WB2020 and used MATLAB fitter to calibrate a truncated log normal distribution (Fig. B5). This distribution was used for both ENF and DBF forests although the ENF forest also included additional spectra from Boreal forests (Miller et al., 1997).



**Fig. B5.** Histogram of  $Br$  values from WB2020 together with fitted truncated log normal distribution with limits [0,4].

### Appendix C. SL2P sensitivity to sampling of RTM priors

SL2P (Weiss and Baret, 2016) uses orthogonal sampling of priors to produce inputs to RTM simulations. Their justification follows:

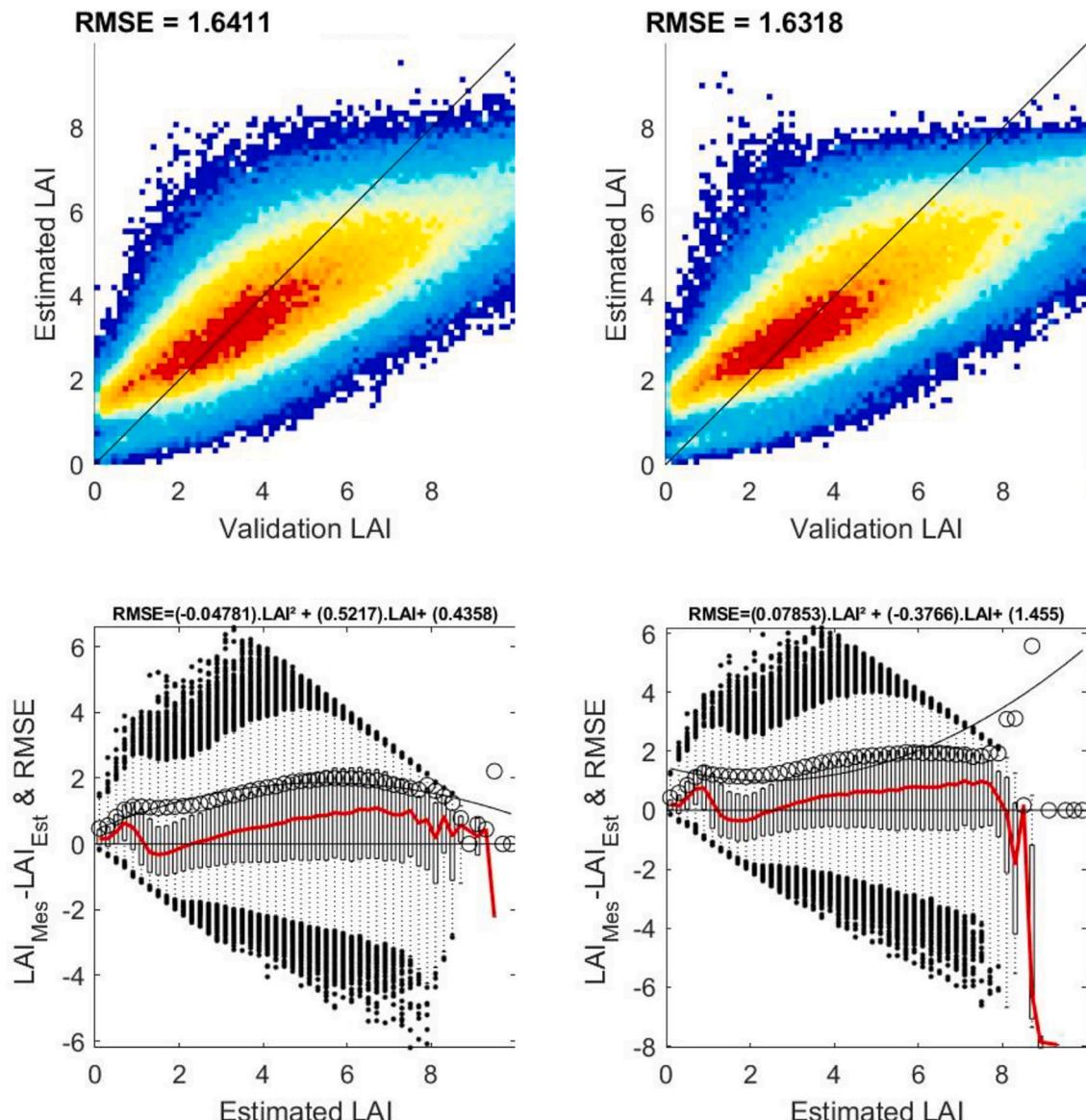
"The sampling scheme is based on a full orthogonal experimental plan. This consists of identifying classes of values for each variable. Then all the combinations of classes are sampled once. Finally, the actual values of each variable are randomly drawn within the range of variation defined by the corresponding class, according to the distribution law specified for the variable considered. This process allows accounting for all the interactions, while having the range of variation for each variable densely and near randomly populated. "

The ATBD for SL2P-CCRS (Fernandes et al., 2021; <https://github.com/rfernand387/SL2P-CCRS/blob/main/Documents/SL2P-CCRS-ATBD.docx>) evaluates 5 sampling schemes and describes why Sobol sampling is preferred. We quote:

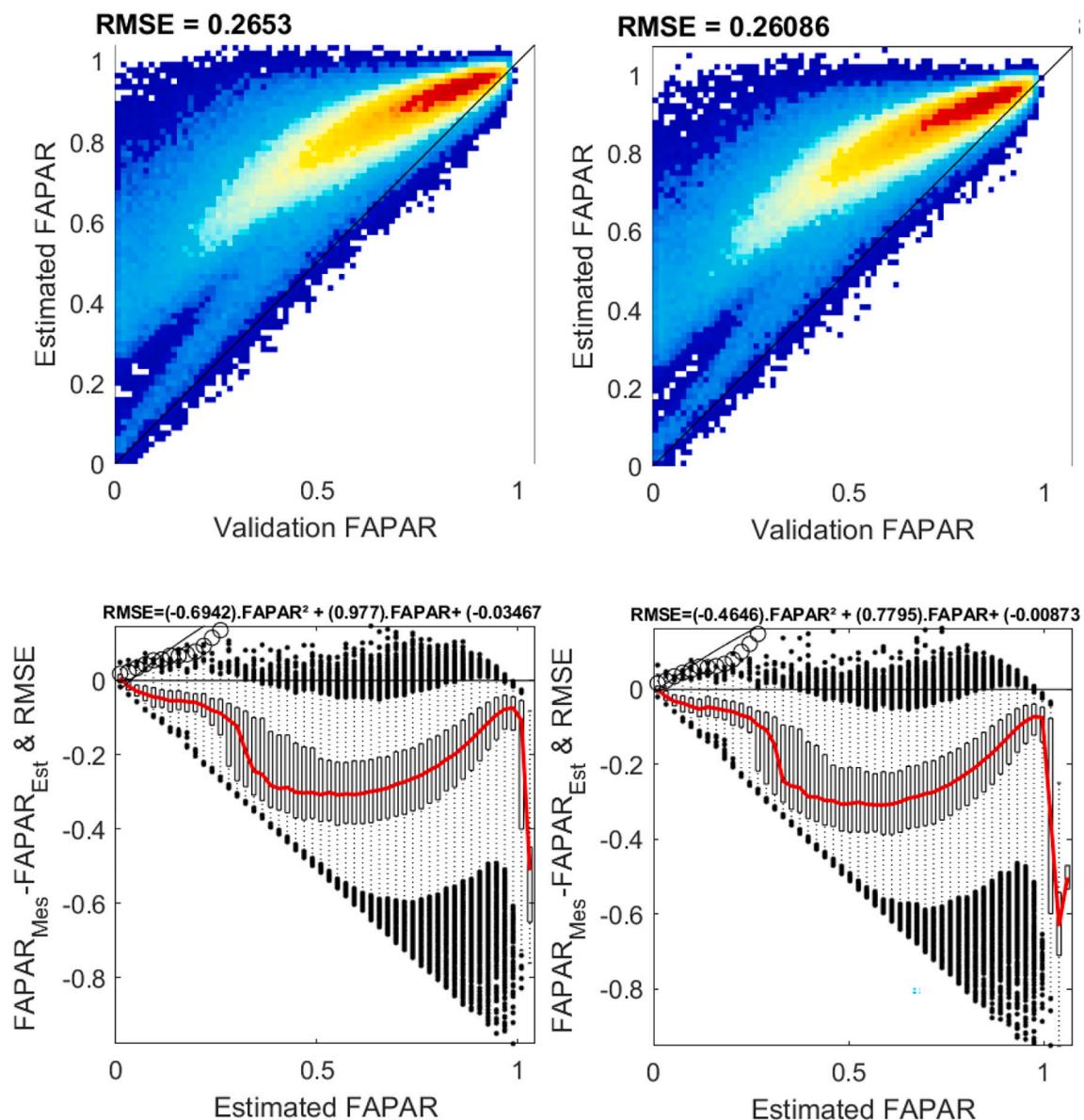
"One of five sampling schemes are provided to sample the joint distribution of either canopy or atmosphere parameters. All of these schemes are unbiased so the calibrated regression will be unbiased if tested with samples from the same distribution. However, the precision of the calibrated network is a function of training sample size  $N$  and the effective dimensionality of the problem  $d$  (which will lie somewhere between 1 and 11 for SL2P-CCRS). Full orthogonal sampling guarantees a reduction rate better than Monte Carlo and potentially as good as most of the other sampling schemes. Full orthogonal has been used for single class applications by SL2P but is problematic if the sample size must vary since increases in sample size must be performed in steps of  $\sim ND/(D - 1)$  and resampling, for example when reusing simulations for another class, must be done carefully to avoid large gaps between samples. Scrambled Halton sampling offers better precision than Monte Carlo sampling when the effective dimensionality of the problem space,  $d$ , is  $< 3$  but does not always guarantee better performance than the best case Full Orthogonal. However, scrambled Sobol sampling gives better precision than Full Orthogonal sampling for  $d < 6$  and better precision than Monte Carlo sampling for  $d < 9$ . Scrambled Sobol sampling is used as default for calibration."

Here we evaluate the sensitivity of SL2P to Orthogonal versus Sobol sampling by comparing cross-validation of SL2P using either scheme with no other changes. SL2P normally uses 41,472 orthogonal samples following WB2020. Sobol sampling requires increments in powers of 2 so 65,336 samples are used to produce the calibration database with Sobol sampling. Cross validation is performed using the ENF calibration database as this dataset is independent of both SL2P variants and includes heterogeneous canopy simulations. For brevity we show the MATLAB diagnostic figures rather than reformatted publication ready figures.

Both population RMSE and the conditional distribution of residuals and their conditional bias is similar for  $LAI < 7$  (Fig. C1) and  $fAPAR < 0.98$  (Fig. C2). Values diverge for above these ranges as the priors distributions have low probability measure at very high LAI and fAPAR. However, the likelihood of observing LAI or fAPAR above this range in-situ is small for north American forests. In summary, negligible differences in SL2P performance was observed between use of 41,472 orthogonal samples and 65,536 Sobol samples. This supports the assumption that differences between SL2P and SL2P-CCRS observed in the manuscript are not related to differences in the sampling scheme or sample size used in WB2020 and the sampling scheme and sample size used for SL2P-CCRS.



**Fig. C1.** LAI cross validation of SL2P with 65,536 Sobol samples (left columns) versus SL2P with 41,472 orthogonal samples (right columns) using ENF simulated reflectance and corresponding LAI. Upper row shows density histograms with highest density indicated in red and lowest in dark blue as well as 1:1 line. Lower columns indicate 50%ile box plots of conditional residuals together with the accuracy (red), uncertainty (hollow circles and black line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. C2.** fAPAR cross validation of SL2P with 65,536 Sobol samples (left columns) versus SL2P with 41,472 orthogonal samples (right columns) using ENF simulated reflectance and corresponding fAPAR. Upper row shows density histograms with highest density indicated in red and lowest in dark blue as well as 1:1 line. Lower columns indicate 50%ile box plots of conditional residuals together with the accuracy (red), uncertainty (hollow circles and black line).

## Appendix D. Validation of SL2P using Land Cover Specific Priors (SL2P-LandCover)

The thematic performance of SL2P and SL2P-CCRS estimates of fAPAR, and LAI were compared with estimates using SL2P with SL2P-CCRS land cover specific priors given in [Table D1](#) and [Table D2](#) for the ENF and DBF classes respectively (SL2P-LandCover). Each SL2P-LandCover algorithm was calibrated using the same sampling scheme, sample size, and neural network architecture as SL2P for deciduous broadleaf forest and evergreen needleleaf forest classes separately. As with SL2P-CCRS, for mixed-forest ESUs the SL2P-LandCover estimate corresponded to the average of the SL2P-LandCover ENF network and SL2P-LandCover DBF network estimates. Otherwise, the network corresponding to the RM ESU land cover class was used. Direct validation was replicated as described in [Section 2.2.2](#) and scatter plots of estimated versus predicted Lai and fAPAR were compared between SL2P, SL2P-CCRS and SL2P and SL2P-LandCover ([Fig. D1](#)).

For LAI, SL2P-LandCover resulted in an increase in negative bias ( $A = -0.92$ ) versus both SL2P ( $A = -0.38$ ) and SL2P-CCRS ( $A = -0.13$ ). The decreased accuracy resulted in SL2P-LandCover having an uncreased uncertainty ( $U = 1.40$ ) and decrease in uncertainty agreement ratio ( $UAR = 0.39$ ) compared to SL2P ( $U = 0.91$ ,  $UAR = 0.51$ ) and SL2P-CCRS ( $U = 0.86$ ,  $UAR = 0.58$ ). Qualitatively, the scatter plot of SL2P-Land Cover showed a similar large bias as SL2P and similar low precision as SL2P-CCRS.

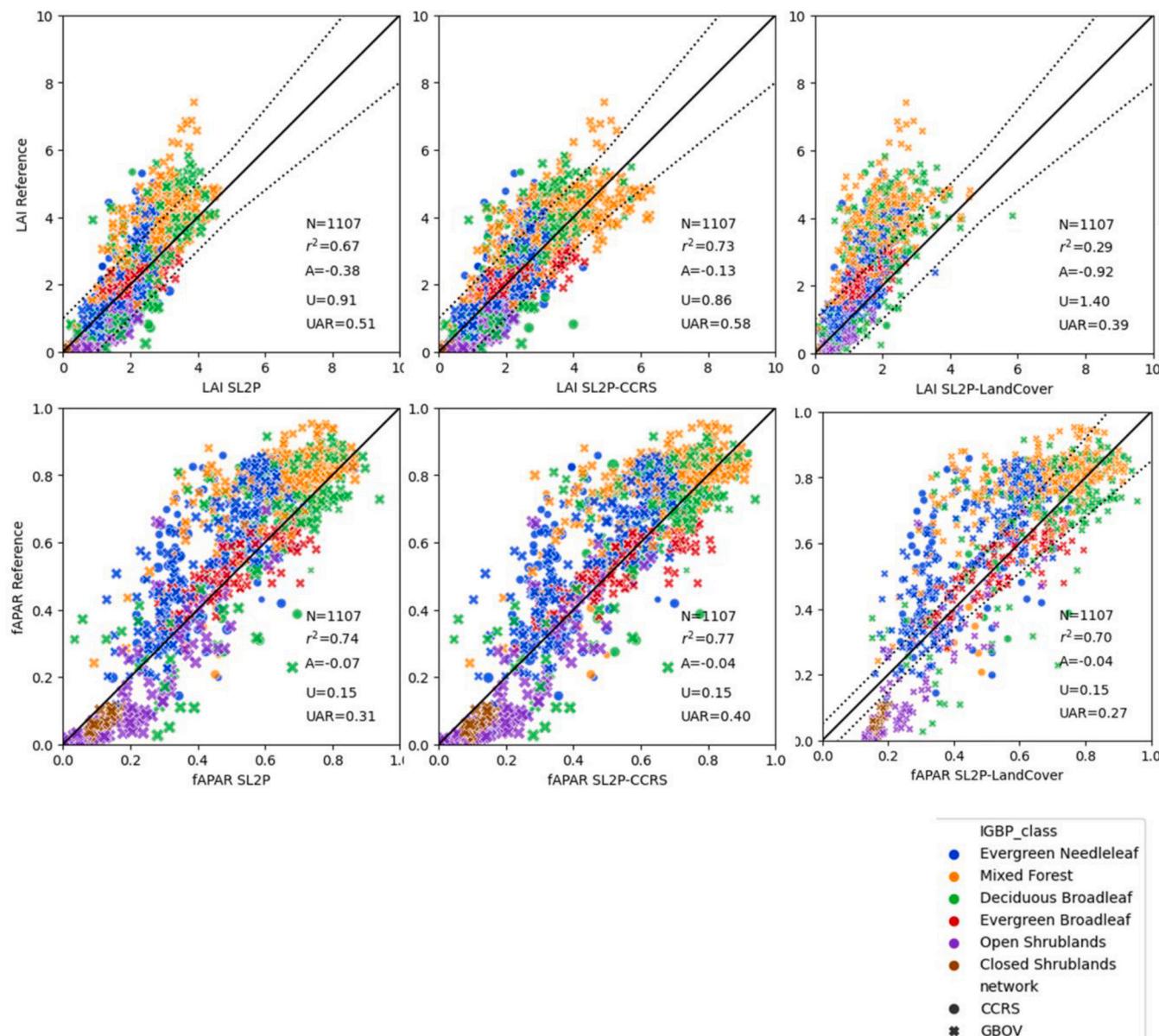
For fAPAR, SL2P-LandCover resulted in no change in slightly lower bias ( $A = -0.04$ ) than SL2P ( $A = -0.07$ ) and the same bias as SL2P-CCRS ( $A = -0.04$ ). The uncertainty of all three algorithms was identical at  $-0.15$ . However, SL2P-Land Cover had much lower UAR (0.27) than SL2P (UAR = 0.31) and SL2P-CCRS (UAR = 0.40). Qualitatively, the scatter plot of SL2P-Land Cover showed lower precision compared to SL2P and SL2P-CCRS for reference fAPAR between 0.3 and 0.8 and a consistent underestimate for reference fAPAR < 0.2.

**Table D1**  
SL2P-LandCover RTM parameter bounds and scaling ranges for deciduous broadleaf forests.

Symbol	$x_{lb}$	$x_{ub}$	$x_{min}(LAI_{lb})$	$x_{max}(LAI_{lb})$	$x_{min}(LAI_{ub})$	$x_{max}(LAI_{ub})$
LAI	0	10	–	–	–	–
CC	1	1	1	1	1	1
ALA	20	60	20	60	20	60
HsD	0.1	0.5	0.1	0.5	0.1	0.5
$\chi$	N/A	N/A	N/A	N/A	N/A	N/A
N	1.1	2.2	1.1	2.2	1.1	2.2
Cab	15	65	15	65	45	60
Cdm	0.005	0.01	0.005	0.01	0.005	0.01
Cw_rel	0.5	0.9	0.5	0.9	0.75	0.90
Cbp	0	0.2	0	0.2	0	0.2
$\gamma$	N/A	N/A	N/A	N/A	N/A	N/A
Bs	0.5	3.5	0.5	3.5	0.5	1.2

**Table D2**  
SL2P-LandCover RTM parameter bounds and scaling ranges for evergreen needleleaf forests.

Symbol	$x_{lb}$	$x_{ub}$	$x_{min}(LAI_{lb})$	$x_{max}(LAI_{lb})$	$x_{min}(LAI_{ub})$	$x_{max}(LAI_{ub})$
LAI	0	10	–	–	–	–
CC	0.01	0.95	0.01	0.52	1.2	1.35
ALA	20	70	20	70	20	70
HsD	0.1	0.5	0.1	0.5	0.1	0.5
$\chi$	N/A	N/A	N/A	N/A	N/A	N/A
N	1.1	2.2	1.1	2.2	1.1	2.2
Cab	15	65	15	65	45	60
Cdm	0.005	0.01	0.005	0.01	0.005	0.01
Cw_rel	0.7	0.9	0.7	0.9	0.75	0.90
Cbp	0	0.2	0	0.2	0	0.2
$\gamma$	N/A	N/A	N/A	N/A	N/A	N/A
Bs	0.5	3.5	0.5	3.5	0.5	1.2



**Fig. D1.** Estimated versus reference LAI (top row) and fAPAR (bottom row) for SL2P, SL2P-CCRS and SL2P-Land Cover for all reference ESUs. Thematic performance statistics corresponding to sample size (N), coefficient of determination ( $r^2$ ), accuracy (A), uncertainty (U) and uncertainty agreement ratio (UAR) are provided for each algorithm.

## References

- Ali, A.M., Darvishzadeh, R., Skidmore, A., Gara, T.W., Heurich, M., 2021. Machine learning methods' performance in radiative transfer model inversion to retrieve plant traits from Sentinel-2 data of a mixed mountain forest. *Int. J. Digit. Earth* 14 (1), 106–120. <https://doi.org/10.1080/17538947.2020.1794064>.
- Banskota, A., Serbin, S., Wynne, R., Thomas, V., Falkowski, M., Kayast, N., Townshend, P., 2015. An LUT-based inversion of DART model to estimate forest LAI from hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8 <https://doi.org/10.1109/JSTARS.2015.2401515>.
- Barclay, H., 2001. Distribution of leaf orientations in six conifer species. *Botany* 79, 389–397.
- Barr, H.D., Weatherley, P.E., 1962. A re-examination of the relative turgidity technique for estimating water deficit in leaves. *Aust. J. Biol. Sci.* 15, 413–428.
- Boren, E.J., Boschetti, L., Johnson, D.M., 2019. Characterizing the variability of the structure parameter in the PROSPECT leaf optical properties model. *Remote Sens.* 11 <https://doi.org/10.3390/rs11101236>.
- Brown, L., Otougu, B., Dash, J., 2019. Estimating forest leaf area index and canopy chlorophyll content with Sentinel-2: an evaluation of two hybrid retrieval algorithms. *Remote Sens.* 11 (15), 1752. <https://doi.org/10.3390/rs11151752>.
- Brown, L.A., Meier, C., Morris, H., Pastor, Brown, 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 Sens. Environ.* 247 <https://doi.org/10.1016/j.rse.2020.111935>.
- Brown, L., Fernandes, R., Djamai, N., Meier, C., Gobron, N., Morris, H., Canisius, C., Bai, G., Lerebourg, C., Lanconelli, C., Clerici, M., Dash, J., 2021a. Validation of baseline and modified Sentinel-2 level 2 prototype processor leaf area index retrievals over the United States. *ISPRS J. Photogramm. Remote Sens.* 175, 71–87. <https://doi.org/10.1016/j.isprsjprs.2021.02.020>.
- Brown, L., 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., 2021b. Fiducial reference measurements for vegetation bio-geophysical variables: an end-to-end uncertainty evaluation framework. *Remote Sens.* 13, 3194. <https://www.mdpi.com/2072-4292/13/16/3194>.
- 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 Ecol. Evol.* 00, 1–12. <https://doi.org/10.1111/2041-210X.14199>.
- Buchhorn, M., Lesiv, M., Tsednbazar, N.-E., Herold, M., Bertels, L., Smets, B., 2020. Copernicus global land cover layers-collection 2. *Remote Sens.* 108 <https://doi.org/10.3390/rs12061044>.

- Chen, J.M., 1996. Canopy architecture and remote sensing of the fraction of photosynthetically active radiation absorbed by boreal conifer forests. *IEEE Trans. Geosci. Rem. Sens.* 34 <https://doi.org/10.1109/36.544559>.
- Chen, J.M., Leblanc, S.G., 1997. A four-scale bidirectional reflectance model based on canopy architecture. *IEEE Trans. Geosci. Remote Sens.* 35 <https://doi.org/10.1109/36.628798>.
- Chen, J.M., Rich, P.M., Gower, S.T., Norman, J.M., Plummer, S., 1997. Leaf area index of boreal forests: theory, techniques, and measurements. *J. Geophys. Res.* 102 (D24) <https://doi.org/10.1029/97JD01107>.
- Commission for Environmental Cooperation, 2022. CEC Terrestrial Ecological Regions of North America, Montreal. Commission for Environmental Cooperation accessed at <http://www.cec.org/north-american-environmental-atlas/north-american-forests-2022/>.
- Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B., Staebler, R.M., 2017. Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Glob. Chang. Biol.* 23 <https://doi.org/10.1111/gcb.13599>.
- Djamai, N., Fernandes, R., 2021. Active learning regularization increases clear sky retrieval rates for vegetation biophysical variables using Sentinel-2 data. *Remote Sens. Environ.* 254 <https://doi.org/10.1016/j.rse.2020.112241>.
- Djamai, N., Fernandes, R., Weiss, M., McNairn, H., Goita, K., 2019. Validation of the sentinel simplified level 2 product prototype processor (SL2P) for mapping cropland biophysical variables using Sentinel-2/MSI and Landsat-8/OLI data. *Remote Sens. Environ.* 225, 416–430. <https://doi.org/10.1016/j.rse.2019.03.020>.
- Donnelly, A., Yu, R., Rehberg, C., 2020. Leaf chlorophyll estimates of temperate deciduous shrubs during autumn senescence using a SPAD-502 meter and calibration with extracted chlorophyll. *Ann. For. Sci.* 77, 30. <https://doi.org/10.1007/s13595-020-00940-6>.
- Estevez, J., Salinero Delgado, M., Berger, K., Pipia, L., Rivera Caicedo, J., Wocher, M., Reyes Munoz, P., Tagliabue, G., Boschetti, M., Verrelst, J., 2022. Gaussian processes retrieval of crop traits in Google earth engine based on Sentinel-2 top-of-atmosphere data. *Remote Sens. Environ.* 273 <https://doi.org/10.1016/j.rse.2022.112958>.
- Fang, H., Baret, F., Plummer, S., Schaepman-Strub, G., 2019. An overview of global leaf area index (LAI): methods, products, validation, and applications. *Rev. Geophys.* 57, 739–799. <https://doi.org/10.1029/2018RG000608>.
- Féret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P.R., Ustin, S. L., le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: advances in the leaf optical properties model 642 separating photosynthetic pigments. *Remote Sens. Environ.* 112 <https://doi.org/10.1016/j.rse.2008.02.012>.
- Féret, J.-B., Gitelson, A.A., Noble, S.D., Jacquemoud, S., 2017. PROSPECT-D: towards modeling leaf optical properties through a complete lifecycle. *Remote Sens. Environ.* 193 <https://doi.org/10.1016/j.rse.2017.03.004>.
- Fernandes, R., Butson, C., Leblanc, S., Litafovic, 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. *Can. J. Remote. Sens.* 29 <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., Weiss, M., Widlowski, J.L., Schaepman-Strub, G., Roman, M., Nickeson, J., 2014. Global leaf area index product validation good practices. In: Schaepman-Strub, G., Plummer, S., Nightingale, J. (Eds.), Best Practice for Satellite-Derived Land Product Validation. Land Product Validation Subgroup (Committee on Earth Observation Satellites Working Group on Calibration and Validation) (2014). <https://doi.org/10.5067/doc.ceoswgcv/lpv/lai.002>.
- Fernandes, R., Maloley, M., Canisius, F., 2018. Relationship between leaf area index and Landsat operational land imager equivalent reduced simple ratio vegetation index for the Athabasca oil sands region, northern Alberta. In: Geomatics Canada, Open File, 39. <https://doi.org/10.4095/308333>.
- Fernandes, R., et al., 2021. LEAF Toolbox. Canada Centre for Remote Sensing accessed at <https://github.com/rfernand387/LEAF-Toolbox/wiki>. on March 15, 2023, 202110.5281/zenodo.4321298.
- Fernandes, R., Brown, L., Canisius, F., Dash, J., He, L., Hong, G., Huang, L., Le, N.Q., MacDougall, C., Meier, C., Darko, P.O., Shah, H., Spafford, L., Sun, L., 2023. Validation of simplified level 2 prototype processor Sentinel-2 fraction of canopy cover, fraction of absorbed photosynthetically active radiation and leaf area index products over North American forests. *Remote Sens. Environ.* 293 <https://doi.org/10.1016/j.rse.2023.113600>.
- Ganguly, S., Nemani, R.R., Zhang, G., Hashimoto, H., Milesi, C., Michaelis, A., Wang, W., Votava, P., Samanta, A., Melton, F., Dungan, J.L., Vermote, E., Gao, F., Knyazikhin, Y., Myinen, R.B., 2012. Generating global leaf area index from Landsat: algorithm formulation and demonstration. *Remote Sens. Environ.* 122 <https://doi.org/10.1016/j.rse.2011.10.032>.
- Gonsamo, A., Chen, J.M., 2014. Improved LAI algorithm implementation to MODIS data by incorporating background, topography, and foliage clumping information. *IEEE Trans. Geosci. Remote* 52. <https://doi.org/10.1109/TGRS.2013.2247405>.
- Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, G., Schmuck, G., 1994. Leaf Optical Properties EXperiment 93 (LOPEX93) European Commission — Joint Research Centre, EUR 16095 EN, 20 pp.
- Hu, Q., Yang, J., Xu, B., Huang, J., Memon, Y., Yin, M.S., Zeng, G., Zhao, Y., Liu, J., 2020. Evaluation of global decametric-resolution LAI, FAPAR and FVC estimates derived from Sentinel-2 imagery. *Remote Sens.* 12 <https://doi.org/10.3390/rs12060912>.
- Huang, J., Zeng, Y., Kuusk, A., Wu, B., Dong, L., Mao, K., Chen, J.J., 2011. Inverting a forest canopy reflectance model to retrieve the overstorey and understorey leaf area index for forest stands. *International J. Remote Sen.* 32 (22), 7591–7611. <https://doi.org/10.1080/01431161.2010.525259>.
- Ishii, H., Hamada, Y., Utsugi, H., 2012. Variation in light-intercepting area and photosynthetic rate of sun and shade shoots of two *Picea* species in relation to the angle of incoming light. *Tree Physiol.* 32 <https://doi.org/10.1093/treephys/tps090>.
- Kang, Y., Ozdogan, M., Gao, F., Anderson, M.C., White, W.A., Yang, Yun, Yang, Yang, Erickson, T.A., 2021. A data-driven approach to estimate leaf area index for Landsat images over the contiguous US. *Remote Sens. Environ.* 258 <https://doi.org/10.1016/j.rse.2021.112383>.
- Kao, R.H., Gibson, C.M., Gallery, R.E., Meier, C.L., Barnett, D.T., Docherty, K.M., Blevins, K.K., Travers, P.D., Azuaje, E., Springer, Y.P., Thibault, K.M., McKenzie, V. J., Keller, M., Alves, L.F., Hinckley, E.-L.S., Parnell, J., Schimel, D., 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. *Ecosphere* 3. <https://doi.org/10.1890/ES12-00196.1>.
- Kattge, J., Boenisch, G., Diaz, S., et al., 2020. TRY plant trait database - enhanced coverage and open access. *Glob. Chang. Biol.* 26 <https://doi.org/10.1111/gcb.14904>.
- Lambin, E.F., Geist, H.J., 2006. In: Lambin, E.F., Geist, H.J. (Eds.), *Land-Use and Land-Cover Change. Local Processes and Global Impacts*, The IGBP Series. Springer-Verlag, Berlin, p. 222.
- Laura Martínez-Ferrer, L., Moreno-Martínez, A., Campos-Taberner, M., García-Haro, J., Muñoz-Marí, J., Running, S.W., Kimball, J., Clinton, N., Camps-Valls, G., 2022. Quantifying uncertainty in high resolution biophysical variable retrieval with machine learning. *Remote Sens. Environ.* 280 <https://doi.org/10.1016/j.rse.2022.113199>.
- Lewis, P., Disney, M., Quaife, T., Knyazikhin, Y., Schull, N.M., 2007. Modelling canopy reflectance with spectral invariants. In: 10th International Symposium on Physical Measurements and Signatures in Remote Sensing, Davos, CHF. In: <https://www.ispr.s.org/proceedings/xxxvi/7-c50/papers/T14.pdf>. accessed on August 21, 2023.
- 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 Sens.* 7 <https://doi.org/10.3390/rs1115494>.
- Liu, W., Baret, F., Gu, X., Tong, Q., Zheng, L., Zhang, B., 2002. Relating soil surface moisture to reflectance. *Remote Sens. Environ.* 81 [https://doi.org/10.1016/S0034-4257\(01\)00347-9](https://doi.org/10.1016/S0034-4257(01)00347-9).
- Miller, J.R., Peter White, H.P., Chen, J., Peddle, D., McDermid, G., Fournier, R.A., Shepherd, P., Rubinstein, I., Freemantle, J., Soffer, R., LeDrew, E., 1997. Seasonal change in understory reflectance of boreal forests and influence on canopy vegetation indices. *J. Geophys. Res.* 102 (D24) <https://doi.org/10.1029/97JD02558>.
- Möttö, M., Stenberg, P., 2008. A simple parameterization of canopy reflectance using photon recollision probability. *Remote Sens. Environ.* 112 <https://doi.org/10.1016/j.rse.2007.08.0021>.
- Müller-Wilm, U., 2018. *Sen2Cor Configuration and User Manual*, 2nd ed. CS, Toulouse, France.
- Myinen, R.B., Williams, D., 1994. On the relationship between FAPAR and NDVI. *Remote Sens. Environ.* 49, 200–211.
- Myinen, R.B., Williams, D.L., 1994. On the relationship between FAPAR and NDVI. *Remote Sens. Environ.* 49 [https://doi.org/10.1016/0034-4257\(94\)90016-7](https://doi.org/10.1016/0034-4257(94)90016-7).
- Nilson, T., 1999. Inversion of gap frequency in forest stands. *Agric. For. Meteorol.* 98–99 [https://doi.org/10.1016/S0168-1923\(99\)00114-8](https://doi.org/10.1016/S0168-1923(99)00114-8).
- North, P.R.J., 1996. Three-dimensional forest light interaction model using a Monte Carlo method. *IEEE Trans. Geosci. Remote Sens.* 34 <https://doi.org/10.1109/36.508411>.
- Pallarés, P.M., Rivera Caicedo, J.P., Belda, S., De Gravé, C., Burriel, H., Moreno, J., Verrelst, J., 2019. Quantifying the robustness of vegetation indices through global sensitivity analysis of homogeneous and forest leaf-canopy radiative transfer models. *Remote Sens.* 11 <https://doi.org/10.3390/rs1120241>.
- Pisek, J., Diaz-Pines, E., Matteucci, G., Noe, S., Rebmann, C., 2022. On the leaf inclination angle distribution as a plant trait for the most abundant broadleaf tree species in Europe. *Agric. For. Meteorol.* 323 <https://doi.org/10.1016/j.agrformet.2022.109030>.
- Pope, G., Treitz, P., 2013. Leaf area index (LAI) estimation in boreal mixedwood forest of Ontario, Canada using light detection and ranging (LiDAR) and WorldView-2 imagery. *Remote Sens.* 5 <https://doi.org/10.3390/rs5105040>.
- Purves, D.W., Lichstein, J.W., Pacala, S.W., 2007. Crown plasticity and competition for canopy space: a new spatially implicit model parameterized for 250 North American tree species. *PLoS One* 2 (9), e870. <https://doi.org/10.1371/journal.pone.0000870>.
- Qi, Y., Li, F., Liu, Z., Jin, G., 2014. Impact of understory on overstorey leaf area index estimation from optical remote sensing in five forest types in northeastern China. *Agric. For. Meteorol.* 198–199, 72–80.
- Rautainen, M., Möttö, M., Yáñez Rausell, L., Homolová, L., Malenovský, Z., Schaepman, M., 2012. A note on upscaling coniferous needle spectra to shoot spectral albedo. *Remote Sens. Environ.* 117, 469–474.
- Rochdi, N., Fernandes, R.A., Chelle, M., 2006. An assessment of needles clumping within shoots when modeling radiative transfer within homogeneous canopies. *Remote Sens. Environ.* 102 <https://doi.org/10.1016/j.rse.2006.02.003>.
- Roy, D.P., Li, J., Zhang, H.K.K., Yan, L., Huang, H.Y., Li, Z.B., 2017. Examination of Sentinel 2A multi-spectral instrument MSI reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* 199 <https://doi.org/10.1016/j.rse.2017.06.019>.
- Shabanov, N., Gastellu-Etchegorry, J.P., 2018. The stochastic Beer-Lambert-Bouguer law for discontinuous vegetation canopies. *J. Quant. Spectrosc. Radiat. Transf.* 214, 18–32.

- Shabanov, N., Gastellu-Etchegorry, J.-P., 2018. The stochastic Beer–Lambert–Bouguer law for discontinuous vegetation canopies. *J. Quant. Spectrosc. Radiat. Transf.* 214 <https://doi.org/10.1016/j.jqsrt.2018.04.021>.
- Shabanov, N.V., Huang, D., Yang, W., Tan, B., Knyazikhin, Y., Myneni, R.B., Ahl, D.E., Gower, S.T., Huete, A.R., Aragao, L.E.O.C., Shimabukuro, Y.E., 2005. Analysis and optimization of the MODIS leaf area index algorithm retrievals over broadleaf forests. *IEEE Trans. Geosci. Remote Sens.* 43 <https://doi.org/10.1109/TGRS.2005.852477>.
- Smolander, S., Stenberg, P., 2003. A method to account for shoot scale clumping in coniferous canopy reflectance models. *Remote Sens. Environ.* 88 <https://doi.org/10.1016/j.rse.2003.06.003>.
- Stenberg, P., 1996. Correction LAI-2000 estimates for the clumping of needles ion shoots of conifers. *Agric. For. Meteorol.* 79, 1–8.
- Stenberg, P., Manninen, T., 2015. The effect of clumping on canopy scattering and its directional properties: a model simulation using spectral invariants. *Int. J. Remote Sens.* 36 (19–20), 5178–5191. <https://doi.org/10.1080/01431161.2015.1049383>.
- Stenberg, P., Palmroth, S., Bond, B., Sprugel, D., Smolander, H., 2001. Shoot structure and photosynthetic efficiency along the light gradient in a Scots pine canopyTree. *Physiology* 21, 805–814.
- Verhoef, W., 1985. Earth observation modeling based on layer scattering matrices. *Remote Sens. Environ.* 17 [https://doi.org/10.1016/0034-4257\(85\)90072-0](https://doi.org/10.1016/0034-4257(85)90072-0).
- Verhoef, W., Bach, H., 2007. Coupled soil-leaf-canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sens. Environ.* 109, 166–182. <https://doi.org/10.1016/j.rse.2006.12.013>.
- Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G., Camps-Valls, G., Moreno, J., 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods—a comparison. *ISPRS J. Photogramm. Remote Sens.* 108, 260–272.
- Weinstein, B.G., Graves, S.J., Marconi, S., Singh, A., Zare, A., Stewart, D., et al., 2021. A benchmark dataset for canopy crown detection and delineation in co-registered airborne RGB, LiDAR and hyperspectral imagery from the national ecological observation network. *PLoS Comput. Biol.* 17 (7), e1009180 <https://doi.org/10.1371/journal.pcbi.1009180>.
- Weiss, M., Baret, F., 2016. S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER, 1.0. ed. Institut National de la Recherche Agronomique, Avignon, France. [https://step.esa.int/docs/extras/ATBD\\_S2ToolBox\\_L2B\\_V1.1.pdf](https://step.esa.int/docs/extras/ATBD_S2ToolBox_L2B_V1.1.pdf).
- Weiss, M., Baret, F., 2020. S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER, 2.0. ed. Institut National de la Recherche Agronomique, Avignon, France. [https://step.esa.int/docs/extras/ATBD\\_S2ToolBox\\_L2B\\_V2.0.pdf](https://step.esa.int/docs/extras/ATBD_S2ToolBox_L2B_V2.0.pdf).
- Widlowski, J.-L., et al., 2007. Third radiation transfer model Intercomparison (RAMI) exercise: documenting progress in canopy reflectance models. *J. Geophys. Res.* 112, D09111. <https://doi.org/10.1029/2006JD007821>.
- Widlowski, J.L., Mio, C., Disney, M., Adams, J., Andredakis, I., Atzberger, C., Brennan, J., Busetto, I., Chelle, M., Ceccherini, G., et al., 2015. The fourth phase of the radiative transfer model intercomparison (RAMI) exercise: actual canopy scenarios and conformity testing. *Remote Sens. Environ.* 169 <https://doi.org/10.1016/j.rse.2015.08.016>.
- WMO, 2022. The 2022 GCOS ECV Requirements, GCOS- No. 245, © World Meteorological Organization, 261pp.
- Yan, G., Hailan, J., Luo, J., Mu, X., Li, F., Qi, J., Hu, R., Xie, D., Zhou, G., 2021. Quantitative evaluation of leaf inclination angle distribution on leaf area index retrieval of coniferous canopies. *J. Remote Sens.* <https://doi.org/10.34133/2021/2708904>.
- Zarco-Tejada, P.J., Miller, J.R., Morales, A., Berjón, A., Agüera, J., 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sens. Environ.* 90 <https://doi.org/10.1016/j.rse.2004.01.017>.